

Mention "Physique" de la Licence de Sciences et Technologies L2

Parcours "Physique Fondamentale" (PF) $Ann\acute{e}~2013\text{-}2014$

Méthodes mathématiques pour physiciens 2 LP207

Jean-Bernard Zuber

Notes disponibles sur Sakai ou sur www.lpthe.jussieu.fr/~zuber/Z_Notes.html



Carl Friedrich Gauss

Carl Friedrich Gauss (1777–1855) domine la science mathématique et physique de son temps. Il a laissé des contributions fondamentales en arithmétique, algèbre, géométrie différentielle, théorie des probabilités, ainsi qu'en électricité, magnétisme, mécanique céleste, etc. Nous rencontrerons son nom à plusieurs reprises dans ce cours, du pivot de Gauss à la courbe gaussienne.

Table des matières

Partie I : Algèbre linéaire

Chapitre 1. Des vecteurs aux matrices		1
1. Rappels sur les vecteurs. Espaces vectoriels		1
1.1. Vecteurs de \mathbb{R}^2 et \mathbb{R}^3		1
1.2. Autres espaces vectoriels		2
2. Indépendance linéaire. Base d'un espace vectoriel		3
2.1. Dépendance et indépendance linéaire $\dots \dots \dots \dots \dots$		3
2.2. Base		4
2.3. Rang d'un système de vecteurs $\dots \dots \dots \dots \dots \dots$		7
2.4. Changement de base. Matrice d'un changement de base		8
3. Applications linéaires		10
3.1. Projection		10
3.2. Application linéaire		11
3.3. Matrice d'une application linéaire		12
4. Matrices		14
4.1. Produit matriciel		14
4.2. Addition et multiplication par un scalaire des matrices		17
4.3. Changement de base pour la matrice d'une application		17
4.4. Autres définitions		18
4.5. Matrices-lignes, matrices-colonnes		20
4.6. Rang d'une matrice		21
5. Vecteurs, tenseurs		22
5.1. Produit tensoriel		22
5.2. Formules de changement de base		23
5.3. Contraction de deux tenseurs. Convention d'Einstein		25
5.4. Repères orthonormés		25
5.5. Exemples physiques. Tenseur d'inertie. Tenseur de conductivité.	Tenseurs of	les
déformations et des contraintes. Tenseur piézoélectrique		26
Chapitre 2. Déterminants		20
1. Rappels sur les permutations de p objets		
1. Trappers our res permutations de ρ objets		49

$2.\ {\rm Fonctions}$ multilinéaires. Fonctions antisymétriques. Fonction déterminant			31
3. Propriétés du déterminant			32
4. Méthodes de calcul			35
4.1. Calcul direct			35
4.2. Combinaisons linéaires des colonnes ou des lignes			36
4.3. Développement par rapport à une ligne ou à une colonne. Mineurs $$.			37
4.4. Méthode du pivot de Gauss $\dots \dots \dots \dots \dots \dots \dots \dots$			38
4.5. Calcul de l'inverse d'une matrice			39
5. Applications des déterminants			40
5.1. Critère d'indépendance linéaire			40
5.2. Équation d'une droite de \mathbb{R}^2 , d'un plan de \mathbb{R}^3			41
5.3. Wronskien			42
5.4. Interprétation géométrique du déterminant. Jacobien			42
Chapitre 3. Systèmes linéaires d'équations algébriques			43
1. Position du problème			
1.1. Système d'équations considéré comme un problème vectoriel			43
1.2. Systèmes d'équations homogènes			44
1.3. Interprétation matricielle			
2. Rang d'un système, déterminant principal			
3. Discussion et résolution. Systèmes de Cramer			
3.1. $p = r \le n$			
3.2. $r < p$. Déterminants caractéristiques			
3.3. Système homogène			
4. Un exemple détaillé			
5. Applications mécaniques. Équilibre statique de solides indéformables			52
6. Applications électriques. Circuits et lois de Kirchhoff			53
Chapitre 4. Valeurs propres, vecteurs propres. Diagonalisation			55
1. Vecteurs et valeurs propres			55
1.1. Définitions de base			
1.2. Valeurs propres d'une matrice singulière			
1.3. Sous-espace propre			
1.4. Polynôme caractéristique			
2 Diagonalisation d'une matrice			60

	Table des	mati	ères	iii
2.1. Détermination des vecteurs propres				. 60
2.2. Diagonalisation. Changement de base				. 60
2.3. Exemples				. 61
2.4. "Triangularisation" d'une matrice. Théorème de Cayley-Ham	ilton .			. 62
3. Conséquences et applications de la diagonalisation				. 63
3.1. Matrices commutantes				. 63
3.2. Puissances et exponentielle d'une matrice				
4. Applications aux systèmes linéaires d'équations différentielles. Osc				
4.1. Systèmes de 2 équations différentielles linéaires				
4.2. Systèmes de n équations				
4.3. Oscillateurs couplés				
Chapitre 5. Matrices symétriques et formes quadratiques .				. 73
1. Formes bilinéaires, formes quadratiques				. 73
1.1. Formes bilinéaires et quadratiques				. 73
1.2. Formes définies positives				. 74
1.3. Représentations matricielles				. 75
2. Réduction d'une forme quadratique				. 76
2.1. Vecteurs orthogonaux, vecteurs orthonormés				. 76
2.2. Procédé d'orthonormalisation de Schmidt				. 76
2.3. Matrices orthogonales				. 77
2.4. Diagonalisation d'une matrice symétrique				. 79
2.5. Réduction d'une forme quadratique				
2.6. Diagonalisation simultanée de deux matrices symétriques com	mutantes	3.		. 81
3. Extension aux formes sesquilinéaires et matrices hermitiennes .				. 81
4. Applications physiques				. 82
4.1. Tenseur d'inertie				. 82
4.2. Tenseur de conductivité				. 84
4.3. Stabilité des extrema d'un potentiel				
Partie II : Probabilités				
Chapitre 6. Événements et probabilités			•	. 87

2. Propriétés et axiomes	91
2.1. Événements composés	91
2.2. Espace d'épreuves fini	92
2.3. Exemples	93
2.4. Espace probabilisé	93
2.5. Probabilités conditionnelles	93
3. Un peu de combinatoire	. 94]
Chapitre 7. Variables aléatoires	97
1. Variables aléatoires. Distributions de v.a	98
1.1. Définition d'une variable aléatoire	98
1.2. Les fonctions importantes attachées à une v.a	99
1.3. Plusieurs variables aléatoires	. 101
2. Moyenne, variance, écart-type, moments	. 101
2.1. Moyenne	. 101
2.2. Variance et écart-type	. 102
2.3. Moments	. 103
3. Fonctions de corrélation de plusieurs v.a	. 104
4. Changement de variable aléatoire	. 105
Chapitre 8. Distributions classiques	. 107
1. Distribution uniforme	. 107
2. Distribution binomiale	. 108
3. Distribution normale $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$. 110
4. Distribution de Poisson	. 113
5. Limites de la loi binomiale $\dots \dots \dots \dots \dots \dots \dots \dots$. 114
6. Quelques exemples concrets	. 116
6.1. L'aiguille de Buffon	. 116
6.2. Distribution de Maxwell des vitesses dans un gaz	. 117
6.3. Désintégrations radioactives	. 118
Chapitre 9. Théorèmes asymptotiques	. 120
1. Préambule	. 120
1.1. Convergence en probabilité $\dots \dots \dots \dots \dots \dots \dots$. 120
1.2. Moyenne arithmétique de N v.a	. 120
2 Loi dos grande nombros	199

	2.1.	Le théorème et sa preuve .												122
	2.2.	Deux illustrations												123
3.	Thé	orème limite central												124
	3.1.	Énoncé et remarques												125
	3.2.	Éléments de preuve du théore	ème											125
	3.3.	Illustrations												126
	3.4.	Évaluations de l'erreur												127
4.	Mar	che aléatoire												128
	4.1.	Marche au hasard sur un rése	eau											131
	4.2.	Processus stochastiques, prop	oriété	de l	Maı	kov	I							132
	4.3.	Limite continue et équation o	de la c	chal	eur									132
	4.4.	Lois d'échelle dans la marche	au ha	asar	rd									135

Partie I: Algèbre linéaire

Chapitre 1. Des vecteurs aux matrices

Ce chapitre est consacré à des rappels sur les vecteurs et les concepts clés d'indépendance linéaire et de bases.

La question du changement de base va amener tout naturellement l'introduction des matrices et leurs propriétés de multiplication.

1. Rappels sur les vecteurs. Espaces vectoriels

1.1. Vecteurs de \mathbb{R}^2 et \mathbb{R}^3

La géométrie du plan et celle de l'espace nous ont familiarisés avec la notion de vecteur. Étant donné le plan noté \mathbb{R}^2 ou l'espace tridimensionnel noté \mathbb{R}^3 , on définit le vecteur \overrightarrow{OM} : c'est le segment orienté reliant l'origine O au point M. Il est donc caractérisé par sa longueur ou module $|\overrightarrow{OM}| = \text{longueur ordinaire du segment } OM$, et sa direction orientée¹. Le vecteur $\overrightarrow{V} = \overrightarrow{OM}$ est alors identifié à toutes ses copies d'origine arbitraire : $\overrightarrow{OM} = \overrightarrow{O'M'}$ si la figure OMM'O' est un parallélogramme, voir figure 1(a).

Les deux opérations suivantes peuvent s'effectuer sur les vecteurs

- multiplication par un nombre réel (un "scalaire") λ : si $\vec{V} = \overrightarrow{OM}$, le vecteur $\lambda \vec{V} = \overrightarrow{OM'}$ a la même direction que \vec{V} , et une orientation identique ou opposée selon le signe de λ , mais une longueur multipliée par $|\lambda|$: $|\overrightarrow{OM'}| = |\lambda| |\overrightarrow{OM}|$ (bien noter la valeur absolue de λ !). On dit que les vecteurs $\lambda \vec{V} = \overrightarrow{OM'}$ et $\vec{V} = \overrightarrow{OM}$ sont colinéaires;
- somme de deux vecteurs de même origine : c'est la fameuse règle du parallélogramme, voir figure 1(b).

¹ Plutôt que la flèche, on utilise souvent dans les livres des caractères gras pour représenter un vecteur **OM** : pas très faciles à faire à la main ou au tableau noir...

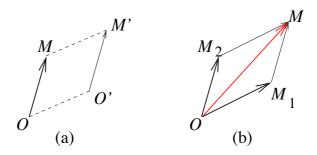


Fig. 1: (a): Équivalence des deux vecteurs \overrightarrow{OM} et $\overrightarrow{O'M'}$. (b): Addition de deux vecteurs $\overrightarrow{OM_1}$ et $\overrightarrow{OM_2}$.

Les règles de calcul combinées de ces deux opérations sont familières

$$(\lambda + \mu)\vec{V} = \lambda \vec{V} + \mu \vec{V}$$

$$\lambda(\vec{V}_1 + \vec{V}_2) = \lambda \vec{V}_1 + \lambda \vec{V}_2$$
(1.1)

(c'est la "distributivité" de la multiplication par un scalaire par rapport à l'addition).

On peut aussi définir le vecteur nul, $\vec{0} = \overrightarrow{OO}$, dont l'addition à tout \overrightarrow{OM} ne le modifie pas, et l'opposé $-\overrightarrow{OM}$ tel que $\overrightarrow{OM} + (-\overrightarrow{OM}) = \vec{0}$. Le vecteur $\vec{0}$ résulte de la multiplication par le scalaire 0 de tout vecteur \overrightarrow{OM} , et $-\overrightarrow{OM}$ de la multiplication par le scalaire -1 du vecteur \overrightarrow{OM} , en cohérence avec les règles (1.1).

L'addition de deux vecteurs quelconques est un autre vecteur, la multiplication d'un vecteur par un scalaire en est un aussi : on dit que l'ensemble des vecteurs est "stable" sous l'effet de ces deux opérations d'addition et de multiplication par un scalaire. Un tel ensemble est par définition un espace vectoriel.

En général, dans un tel espace, on peut construire toutes les *combinaisons linéaires* de n vecteurs $\vec{V}_1, \dots, \vec{V}_n$,

$$\forall \lambda_i \in \mathbb{R}$$
 $\lambda_1 \vec{V}_1 + \lambda_2 \vec{V}_2 + \dots + \lambda_n \vec{V}_n$ est un vecteur.

La notion de vecteur et toutes ces opérations trouvent bien sûr leur application en géométrie (translations etc), en mécanique (déplacements, vitesses, accélérations, forces, moment cinétique, etc), mais aussi en électromagnétisme (champs et moments électriques et magnétiques...), en optique (vecteurs d'onde), etc.

1.2. Autres espaces vectoriels

De nombreux concepts d'origine géométrique impliquant des vecteurs peuvent s'étendre à des situations plus générales impliquant d'autres objets mathématiques dotés de propriétés de même nature. Par exemple, on rencontrera les espaces vectoriels des

- fonctions réelles d'une variable $x \in [a, b]$ (intervalle qui peut être infini), par exemple continues, ou dérivables ;
- polynômes d'une variable x;
- fonctions périodiques de période T;
- solutions d'une équation différentielle linéaire homogène ("sans second membre"), etc.

Pour chacun de ces objets, les notions d'addition et de multiplication par un scalaire réel² vérifiant les règles de calcul (1.1) sont naturelles et ne "font pas sortir de l'ensemble". Chacun de ces ensembles de fonctions est donc un espace vectoriel, et les raisonnements que nous allons développer dans la suite s'y appliquent.

Ces extensions jouent aussi un rôle dans de nombreuses situations de physique : l'étude des systèmes dynamiques linéaires qu'on rencontre en Mécanique ou en Électricité conduit naturellement à des systèmes d'équations différentielles linéaires dont les solutions forment un espace vectoriel. En Mécanique Quantique, la description des états d'un système physique se fait en termes de vecteurs abstraits, dont une réalisation est celle des "fonctions d'onde", fonctions à valeurs complexes des variables de position et de temps, dont l'interprétation est intimement liée au caractère probabiliste de la physique quantique.

2. Indépendance linéaire. Base d'un espace vectoriel

2.1. Dépendance et indépendance linéaire

Nous continuerons d'utiliser la notation \vec{V} pour désigner un vecteur même si c'est dans un contexte plus général comme ceux discutés au § 1.2. Quand on discute de vecteurs, deux notions **fondamentales** sont celles de dépendance et d'indépendance linéaire

Définition : Les p vecteurs $\vec{V}_1, \dots \vec{V}_p$ sont linéairement dépendants (ou "forment un système lié") s'il existe un ensemble de p scalaires (nombres réels) $\lambda_1, \dots, \lambda_p$ non tous nuls tels que

$$\lambda_1 \vec{V}_1 + \dots + \lambda_p \vec{V}_p = \vec{0} \tag{2.1}$$

ce qu'on abrégera désormais par la notation $\sum_{i=1}^p \lambda_i \vec{V}_i = 0$.

² Dans toutes ces notes, on considérera sauf exception le cas d'espaces vectoriels réels, où les nombres "scalaires" considérés sont réels. Il n'y a pas de difficulté majeure à étendre la discussion au cas complexe. Certaines situations rencontrées en physique, comme l'étude des systèmes oscillants, ou la mécanique quantique, peuvent y conduire. Voir l'exercice de TD sur les matrices d'impédance ou d'admittance de quadrupôles.

À l'inverse,

Définition : Les n vecteurs $\vec{V}_1, \cdots \vec{V}_p$ sont linéairement indépendants (ou "forment un système libre") s'il n'existe pas d'ensemble de p scalaires (nombres réels) $\lambda_1, \cdots, \lambda_p$ non tous nuls tels que $\sum_{i=1}^p \lambda_i \vec{V}_i = 0$, autrement dit si

$$\sum_{i=1}^{p} \lambda_i \vec{V}_i = 0 \quad \Longrightarrow \quad \lambda_i = 0 \quad \forall i = 1, \dots, n \quad . \tag{2.2}$$

Exemples

- dans l'espace \mathbb{R}^2 ou \mathbb{R}^3 , (ou plus généralement \mathbb{R}^d , $d \geq 2$) les vecteurs \vec{V}_1 , \vec{V}_2 non colinéaires sont linéairement indépendants (preuve par l'absurde : si $\lambda_1 \vec{V}_1 + \lambda_2 \vec{V}_2 = 0$ avec par exemple $\lambda_2 \neq 0$, on peut écrire $\vec{V}_2 = -\lambda_1 \lambda_2^{-1} \vec{V}_1$ ce qui montre bien que \vec{V}_1 et \vec{V}_2 sont colinéaires, contrairement à l'hypothèse) ; les vecteurs \vec{V}_1 , \vec{V}_2 et $\vec{V}_3 = \lambda_1 \vec{V}_1 + \lambda_2 \vec{V}_2$ sont linéairement dépendants. En effet on peut écrire $\lambda_1 \vec{V}_1 + \lambda_2 \vec{V}_2 \vec{V}_3 = 0$. Trois vecteurs non nuls \vec{V}_1 , \vec{V}_2 et \vec{V}_3 sont linéairement indépendants si et seulement s'ils ne sont pas coplanaires ;
- interprétation géométrique des combinaisons linéaires $\lambda \vec{V}_1 + \mu \vec{V}_2$ dans \mathbb{R}^3 comme ensemble des vecteurs du plan contenant \vec{V}_1 et \vec{V}_2 ;
- les polynômes 1, x et x^2 sont linéairement indépendants ; les polynômes 1, x, ax + b ne sont pas indépendants, quels que soient $a, b \in \mathbb{R}$;
- Dans l'espace vectoriel des fonctions périodiques de période 2π , discuter l'indépendance linéaire des fonctions 1, $\cos x$, $\cos 2x$; ou encore de 1, $\cos^2 x$, $\cos 2x$;
- Dans l'espace vectoriel des solutions de l'équation différentielle $\ddot{x} + x = 0$, les solutions e^{ix} , e^{-ix} , $\cos x$ ne sont pas indépendantes, pourquoi ?

2.2. Base

Définition : Dans un espace vectoriel, on dit qu'un ensemble de n vecteurs $\vec{e_1}$, $\vec{e_2}$, \cdots , $\vec{e_n}$ est une base si ces vecteurs sont linéairement indépendants et si tout autre vecteur \vec{X} de l'espace peut être écrit comme combinaison linéaire des $\vec{e_i}$

$$\forall \vec{X} \quad \exists x_1, x_2, \cdots, x_n \quad \vec{X} = \sum x_i \vec{e_i} .$$

On appelle alors les x_i composantes du vecteur \vec{X} dans la base $\{\vec{e}_i\}$.

Noter que pour un vecteur \vec{X} donné, ces composantes sont uniques, comme conséquence de l'indépendance linéaire des \vec{e}_i . En effet s'il existait deux ensembles de

composantes x_i et x_i' telles que $\vec{X} = \sum_{i=1}^n x_i \vec{e}_i$ et $\vec{X} = \sum_{i=1}^n x_i' \vec{e}_i$, en soustrayant membre à membre on aurait

$$\sum (x_i - x_i')\vec{e}_i = 0 \quad \stackrel{(2.2)}{\Longrightarrow} \quad \forall i, \ x_i - x_i' = 0, \ \text{donc} \ x_i = x_i' \ ,$$

ce qui montre l'unicité des x_i .

Exemples.

- Une base dans l'espace \mathbb{R}^d est encore appelée $rep\`ere$: elle permet de repérer tout point M de l'espace (ou tout vecteur \overrightarrow{OM}) par ses coordonnées (ou composantes) dans ledit repère. Voir Fig. 2.
- Les monômes 1, x, x^2, \dots, x^n forment une base de l'ensemble des polynômes de degré inférieur ou égal à n

$$P(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$$

avec les composantes a_i identifiées aux coefficients du polynôme.

– Montrer que $\{e^{ix}, e^{-ix}\}$ et $\{\cos x, \sin x\}$ sont deux bases de solutions de l'équation différentielle $\ddot{x} + x = 0$.

Le choix de base n'est en effet jamais unique. Par exemple, si $\vec{e}_1, \ \vec{e}_2, \cdots, \vec{e}_n$ forment une base, il en est de même de $\lambda_1 \vec{e}_1, \ \lambda_2 \vec{e}_2, \cdots, \lambda_n \vec{e}_n$, pour tous $\lambda_i \neq 0$; ou de $\vec{e}_1 + \vec{e}_2, \vec{e}_1 - \vec{e}_2, \vec{e}_3, \cdots, \vec{e}_n$, etc. Nous allons bientôt disposer d'un critère disant quels changements de base sont possibles...

Théorème 1 : Si on connaît une base de n vecteurs, tout système de p vecteurs, avec p > n, est nécessairement linéairement dépendant.

Preuve par l'absurde. Soit $\vec{e}_i,\ i=1,\cdots,n$ une base, soit $\vec{X}_j,\ j=1,\cdots,p$ un système de p>n vecteurs que nous supposons indépendants. Montrons alors que les $\vec{X}_i,\ 1\leq i\leq n$ forment une autre base, et que les $\vec{X}_j,\ n+1\leq j\leq p$ peuvent donc s'exprimer en termes des $\vec{X}_i,\ 1\leq i\leq n$, en contradiction avec l'hypothèse d'indépendance. Les \vec{e} formant une base, \vec{X}_1 peut s'écrire $\vec{X}_1=\sum_{1\leq j\leq n}x_{1j}\vec{e}_j$ avec des x_{1j} non tous nuls (sans quoi \vec{X}_1 serait nul et le système des vecteurs \vec{X} pas indépendants). Supposons par exemple que $x_{11}\neq 0$. Cela permet d'écrire \vec{e}_1 comme combinaison linéaire de \vec{X}_1 et des $\vec{e}_j,\ 2\leq j\leq n$. Les vecteurs $\vec{X}_1,\vec{e}_j,\ 2\leq j\leq n$ forment donc une nouvelle base. Itérons alors le raisonnement : \vec{X}_2 peut s'exprimer comme combinaison de ces vecteurs: $\vec{X}_2=x_{21}\vec{X}_1+\sum_{j=2}^nx_{2j}\vec{e}_j$ avec des coefficients x_{2j} non tous nuls sans quoi \vec{X}_1 et \vec{X}_2 seraient linéairement dépendants, contrairement à l'hypothèse. Supposons par exemple $x_{22}\neq 0$, ce qui permet d'exprimer \vec{e}_2 en termes de $\vec{X}_1,\vec{X}_2,\vec{e}_3,\cdots,\vec{e}_n$, qui constituent donc une nouvelle base. Etc. Après n itérations de ce raisonnement, on arrive à la conclusion que $\vec{X}_1,\cdots,\vec{X}_n$ forment une base, et donc que les vecteurs $\vec{X}_{n+1},\cdots,\vec{X}_p$ peuvent s'exprimer comme combinaisons linéaires des n premiers, contrairement à l'hypothèse. Le théorème est démontré.

Corollaire 1 : Le nombre d'éléments d'une base est indépendant du choix de cette base.

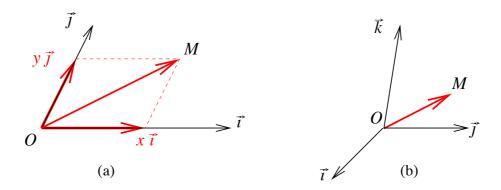


Fig. 2: (a) : repère dans l'espace \mathbb{R}^2 ; (b) : repère dans l'espace \mathbb{R}^3 .

En effet supposons que nous ayons construit deux bases distinctes $\vec{e_i}$, $i=1,\dots n$ et $\vec{f_i}$, $i=1,\dots n'$. D'après le Théorème précédent, $n' \leq n$ (indépendance des $\vec{f_i}$) et $n \leq n'$ (indépendance des $\vec{e_i}$). Donc n=n'.

Définition : Ce nombre d'éléments de toute base est appelé dimension de l'espace (vectoriel).

Exemples.

- a) En géométrie élémentaire du plan, on définit des $rep\`eres$ de l'espace \mathbb{R}^2 , constitué de 2 vecteurs notés \vec{i} , \vec{j} , tels que tout vecteur \overrightarrow{OM} s'exprime comme $\overrightarrow{OM} = x\vec{i} + y\vec{j}$, et les nombres x et y sont les composantes du vecteur \overrightarrow{OM} ou encore les coordonn'ees du point M dans le repère \vec{i} , \vec{j} , voir Fig. 2. Dans l'espace \mathbb{R}^3 , cette construction s'étend en introduisant un 3ème vecteur \vec{k} dans le repère, et donc une troisième composante z pour tout vecteur \vec{X} . De façon générale, \mathbb{R}^n est un espace de dimension n et vice versa, tout espace de dimension n peut être considéré comme identifiable ("isomorphe") à \mathbb{R}^n .
- b) Dans l'espace des polynômes de degré strictement inférieur à n, les n polynômes $1, x, x^2, \dots x^{n-1}$ constituent une base. La dimension de cet espace est donc égale à n.
- c) Dans l'espace vectoriel des fonctions périodiques de période 2π continues, les fonctions 1 et $\cos px$, $\sin px$ pour tout p entier positif constituent une base, selon les théorèmes sur les développements de Fourier (cf cours LP206). Cet espace est de dimension infinie!

On se restreindra dans toute la suite de ce cours à des situations où la dimension de l'espace est finie.

2.3. Rang d'un système de vecteurs

Considérons p vecteurs $\vec{X}_1, \dots, \vec{X}_p$ d'un espace vectoriel E de dimension n. Ces p vecteurs engendrent un espace vectoriel F qui est l'ensemble de toutes les combinaisons linéaires de

la forme $\lambda_1 \vec{X}_1 + \dots + \lambda_p \vec{X}_p$. (Exercice: vérifier que cet ensemble F est en effet un espace vectoriel, toutes les combinaisons linéaires de vecteurs de cette forme étant elles-mêmes de cette forme.) On dit que F est un sous-espace vectoriel de E.

Quelle est la dimension r de cet espace F? Elle est certainement inférieure ou égale à celle (notée n) de l'espace E, par le Théorème 1. Montrons qu'elle est aussi inférieure ou égale au nombre p des vecteurs \vec{X}_i . Si ces p vecteurs \vec{X}_i sont linéairement indépendants, ils constituent une base de l'espace F (par définition d'une base), et la dimension de F est donc p. Si les p vecteurs \vec{X}_i sont linéairement dépendants, la dimension de F est strictement inférieure à p. En général

Définition : On appelle $rang\ r$ d'un système de p vecteurs la dimension de l'espace F qu'ils engendrent.

Nous venons de démontrer les deux inégalités (la notation $\#\{\cdot\}$ signifie "nombre de")

$$\dim F = r \le p = \#\{\vec{X}_i\}$$

$$\le n = \dim E .$$
(2.3)

▷ Compléter un système de vecteurs en une base

Soit un système de vecteurs donnés \vec{V}_i de rang r dans un espace E de dimension n: cela signifie qu'on peut extraire du système r vecteurs linéairement indépendants, supposons que ce sont les r premiers $\vec{V}_1, \cdots, \vec{V}_r$. Si r=n, ils constituent une base de E. Si r< n, il existe un vecteur X_1 de E qui n'appartient pas au sous-espace F engendré par les \vec{V} . On peut l'ajouter à $\vec{V}_1, \cdots, \vec{V}_r$ pour obtenir un système de r+1 vecteurs indépendants. On itère le raisonnement : si r+1=n, on a une base de E, sinon, etc. Au final, on a construit une base de E en complétant le système $\vec{V}_1, \cdots, \vec{V}_r$ par n-r vecteurs linéairement indépendants dans le "sous-espace supplémentaire" de F dans E. Ce principe de construction est utilisé très fréquemment.

Exemple. Considérons l'espace vectoriel E des polynômes en x de degré inférieur ou égal à 3. Sa dimension est 4. Le système de "vecteurs" $1 + x^2$ et $1 + x^3$ est de rang 2. On peut le compléter en une base de E en lui ajoutant par exemple 1 et x.

2.4. Changement de base. Matrice d'un changement de base

Comme on l'a dit, le choix de base n'est pas unique. Considérons donc un espace vectoriel E où on a construit deux bases distinctes $\vec{e_i}$ et $\vec{f_i}$, $i=1,\dots n$. Les $\vec{e_i}$ formant une base, on peut exprimer tout vecteur, et en particulier les $\vec{f_j}$ comme combinaisons linéaires des $\vec{e_i}$

$$\vec{f}_j = \sum_{i=1}^n \vec{e}_i a_{ij} \qquad j = 1, \dots, n$$
 (2.4)

avec un ensemble de nombres a_{ij} . On représente un tel ensemble comme un tableau carré A à n lignes et n colonnes appelé matrice. L'indice de gauche i est l'indice de ligne, prenant des valeurs de 1 à n; celui de droite j est celui de colonne, et l'élément a_{ij} est ce qu'on lit à la croisée de la i-ème ligne et de la j-ième colonne, et on écrit aussi $(A)_{ij} = a_{ij}$. Noter que cet élément a_{ij} est par définition la i-ième composante (dans la base \vec{e}) de \vec{f}_j . Donc

• La j-ième colonne de la matrice est faite des composantes du vecteur \vec{f}_j dans la base \vec{e} . Par exemple, si la dimension est n=3

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} .$$

Soit maintenant \vec{X} un vecteur quelconque de l'espace vectoriel E précédent. Dans les deux bases $\{\vec{e_i}\}$ et $\{\vec{f_i}\}$, le vecteur \vec{X} a deux décompositions

$$\vec{X} = \sum_{i} x_i \vec{e_i}$$

$$= \sum_{j} y_j \vec{f_j} = \sum_{i,j} y_j \vec{e_i} a_{ij} ,$$

où on a utilisé la formule (2.4). En identifiant les composantes sur $\vec{e_i}$ de ces deux expressions, on trouve

$$x_i = \sum_j a_{ij} y_j , \qquad (2.5)$$

formule qui exprime la transformation des composantes d'une base à l'autre. Bien observer les différences entre (2.4) et (2.5): dans (2.4), on a exprimé les nouveaux vecteurs de base comme combinaisons linéaires des anciens, tandis que dans (2.5), on exprime les anciennes composantes en termes des nouvelles, et la sommation sur les indices diffère dans les deux formules: on somme sur l'indice i dans la première, sur l'indice j dans la seconde.

Exemple. Dans l'espace \mathbb{R}^2 , considérons le changement de base $(\vec{e}_1,\ \vec{e}_2) \to (\vec{f}_1 =$ $\vec{e}_1 + \vec{e}_2$, $\vec{f}_2 = 2\vec{e}_1 - \vec{e}_2$). Sa matrice est donc, suivant (2.4) $\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$ et si on écrit un vecteur \vec{X} dans les deux bases $\vec{X} = x_1 \vec{e}_1 + x_2 \vec{e}_2$ et

$$\vec{X} = y_1 \vec{f_1} + y_2 \vec{f_2} = y_1 (\vec{e_1} + \vec{e_2}) + y_2 (2\vec{e_1} - \vec{e_2}) = (y_1 + 2y_2)\vec{e_1} + (y_1 - y_2)\vec{e_2}$$

en identifiant les composantes sur \vec{e}_1 et \vec{e}_2 , on trouve : $x_1 = y_1 + 2y_2$, $x_2 = y_1 - y_2$, qui est en accord avec (2.5).

▷ Composition de 2 changements de base et multiplication matricielle.

Supposons maintenant qu'on effectue successivement deux changements de base. D'abord $\{\vec{e}\} \to \{\vec{f}\}$ avec une matrice A de changement de base, puis $\{\vec{f}\} \to \{\vec{g}\}$ avec la matrice B. On écrit donc

$$ec{f_j} = \sum_{i=1}^n ec{e_i} a_{ij} \qquad ec{g}_k = \sum_{j=1}^n ec{f_j} b_{jk}$$

soit, en reportant la première expression de \vec{f}_j dans la deuxième expression

$$\vec{g}_k = \sum_{j=1}^n \sum_{i=1}^n \vec{e}_i a_{ij} b_{jk} = \sum_{i=1}^n \vec{e}_i c_{ik}$$

où le coefficient c_{ik} qui donne la décomposition des \vec{g}_k dans la base des \vec{e}_i s'identifie à

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} . (2.6)$$

Cela introduit de façon naturelle la multiplication des deux matrices A et B.

Définition : Le produit des deux matrices carrées $n \times n$ A et B est la matrice carrée C dont les éléments sont donnés par l'expression (2.6).

$$A \cdot B = C = (c_{ij})$$
 $c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$. (2.7)

Il faut bien retenir les opérations qu'il faut faire pour construire le produit de deux matrices : selon (2.6), l'élément (i, j) de la matrice produit s'obtient par la somme des produits des éléments de la ligne i de la matrice de gauche par ceux de la colonne j de la matrice de droite. On peut éventuellement visualiser ces opérations par une disposition en équerre des trois matrices, illustrée ici dans le cas n=2

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

Exemples. Vérifier que

$$\begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ -4 & 5 \end{pmatrix}$$

et que

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix} .$$

Que valent $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$?

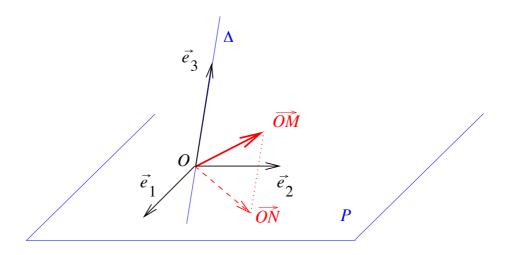


Fig. 3: Projection d'un vecteur \overrightarrow{OM} de l'espace \mathbb{R}^3 dans le plan \mathbb{R}^2 .

3. Applications linéaires

3.1. Projection

Considérons dans l'espace ordinaire à trois dimensions (\mathbb{R}^3) un plan P et un axe Δ n'appartenant pas à ce plan, coupant P en O. On définit la projection $\mathcal{P}(M)$ de tout point M sur P parallèlement à la direction Δ comme suit : une droite passant par M et parallèle à Δ coupe P en $N: N = \mathcal{P}(M)$ est le projeté de M, ou encore $\overrightarrow{ON} = \mathcal{P}(\overrightarrow{OM})$, voir fig. 3.

On peut redire cela dans le langage des espaces vectoriels : on choisit une base \vec{e}_1, \vec{e}_2 de l'espace vectoriel \mathbb{R}^2 des vecteurs de P d'origine O. Selon le principe expliqué au § 2.3, cette base est complétée en une base de \mathbb{R}^3 par un troisième vecteur \vec{e}_3 , qu'on choisit porté par l'axe Δ . Pour tout point M, le vecteur \overrightarrow{OM} étant écrit d'une manière unique sous la forme $\overrightarrow{OM} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$, on a $\overrightarrow{ON} = \mathcal{P}(\overrightarrow{OM}) = x_1\vec{e}_1 + x_2\vec{e}_2$.

Dans cette nouvelle façon de voir les choses, il est clair que l'opération de projection \mathcal{P} est linéaire, ce qui signifie que quels que soient les coefficients λ et λ' et les vecteurs $\vec{X} = \overrightarrow{OM}$ et $\vec{X}' = \overrightarrow{OM'}$

$$\mathcal{P}(\lambda \vec{X} + \lambda' \vec{X}') = \lambda \mathcal{P}(\vec{X}) + \lambda' \mathcal{P}(\vec{X}') , \qquad (3.1)$$

et c'est une opération qui envoie tout vecteur de l'espace de départ \mathbb{R}^3 dans l'espace \mathbb{R}^2 .

3.2. Application linéaire

Définition : En général on dira qu'une application \mathcal{A} d'un espace vectoriel E dans un espace vectoriel F est une application linéaire (on dit aussi un opérateur linéaire) si elle satisfait la relation (3.1) (avec \mathcal{P} changée en \mathcal{A}) pour tous $\lambda, \lambda' \in \mathbb{R}$, et tous $\vec{X}, \vec{X}' \in E$, soit

$$\mathcal{A}(\lambda \vec{X} + \lambda' \vec{X}') = \lambda \mathcal{A}(\vec{X}) + \lambda' \mathcal{A}(\vec{X}') . \tag{3.1}$$

Exemples.

- 1. Comme on vient de voir, la projection sur un sous-espace est une application linéaire.
- 2. L'espace d'arrivée peut bien sûr aussi coïncider avec l'espace de départ E, on parle alors de transformation linéaire (ou d'"endomorphisme") de l'espace E. Par exemple, une rotation dans l'espace euclidien \mathbb{R}^2 (ou \mathbb{R}^3) est une transformation linéaire de cet espace. Exercice : vérifier que la relation (3.1)′ est bien satisfaite par une rotation dans le plan (pour la somme, se rappeler la règle du parallélélogramme).
- 3. Une application linéaire de E dans l'espace \mathbb{R} des réels est appelée forme linéaire sur E. Par exemple, soit E = C([a,b]) l'espace des fonctions continues réelles sur l'intervalle [a,b]. L'application $f \mapsto \int_a^b f(x) \, \mathrm{d}x$ est une application linéaire de C([a,b]) dans l'espace \mathbb{R} des réels, c'est donc une forme linéaire sur C([a,b]).
- 4. La dérivation est une application linéaire dans l'espace des fonctions dérivables, ce qu'on utilisera plus bas dans l'étude des équations différentielles linéaires.

⊳ Noyau, image et rang d'une application linéaire

Trois définitions:

- 1. On appelle noyau d'une application linéaire \mathcal{A} de E dans F l'ensemble des $\vec{X} \in E$ tels que $\mathcal{A}(\vec{X}) = 0$. Exercice : en utilisant (3.1)', montrer que ce noyau est un espace vectoriel, donc un sous-espace de E.
- 2. On appelle image de l'application \mathcal{A} l'ensemble $\mathcal{A}(E)$. Exercice : en utilisant (3.1)', montrer que c'est un espace vectoriel, donc un sous-espace vectoriel de F.
- 3. Le rang de l'application linéaire est la dimension de l'espace image : $\operatorname{rang}(A) = \dim(A(E))$.

Notant $n = \dim E$, nous avons alors l'important

Théorème 2 : dim (noyau(A)) + rang(A) = n .

Preuve. Supposons le noyau de \mathcal{A} de dimension s (éventuellement nul, si le noyau se réduit à $\vec{0}$). Soit $\vec{e}_1, \cdots, \vec{e}_s$ une base du noyau (éventuellement vide si s=0) et complétons-la comme expliqué au § 2.3 par n-s vecteurs $\vec{e}_{s+1}, \cdots, \vec{e}_n$ pour avoir une base de E. Puisque $\mathcal{A}(\vec{e}_i)=0$ pour $i=1,\cdots,s$, pour tout \vec{X} , $\mathcal{A}(\vec{X})$ est combinaison linéaire de $\mathcal{A}(\vec{e}_{s+1}), \cdots, \mathcal{A}(\vec{e}_n)$, qui engendrent donc l'image $\mathcal{A}(E)$. Montrons que ces vecteurs $\mathcal{A}(\vec{e}_{s+1}), \cdots, \mathcal{A}(\vec{e}_n)$ sont linéairement indépendants. S'il n'en était pas ainsi, il existerait des λ_i non tous nuls tels que $\sum_{i=s+1}^n \lambda_i \mathcal{A}(\vec{e}_i) = \mathcal{A}(\sum_{i=s+1}^n \lambda_i \vec{e}_i) = 0$, et alors $\sum_{i=s+1}^n \lambda_i \vec{e}_i$ appartiendrait au noyau de \mathcal{A} , ce qui est en contradiction avec la définition des vecteurs $\vec{e}_{s+1}, \cdots, \vec{e}_n$ comme formant une base du sous-espace supplémentaire du noyau. On a bien montré que la dimension de $\mathcal{A}(E)$, c'est-à-dire le rang de \mathcal{A} , est égale à n-s, ce qui établit le Théorème.

Exemple: Quelle est l'image, quel est le noyau d'une projection de \mathbb{R}^3 dans \mathbb{R}^2 comme celle considérée au § 3.1 ? Même question avec une rotation dans \mathbb{R}^3 .

3.3. Matrice d'une application linéaire

Considérons une application linéaire \mathcal{A} d'un espace vectoriel E de dimension n dans un espace vectoriel F de dimension p, et supposons qu'on connaisse une base dans chacun de ces espaces : $\{\vec{e}_i\}$, $i=1,\dots n$ dans E et $\{f_j\}$, $j=1,\dots,p$ dans F. On va montrer que l'on connaît l'application linéaire si et seulement si ("ssi") on sait comment les transformés des \vec{e}_i s'expriment en termes des \vec{f}_j

$$\mathcal{A}(\vec{e}_i) = \sum_{i=1}^p \vec{f}_j a_{ji} , \qquad (3.2)$$

c'est-à-dire si on connaît la matrice a_{ji} de l'application linéaire \mathcal{A} dans les deux bases \vec{e} et \vec{f} . Cette matrice est un tableau $p \times n$ (c'est-à-dire à p lignes et n colonnes) puisque le premier indice (indice de lignes) $j = 1, \dots p$, tandis que le deuxième (indice de colonnes) prend les valeurs $i = 1, \dots n$.

³ Bien noter que a_{ji} est la composante sur \vec{f}_j de $\mathcal{A}(\vec{e}_i)$.

Soit un vecteur $\vec{X} = \sum_{i=1}^{n} x_i \vec{e_i}$ de l'espace E et $\vec{Y} = \sum_{j=1}^{p} y_j \vec{f_j}$ son transformé (ou son "image") par l'application \mathcal{A} . Par linéarité et en utilisant (3.2), on sait que

$$\vec{Y} = \mathcal{A}(\vec{X}) = \mathcal{A}(\sum_{i=1}^{n} x_i \vec{e_i}) = \sum_{i=1}^{n} x_i \mathcal{A}(\vec{e_i}) = \sum_{i=1}^{n} \sum_{j=1}^{p} x_i a_{ji} \vec{f_j}$$

donc en identifiant la composante de \vec{Y} sur le vecteur \vec{f}_i

$$y_j = \sum_{i=1}^n a_{ji} x_i \ . \tag{3.3}$$

Supposons maintenant qu'on effectue deux applications linéaires successivement : \mathcal{A} de l'espace E dans l'espace F, puis \mathcal{B} de l'espace F dans un espace G. On note \vec{Z} l'image de \vec{X} par cette composition des deux applications, et on écrit $\vec{Y} = \mathcal{A}\vec{X}$, $\vec{Z} = \mathcal{B}\vec{Y} = (\mathcal{B}\mathcal{A})\vec{X}$. Supposons l'espace G de dimension q et muni d'une base \vec{g}_k , $k = 1, \dots, q$: un vecteur \vec{Z} de G s'écrit $\vec{Z} = \sum_{k=1}^q z_k \vec{g}_k$. On s'est donné la matrice a_{ij} de \mathcal{A} dans les bases \vec{e} et \vec{f} , et la matrice b_{jk} de \mathcal{B} dans les bases \vec{f} et \vec{g} . Quelle est la matrice de la composée $\mathcal{B}\mathcal{A}$ dans les bases \vec{e} et \vec{g} ? Le calcul est très simple

$$\vec{Z} = (\mathcal{B}\mathcal{A})\vec{X} = \mathcal{B}(\mathcal{A}(\vec{X})) = \sum_{i=1}^{n} \sum_{j=1}^{p} x_i a_{ji} \mathcal{B}(\vec{f}_j)$$
$$= \sum_{i=1}^{n} \sum_{j=1}^{p} \sum_{k=1}^{q} x_i a_{ji} b_{kj} \vec{g}_k$$

donc en identifiant le coefficient de \vec{q}_k

$$z_k = \sum_{i=1}^n \sum_{j=1}^p b_{kj} a_{ji} x_i = \sum_{i=1}^n (B \cdot A)_{ki} x_i$$

avec le produit des deux matrices $B = (b_{ij})$ et $A = (a_{ij})$ dans cet ordre défini par

$$(B \cdot A)_{ki} = \sum_{j=1}^{p} b_{kj} a_{ji} .$$
(3.4)

Exemple:

$$\underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \end{pmatrix}}_{B} \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 3 \\ 4 & -1 & 2 \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} 17 & -2 & 13 \\ -10 & 2 & -7 \end{pmatrix}}_{B:A}.$$

Plusieurs remarques importantes

- nous avons été menés à cette définition de la multiplication de deux matrices de façon naturelle, en suivant les conséquences de la linéarité des applications ;
- cette définition étend au cas de "matrices rectangulaires" la définition (2.6) de la multiplication de matrices carrées ;
- bien noter qu'il n'est cohérent de multiplier une matrice B par une matrice A que si le nombre de colonnes de B égale celui des lignes de A (noté ici p);
- bien noter aussi que l'ordre des opérations a été dicté par la composition des applications. Même si on n'a à faire qu'à des matrices carrées, n=p=q, il n'est pas vrai en général que $B \cdot A = A \cdot B$.

Le produit matriciel n'est en général pas commutatif!

- L'application \mathcal{A} suivie de \mathcal{B} a mené à la composée $\mathcal{B}\mathcal{A}$ et au produit matriciel $B \cdot A$: il faut garder à l'esprit que l'action des opérateurs \mathcal{B} , \mathcal{A} , s'effectue sur un objet à droite, et donc dans le produit $\mathcal{B}\mathcal{A}$, l'action de \mathcal{A} précède celle de \mathcal{B} .
- \triangleright Exemple. Considérons les deux matrices 2×2

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} . \tag{3.5}$$

Calculer les deux produits $A \cdot B$ et $B \cdot A$ et vérifier qu'ils sont différents.

Cette propriété de non-commutation va jouer un rôle important dans la suite. Elle implique par exemple que les identités familières du calcul algébrique usuel doivent être reconsidérées. Ainsi pour deux matrices A et B, l'"identité du binôme" se lit

$$(A+B)^{2} = A^{2} + A \cdot B + B \cdot A + B^{2}$$
(3.6)

et non $A^2 + 2A \cdot B + B^2$.

Exercice: comment s'écrit $(A+B)^3$?

4. Matrices

4.1. Produit matriciel

Récapitulons ce que nous avons appris sur les matrices.

– Le produit de deux matrices B et A dans cet ordre, B de dimensions $q \times p$ et A de dimensions $p \times n$ est défini par l'expression (3.4).

- Si les matrices sont carrées, ce produit n'est en général pas commutatif : $A \cdot B \neq B \cdot A$ en général.
- Le produit matriciel est associatif, ce qui veut dire qu'on peut regrouper à sa guise les facteurs d'un produit (sans modifier leur ordre !), $A \cdot (B \cdot C) = (A \cdot B) \cdot C$, et cela est fort utile dans les calculs...

La matrice identit'e de dimension n est par définition la matrice ayant des 1 sur sa diagonale, des zéros partout ailleurs

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} .$$

Elle a la particularité d'être l'élément neutre de la multiplication matricielle : si A est une matrice quelconque $p \times n$ et \mathbb{I}_n , resp. \mathbb{I}_p les matrices identité dans les dimensions indiquées

$$\mathbb{I}_p \cdot A = A \cdot \mathbb{I}_n \ .$$

L'élément de matrice (i, j) de la matrice \mathbb{I}_n s'écrit $(\mathbb{I}_n)_{ij} = \delta_{ij}$ avec le symbole de Kronecker $\delta_{ij} = 0$ ou 1 selon que $i \neq j$ ou i = j.

Une autre matrice d'usage courant est la matrice nulle en dimension n, dont tous les éléments sont nuls. Elle est notée 0_n , ou simplement 0 quand cela ne prête pas à ambiguïté.

$$0_n = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} .$$

Elle a la propriété que son produit par toute matrice redonne la matrice 0:

$$\forall A, \quad A \cdot 0_n = 0_n \cdot A = 0$$
.

Une matrice carrée $n \times n$ A **peut** avoir une inverse A^{-1} telle que

$$A \cdot A^{-1} = A^{-1} \cdot A = \mathbb{I}_n$$
.

soit encore $(A \cdot A^{-1})_{ij} = \sum_{k} (A)_{ik} (A^{-1})_{kj} = \delta_{ij}$. Une telle matrice A est dite inversible.

Mais **attention!**, une matrice peut aussi ne pas avoir d'inverse! On dit alors qu'elle est *non-inversible*, ou encore *singulière*. On va étudier plus bas quel critère doit satisfaire une matrice pour être inversible, puis comment calculer son inverse dans ce cas.⁴

Exemples. L'inverse de la matrice A de (3.5) est la matrice A elle-même ; même chose pour la matrice B de (3.5) ; la matrice inverse de $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ est $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$. D'une façon générale, toute matrice de changement de base, cf (2.4), doit être inversible, puisqu'on doit pouvoir exprimer les $\vec{e_i}$ en termes des $\vec{f_j}$ et vice versa et ce sont les matrices A et A^{-1} qui effectuent ces changements de base : $\vec{f_j} = \sum_i \vec{e_i} a_{ij}$, $\vec{e_i} = \sum_j \vec{f_j} (A^{-1})_{ji}$, si A est la matrice (a_{ij}) .

En revanche, la matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ est non-inversible. En effet cherchons une matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ telle que

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix} \stackrel{?}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \,.$$

Il est clair qu'on se heurte à une impossibilité, celle de reproduire le 1 inférieur de la diagonale de \mathbb{I}_2 . Vérifier de même qu'il n'existe pas d'inverse à gauche.

Plus étonnant encore : il peut exister des matrices A et B non nulles telles que $A \cdot B = 0$. Par exemple, vérifier que la matrice $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ est de carré nul (on dit aussi qu'elle est nilpotente)

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

Pour finir cette discussion des propriétés surprenantes du calcul matriciel (liées à la non-commutation en général de deux matrices), considérons deux matrices carrées inversibles A et B. On vérifie que le produit $B \cdot A$ est lui-même inversible avec comme inverse

$$(B \cdot A)^{-1} = A^{-1} \cdot B^{-1} , \qquad (4.1)$$

puisqu'on voit immédiatement que

$$(B \cdot A)^{-1} \cdot (B \cdot A) = A^{-1} \cdot B^{-1} \cdot B \cdot A = A^{-1} \cdot A = \mathbb{I}$$

⁴ On peut se demander si la matrice inverse à gauche et la matrice inverse à droite sont bien toujours égales. On montrera plus bas (Théorème à la fin de ce chapitre) qu'elles existent ou n'existent pas simultanément. Soient A_g^{-1} et A_d^{-1} telles que $A_g^{-1}.A = \mathbb{I}$ et $A.A_d^{-1} = \mathbb{I}$. Un calcul simple montre alors que $A_g^{-1} = A_g^{-1}.(A.A_d^{-1}) = (A_g^{-1}.A).A_d^{-1} = A_d^{-1}$, donc A_g^{-1} et A_d^{-1} sont bien égales !

(où on a utilisé l'associativité du produit). Bien noter sur (4.1) que l'inversion a renversé l'ordre des facteurs A et B.

4.2. Addition et multiplication par un scalaire des matrices

Pour deux matrices A et B de mêmes dimensions $p \times n$, on peut définir leur somme A + B, qui est la matrice $p \times n$ d'élément (i, j) donné par

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

on peut aussi définir le produit λA de A par un scalaire λ , avec $(\lambda A)_{ij} = \lambda a_{ij}$. (Cela signifie que les matrices $p \times n$ forment elles-mêmes un espace vectoriel.)

Ces opérations de somme et produit par un scalaire sont compatibles avec le produit matriciel en ce sens que

$$(\lambda_1 A_1 + \lambda_2 A_2) \cdot B = \lambda_1 A_1 \cdot B + \lambda_2 A_2 \cdot B$$
$$A \cdot (\lambda_1 B_1 + \lambda_2 B_2) = \lambda_1 A \cdot B_1 + \lambda_2 A \cdot B_2,$$

c'est la "distributivité" de la multiplication matricielle (à gauche, resp. à droite) par rapport à l'addition.

4.3. Changement de base pour la matrice d'une application

Soit \mathcal{A} un opérateur (ou application) linéaire d'un espace vectoriel E dans un espace F. Supposons que dans une base $\vec{e_i}$ de E et une base $\vec{f_j}$ de F, l'opérateur \mathcal{A} est représenté par une matrice A. Si on change de bases dans les espaces E et F, comment cette matrice est-elle changée? La réponse est obtenue aisément en combinant les formules (3.2) et (2.4).

Soit \vec{e}' la nouvelle base de E, \vec{f}'_j celle de F. En modifiant un peu les notations de (2.4), soient V et W les matrices de ces changements de base

$$ec{e}'_{i} = \sum_{i'} \vec{e}_{i'} V_{i'i}$$
 $ec{f}'_{j} = \sum_{j'} \vec{f}_{j'} W_{j'j}$.

Les matrices V et W doivent être inversibles, puisque les anciens vecteurs de base peuvent s'exprimer dans les nouveaux. En particulier

$$\vec{f}_{j'} = \sum_{j} \vec{f}'_{j} W_{jj'}^{-1}. \tag{4.2}$$

On écrit alors

$$\mathcal{A}(\vec{e}_i') = \sum_{i'} \mathcal{A}(\vec{e}_{i'}) V_{i'i} \quad \text{par linéarité de } \mathcal{A}$$

$$= \sum_{i'j'} \vec{f}_{j'} a_{j'i'} V_{i'i} \quad \text{par (3.2)}$$

$$= \sum_{i'j'j} \vec{f}_j' W_{jj'}^{-1} a_{j'i'} V_{i'i} \quad \text{par (4.2)}$$

$$= \sum_{i} \vec{f}_j' (A')_{ji} \quad \text{par définition de } A'$$

d'où il découle que la nouvelle matrice A' exprimant l'opérateur A dans les nouvelles bases est reliée à la matrice originale par

$$A' = W^{-1}AV. (4.3)$$

Les deux matrices A et $A' = W^{-1}AV$ qui représentent le même opérateur \mathcal{A} dans des bases différentes sont dites équivalentes. (Inversement toute paire de matrices (A, A') de cette forme peut s'interpréter comme représentant le même opérateur \mathcal{A} dans deux bases différentes.)

Dans le cas particulier où E = F et où l'opérateur \mathcal{A} est donc un opérateur linéaire de E dans lui-même (un "endomorphisme"), ces expressions se réduisent à

$$A' = V^{-1}AV \tag{4.4}$$

et les matrices A et $A' = V^{-1}AV$ sont dites semblables.

Exemple: Comment se transforme la matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ par le changement de base $(\vec{e}_1, \vec{e}_2) \mapsto (\vec{e}_2, \vec{e}_1)$? par $(\vec{e}_1, \vec{e}_2) \mapsto (\vec{e}_1 + \vec{e}_2, 2\vec{e}_1 - \vec{e}_2)$? Utiliser les résultats des exemples du § 2.4.

4.4. Autres définitions et propriétés

 \triangleright Transposée d'une matrice. Soit $A = (a_{ij})$ une matrice $p \times n$, avec $i = 1, \dots, p, j = 1, \dots, n$. La matrice transposée, notée A^T (quelquefois aussi tA), est la matrice $n \times p$ dans laquelle les lignes et les colonnes ont été échangées. Autrement dit

$$(A^T)_{ij} = a_{ji} , \quad \text{avec} \quad i = 1, \dots, \quad j = 1, \dots p .$$
 (4.5)

Attention que la transposition renverse l'ordre des facteurs d'un produit

$$(B \cdot A)^T = A^T \cdot B^T . (4.6)$$

En effet cela résulte de l'expression du produit et de la définition (4.5)

$$(A^T \cdot B^T)_{ij} = \sum_k (A^T)_{ik} (B^T)_{kj} = \sum_k a_{ki} b_{jk} = \sum_k b_{jk} a_{ki} = (B \cdot A)_{ji} = ((B \cdot A)^T)_{ij} .$$

Si A est carrée et inversible, sa transposée est aussi inversible. Montrons que l'inverse de la transposée est la transposée de l'inverse

$$(A^T)^{-1} = (A^{-1})^T . (4.7)$$

En effet, en utilisant (4.6)

$$(A^{-1})^T \cdot A^T = (A \cdot A^{-1})^T = \mathbb{I}^T = \mathbb{I}$$

comme annoncé.

Exercice: Calculer $((A \cdot B^{-1} \cdot C)^{-1})^T$.

ightharpoonup Matrice symétrique

Une matrice carrée A égale à sa transposée, donc telle que $\forall i, j, \ a_{ij} = a_{ji}$ est dite symétrique. Ses éléments sont symétriques par rapport à la diagonale, ainsi $\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$.

Exercice. Montrer que pour toute matrice A de dimensions $p \times n$, les matrices $A \cdot A^T$ et $A^T \cdot A$ sont carrées et symétriques.

ightharpoonup Trace d'une matrice carrée : par définition la trace d'une matrice carrée est la somme de ses éléments diagonaux :

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii} \ . \tag{4.8}$$

Calculons la trace du produit $A \cdot B$ de deux matrices carrées. On a

$$\operatorname{tr}(A \cdot B) = \sum_{i} (A \cdot B)_{ii} = \sum_{ij} a_{ij} b_{ji} = \sum_{ij} b_{ji} a_{ij} = \sum_{j} (B \cdot A)_{jj} = \operatorname{tr}(B \cdot A)$$

donc

$$tr A \cdot B = tr B \cdot A , \qquad (4.9)$$

c'est l'importante propriété de cyclicité de la trace.

Corollaire 2 : Deux matrices semblables (au sens de (4.4)) ont même trace. En effet $\operatorname{tr} A' = \operatorname{tr} (V^{-1} \cdot A \cdot V) = \operatorname{tr} (V \cdot V^{-1} \cdot A) = \operatorname{tr} (\mathbb{I} \cdot A) = \operatorname{tr} A.$

Exercice. Soient A, B, C trois matrices carrées de même dimension. Les traces tr(A.B.C), tr(B.C.A), tr(C.B.A) sont-elles égales ?

4.5. Matrices-lignes, matrices-colonnes

Soit \vec{X} un vecteur d'un espace de dimension n, et soient x_i , $i=1,\dots,n$ ses composantes dans une certaine base. Par définition la matrice-colonne X est la matrice à une seule colonne $X_{i1} = x_i$. Autrement dit (dans la base considérée), on peut représenter le vecteur \vec{X} par la matrice $n \times 1$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} . \tag{4.10}$$

Noter qu'avec ces notations, la relation (3.3) entre les composantes du vecteur \vec{X} et celle de $\vec{Y} = \mathcal{A}\vec{X}$ s'exprime par un produit matriciel

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = A \cdot X . \tag{4.11}$$

Exercice. Quelle est la matrice-colonne représentant le vecteur de base \vec{e}_i ?

Il peut être commode de considérer aussi des vecteurs à une seule ligne, appelés matrices-lignes. Ainsi la transposée de X de l'équ. (4.10) est une matrice-ligne

$$X^T = (x_1 \ x_2 \ \cdots x_n) \ . \tag{4.12}$$

Toute matrice $p \times n$ peut être considérée comme la juxtaposition de n matrices-colonnes $\mathbf{A}^j, j = 1, \dots, n$, à p composantes $(\mathbf{A}^j)_i = a_{ij}, i = 1, \dots, p$

$$A = (a_{ij}) = (\mathbf{A}^1 \ \mathbf{A}^2 \ \cdots \ \mathbf{A}^n) = \begin{pmatrix} \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{p1} \end{pmatrix} \quad \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{p2} \end{pmatrix} \quad \cdots \quad \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{pn} \end{pmatrix}$$
(4.13)

Les vecteurs \mathbf{A}^j , $j=1,\dots,n$, sont les *vecteurs-colonnes* de la matrice A. On peut aussi représenter cette même matrice comme une juxtaposition de p matrices-lignes \mathbf{A}_i , $i=1,\dots,p$, à n composantes $(\mathbf{A}_i)^j=a_{ij},\ j=1,\dots,n$

$$A = \begin{pmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_p \end{pmatrix} = \begin{pmatrix} (a_{11} \ a_{12} \ \cdots \ a_{1n}) \\ (a_{21} \ a_{22} \ \cdots \ a_{2n}) \\ \vdots \\ (a_{p1} \ a_{p2} \ \cdots \ a_{pn}) \end{pmatrix} .$$

Les vecteurs \mathbf{A}_i , $i = 1, \dots, p$, sont les vecteurs-lignes de la matrice A. (Noter la position des indices sur \mathbf{A}^j et \mathbf{A}_i .)

Ces définitions vont nous être utiles pour discuter l'inversibilité d'une matrice.

Interprétation des vecteurs colonnes.

Soit A la matrice d'une application \mathcal{A} dans deux bases $\vec{e_i}$ et $\vec{f_j}$, cf (3.2). Selon une observation déjà faite plus haut (note en bas de page suivant (3.2)), l'élément a_{ji} de la matrice A est la composante sur $\vec{f_j}$ de $\mathcal{A}(\vec{e_i})$, que nous notons $\mathcal{A}(\vec{e_i})_j$:

$$(\mathcal{A}(\vec{e_i}))_j = a_{ji} = (\mathbf{A}^i)_j \qquad i = 1, \dots, p \qquad j = 1, \dots, n.$$

Autrement dit, les vecteurs colonnes \mathbf{A}^i sont les images des vecteurs de base \vec{e}_i par l'application $\mathcal{A}: \mathbf{A}^i = \mathcal{A}(\vec{e}_i)$ et on peut écrire comme en (4.13)

$$A = (\mathcal{A}(\vec{e}_1) \, \mathcal{A}(\vec{e}_2) \, \cdots \, \mathcal{A}(\vec{e}_n)) \ .$$

4.6. Rang d'une matrice

Lemme : Pour toute matrice $m \times n$, le rang du système de vecteurs-colonnes égale celui de ses vecteurs-lignes.

Preuve : Soit r le rang du système des vecteurs-colonnes. Le rang d'un système de vecteurs étant inchangé si on permute ces vecteurs, on peut toujours supposer que ce sont les r premiers vecteurs-colonnes $\mathbf{A}^i, i=1,\cdots,r$ qui sont indépendants. Les n-r suivants sont donc des combinaisons linéaires des r premiers

$$\mathbf{A}^{r+s} = \sum_{j=1}^{r} \lambda_j^s \mathbf{A}^j \qquad \forall s = 1, \dots, n-r.$$
 (4.14)

Les *i*-ièmes composantes des vecteurs-colonnes forment la *i*-ième ligne de la matrice. Les relations (4.14) permettent d'exprimer les n-r derniers éléments de chaque ligne comme combinaison linéaire des r premiers, avec des coefficients indépendants de la ligne considérée

$$\forall i = 1, \dots, m$$
 $(\mathbf{A}^{r+s})_i = a_{i\,r+s} = \sum_{j=1}^r \lambda_j^s a_{ij}$ (4.15)

Si on écrit le *i*-ème vecteur-ligne comme $\mathbf{A}_i = \sum_{k=1}^n a_{ik} \mathbf{e}_k$ avec \mathbf{e}_k le vecteur-ligne de composantes $(\mathbf{e}_k)^j = \delta_{kj}$, (ce qui est bien équivalent à $(\mathbf{A}_i)^j = a_{ij}$), (4.15) permet d'écrire

$$\forall i = 1, \dots, m \qquad \mathbf{A}_i = \sum_{j=1}^n a_{ij} \mathbf{e}_j = \sum_{j=1}^r a_{ij} \mathbf{e}_j + \sum_{s=1}^{n-r} \sum_{j=1}^r \lambda_j^s a_{ij} \mathbf{e}_{r+s}$$
$$= \sum_{j=1}^r a_{ij} \left(\mathbf{e}_j + \sum_{s=1}^{n-r} \lambda_j^s \mathbf{e}_{r+s} \right) = \sum_{j=1}^r a_{ij} \mathbf{e}_j',$$

avec pour $j=1,\cdots,r,$ $\mathbf{e}'_j=\mathbf{e}_j+\sum_{s=1}^{n-r}\lambda_j^s\mathbf{e}_{r+s}.$ On a donc montré que les m vecteurs-lignes de A sont combinaisons linéaires des r vecteurs $\mathbf{e}'_j,$ $j=1,\cdots,r$. Leur rang r' est donc inférieur ou égal à r. Mais on peut reprendre ce raisonnement en échangeant le rôle des lignes et des colonnes, ce qui conduit à la conclusion que $r\leq r'$. On conclut que r=r', c.q.f.d.

Définition : Le rang d'une matrice est le rang de son système de vecteurs-colonnes, ou celui de son système de vecteurs-lignes.

Théorème 3 : Une matrice carrée $n \times n$ est inversible si et seulement si son rang est égal à n.

Preuve. Supposons que le rang des vecteurs colonnes $\vec{e}'_i = \mathbf{A}^i = \mathcal{A}(\vec{e}_i)$ égale n. Cela signifie que l'on peut les considérer comme une base de l'espace E, et donc que les \vec{e}_i peuvent s'exprimer comme combinaisons linéaires de ces \vec{e}'_i . Mais cela signifie que l'application inverse \mathcal{A}^{-1} existe puisque $\vec{e}_i = \mathcal{A}^{-1}(\vec{e}'_i)$, donc aussi la matrice inverse : $\vec{e}_i = \sum_j \vec{e}'_j (A^{-1})_{ji}$. La réciproque est claire : l'inversibilité de A garantit que les \vec{e}_i peuvent s'exprimer linéairement en termes des \vec{e}'_i qui forment donc une base, et la matrice est bien de rang n.

5. Vecteurs, tenseurs

Dans ce paragraphe nous introduisons de nouveaux objets de l'algèbre linéaire, les tenseurs. Nous allons nous intéresser en particulier aux transformations de ces tenseurs sous l'effet des changements de base. Nous montrerons finalement à travers quelques exemples que ces tenseurs sont très utiles au physicien et à l'ingénieur.

5.1. Produit tensoriel

Soit E un espace vectoriel de dimension n, et $\{\vec{e}_i\}$ une base de cet espace $(i=1,\dots,n)$. On construit alors une collection de n^2 objets notés $\vec{e}_i \otimes \vec{e}_j$ et on considère l'espace vectoriel de leurs combinaisons linéaires

$$T = \sum_{i=1}^{n} \sum_{j=1}^{n} t^{ij} \vec{e}_i \otimes \vec{e}_j \qquad t^{ij} \in \mathbb{R}.$$
 (5.1)

Cet espace est appelé produit tensoriel de l'espace E avec lui-même et noté $E \otimes E$. Tout élément T de $E \otimes E$ (donc tout vecteur de cet espace) est appelé tenseur (de rang 2, voir plus bas). En particulier, étant donnés deux vecteurs $\vec{X}_1 = \sum_i x_1^i \vec{e}_i$ et $\vec{X}_2 = \sum_j x_2^j \vec{e}_j$ de E, leur produit tensoriel $\vec{X}_1 \otimes \vec{X}_2$ est le tenseur

$$\vec{X}_1 \otimes \vec{X}_2 = \sum_{i=1}^n \sum_{j=1}^n x_1^i x_2^j \, \vec{e}_i \otimes \vec{e}_j \,. \tag{5.2}$$

Remarques.

- (a) On peut plus généralement définir d'une façon analogue le produit tensoriel de deux espaces différents $E \otimes F$, mais nous n'en aurons pas besoin dans la suite.
- (b) Les mathématiciens ont une définition plus formelle du produit tensoriel $E \otimes E$ qui ne repose pas sur des choix de base. L'approche suivie ici est pragmatique et vise aux applications physiques.

5.2. Formules de changement de base

Nous commençons par récrire les formules (2.5) et (4.3) avec des notations un peu modifiées. Soient à nouveau E et F deux espaces vectoriels de dimensions respectives n et p, et \mathcal{A} une application linéaire de E dans F. Soit \vec{X} un vecteur de E, x^i ses composantes dans une base \vec{e}_i que nous regroupons en une matrice-colonne X, et x'^i dans une autre base $\vec{e}_i' = \sum_{i'} \vec{e}_{i'} V_{i'}^{i'}$, avec X' la matrice-colonne des x'^i . De même tout vecteur \vec{Y} de F est repéré par ses composantes y^j et y'^j , $j = 1, \dots p$, dans deux bases \vec{f}_j et $\vec{f}_j' = \sum_{j'} \vec{f}_{j'} W_j^{j'}$, formant des matrices-colonnes Y et Y'. Enfin l'application linéaire A est représentée dans les bases $\{\vec{e}, \vec{f}\}$ et $\{\vec{e}', \vec{f}'\}$ respectivement par des matrices A et A', comme on a vu plus haut au $\S 4.3$. On note que nous avons placé les indices tantôt en haut (indice des composantes des vecteurs, indice de ligne des matrices), tantôt en bas (indice des vecteurs de base, indice de colonne des matrices). Les formules (2.5) et (4.3) peuvent se récrire selon

$$\vec{e}_{i}' = \sum_{i'} \vec{e}_{i'} V_{i}^{i'} \qquad \vec{f}_{j}' = \sum_{j'} \vec{e}_{j'} W_{j}^{j'}$$

$$x'^{i} = \sum_{i'} (V^{-1})_{i'}^{i} x^{i'} \qquad \text{soit} \quad X' = V^{-1} X$$

$$y'^{j} = \sum_{i'} (W^{-1})_{j'}^{j} y^{j'} \qquad \text{soit} \quad Y' = W^{-1} Y$$

$$A'^{i}_{j} = (W^{-1})_{k}^{i} A_{l}^{k} V_{j}^{l} \qquad \text{soit} \quad A' = W^{-1} A V, \text{ cf } (4.3)$$

Bien noter la cohérence des notations : (1) les sommations se font toujours sur une paire d'indices l'un supérieur, l'autre inférieur ; (2) les objets avec un indice supérieur (resp. inférieur) se transforment différemment, par multiplication à gauche par une matrice V^{-1} ou W^{-1} pour les premiers, par multiplication à droite par la matrice V ou W pour les seconds.

On note aussi que le vecteur $\vec{Y} = \mathcal{A}(\vec{X})$ a pour composantes Y = AX dans les bases $\{\vec{e}, \vec{f}\}$ et $Y' = A'X' = W^{-1}AV.V^{-1}X = W^{-1}AX = W^{-1}Y$ dans la base $\{\vec{e}', \vec{f}'\}$, donc qu'il se transforme bien comme un vecteur de l'espace F, comme il se doit. En particulier si on n'effectue qu'un changement de la base \vec{e} en \vec{e}' (sans changer la base \vec{f}), ses composantes

Y ne changent pas, en vertu de la compensation de V et V^{-1} dans les transformations de A et de X.

o Vecteur gradient. Vecteurs contravariants et covariants

Il est souvent important d'effectuer des dérivations par rapport aux composantes x^i d'un vecteur \vec{X} et de bien voir comment ces opérations se transforment par changement de base. Cela résulte des règles du calcul différentiel. Si $x'^i = \sum_{i'} (V^{-1})^i_{i'} x^{i'}$ ou encore $x^i = \sum_{i'} V^i_{i'} x'^{i'}$, on écrit la différentielle de toute fonction f des x soit en termes des dx soit en termes des dx' selon

$$df = \sum_{i} \left(\frac{\partial}{\partial x^{i}} f \right) dx^{i} = \sum_{i,i'} \left(\frac{\partial}{\partial x^{i}} f \right) V_{i'}^{i} dx'^{i} = \sum_{i'} \left(\frac{\partial}{\partial x'^{i'}} f \right) dx'^{i'}$$

sur lequel on lit

$$\frac{\partial}{\partial x'^{i'}} = \sum_{i} \frac{\partial}{\partial x^{i}} V^{i}_{i'} \,. \tag{5.4}$$

On voit que le vecteur gradient $\vec{\nabla}$ de composantes $\nabla_i = \frac{\partial}{\partial x^i}$ ne se transforme pas comme un vecteur ordinaire \vec{X} . La position "inférieure" de l'indice de ∇_i est justifiée par cette loi de transformation.

On appelle vecteur contravariant un vecteur (comme \vec{X}) se transformant par multiplication à gauche pas la matrice V^{-1} et vecteur covariant un vecteur (comme $\vec{\nabla}$) se transformant par multiplication à droite pas la matrice V. Et on note que les vecteurs contravariants ont des composantes à indices supérieurs, les covariants des composantes à indices inférieurs.

o Transformations des tenseurs

Examinons maintenant comment se transforment les tenseurs introduits au § 5.1. Pour le produit tensoriel $T = \vec{X}_1 \otimes \vec{X}_2$ dont les composantes sont $T^{ij} = x_1^i x_2^j$, la transformation est

$$T'^{ij} = x_1'^i x_2'^j = \sum_{i',j'} (V^{-1})^i{}_{i'} (V^{-1})^j{}_{j'} T^{i'j'} \,,$$

et on parlera de $T = \vec{X}_1 \otimes \vec{X}_2$ ou plus généralement de tout vecteur de $E \otimes E$ comme d'un tenseur de rang 2 deux fois contravariant. À l'inverse, on pourrait aussi considérer des tenseurs $U = (U_{ij})$ deux fois covariants, c'est-à-dire se transformant selon

$$U'_{ij} = \sum_{i',j'} U_{i'j'} V^{i'}_{i} V^{j'}_{j}.$$

Au vu de (5.3) où on fait W = V, on voit que la matrice d'une application linéaire de E dans E se transforme comme un tenseur $A = (A^i{}_j)$ de rang 2 (= à deux indices) une fois contravariant et une fois covariant

$$A'^{i}_{j} = \sum_{i',j'} (V^{-1})^{i}_{i'} A^{i'}_{j'} V^{j'}_{j}.$$

Plus généralement on pourra envisager des tenseurs de rang plus élevé, "mixtes" avec deux types d'indices $T^{i_1\cdots i_r}_{j_1\cdots j_s}$, c'est-à-dire r fois contravariants et s fois covariants, se transformant donc selon

$$T_{j_1\cdots j_s}^{i_1\cdots i_r} = \sum_{\substack{i'_1,\cdots,i'_r\\j'_1,\cdots,j'_s}} \prod_{a=1}^r (V^{-1})^{i_a}_{i'_a} T_{j'_1\cdots j'_s}^{i'_1\cdots i'_r} \prod_{b=1}^s V_{j_b}^{j'_b}.$$

Ces formules peuvent sembler compliquées avec leurs multiples indices et sommations, mais leur structure est très simple à comprendre et à mémoriser, grâce aux règles sur les positions des indices. Par commodité, on parlera de T comme d'un tenseur de rang (r,s).

L'étude des tenseurs et de leurs lois de transformation sont importantes en physique : en mécanique et électricité, voir paragraphe suivant, et aussi en Relativité Restreinte et plus encore en Relativité Générale (la théorie de la gravitation d'Einstein).

5.3. Contraction de deux tenseurs. Convention d'Einstein

Supposons qu'on s'est donné deux tenseurs T et U, de rangs (2,0) et (0,1), dotés de composantes T^{ij} et $U_k,\,i,j,k=1,\ldots,n$. (U est donc un vecteur covariant.) On définit alors le nouveau tenseur T.U, contracté de T et U sur l'indice j par

$$(T.U)^i = \sum_j T^{ij} U_j \tag{5.5}$$

et on vérifie immédiatement que T.U se transforme comme un tenseur de rang (1,0) (donc un vecteur contravariant).

Plus généralement si on s'est donné deux tenseurs T et U, de rangs (r_1, s_1) et (r_2, s_2) , on peut les contracter sur plusieurs paires d'indices, ce qui produit un autre tenseur. Par exemple $\sum_{jk} T_k^{ij} U_{jm}^{k\ell}$ est un tenseur de rang (1, 1).

Une convention de sommation implicite sur les paires d'indices répétés ("convention d'Einstein") permet de simplifier notablement les écritures. Ainsi on récrit simplement (5.5) sous la forme $(T.U)^i = T^{ij}U_j$, la sommation sur j étant implicite.

5.4. Repères orthonormés

Anticipant sur ce qui sera étudié au Chap. 5, mais en se basant sur les notions introduites en géométrie euclidienne et en mécanique, on suppose qu'on se restreint à des bases (ou des repères) orthonormées, telles que les produits scalaires des vecteurs de la base satisfassent $\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$. Les changements de base autorisés sont alors réalisés par des matrices V qui ont la propriété que

$$\delta_{ij} = \vec{e}'_i \cdot \vec{e}'_j = \sum_{i',j'} \vec{e}_{i'} \cdot \vec{e}_{j'} V_i^{i'} V_j^{j'} = \sum_{i',j'} \delta_{i'j'} V_i^{i'} V_j^{j'} = \sum_{i'} V_i^{i'} V_j^{i'} (V^T)^i_{i'} V_j^{i'} = (V^T \cdot V)^i_{j},$$

soit $V^T.V = \mathbb{I}$, ou encore $V^{-1} = V^T$. On appelle matrice orthogonale une matrice satisfaisant cette propriété. Pour des changements de base par une matrice orthogonale, on voit aisément qu'il n'y a plus lieu de distinguer tenseurs covariants et contravariants : ainsi pour un tenseur contravariant

$$x'^{i} = \sum_{i'} (V^{-1})^{i}{}_{i'} x^{i'} = \sum_{i'} (V^{T})^{i}{}_{i'} x^{i'} = \sum_{i'} x^{i'} (V)^{i'}{}_{i}$$

et (x^i) se transforme aussi comme un vecteur covariant (pour ces transformations orthogonales). On dit familièrement qu'on peut "monter et descendre" librement les indices des tenseurs, dans le cadre de la géométrie euclidienne, et à condition de travailler dans un repère orthonormé. (Noter qu'il n'en est plus de même dans la géométrie "pseudo-euclidienne" de l'espace de Minkowski de la Relativité Restreinte, où monter ou descendre un indice peut s'accompagner d'un changement de signe...)

- 5.5. Exemples physiques. Tenseur d'inertie. Tenseur de conductivité. Tenseurs des déformations et des contraintes. Tenseur piézoélectrique
- 1. En Électricité, dans l'étude du transport des charges électriques dans un conducteur soumis à un champ électrique, on écrit une loi de proportionnalité du vecteur courant \vec{j} au champ électrique \vec{E} , $\vec{j} = \sigma \vec{E}$, σ étant la conductivité. (Cette loi équivaut à la célèbre loi d'Ohm entre intensité, tension appliquée et résistance.) Mais cette loi n'est valable que dans un milieu isotrope dans lequel \vec{j} et \vec{E} sont colinéaires. En général, dans un milieu non isotrope, on doit seulement supposer que la relation $\vec{E} \mapsto \vec{j}$ est une application linéaire. (Cela est une approximation physique justifiée aux faibles champs. Des champs élevés peuvent donner lieu à des phénomènes non linéaires.) On écrit donc $\vec{j} = \underline{\sigma}\vec{E}$ où la notation $\underline{\sigma}$ souligne (c'est le cas de le dire) le caractère tensoriel de la conductivité. Le tenseur de conductivité $\underline{\sigma}$ est un tenseur de rang 2, une fois contravariant, une fois covariant, c'est-à-dire une matrice! ce qui est naturel puisqu'il exprime la transformation linéaire du vecteur \vec{E} en le vecteur \vec{j} .
- 2. En Mécanique, considérons un solide indéformable dont on connaît la densité de masse volumique $\rho(\vec{r})$. On définit le tenseur d'inertie

$$I_{ij} = \int d^3r \rho(\vec{r}) \left(r^2 \delta_{ij} - r_i r_j\right). \tag{5.6}$$

Ce tenseur joue un grand rôle dans la description des mouvements de rotation du solide et dans la relation entre le vecteur de rotation instantanée et le moment angulaire. On y reviendra en détail au Chap. 5.

- 3. Considérons un milieu déformable (solide "mou" ou fluide). Un point donné M est repéré par des coordonnées x^i . Sous l'effet d'une déformation, le point M est transformé en M' et le vecteur déplacement $\vec{u} = \overrightarrow{MM'}$ a des composantes $u^i(x)$. Noter que les u^i varient selon le point M, d'où la dépendance en $x = (x^i)$. Il est préférable de se placer dans la suite dans un repère orthonormé, et on ne distinguera plus dans le reste de ce paragraphe les indices supérieurs et inférieurs. On définit alors le tenseur des déformations $\varepsilon = (\vec{\nabla} \otimes \vec{u})_S$, où le "S" signifie qu'on l'a symétrisé, c'est-à-dire qu'il a pour composantes $\varepsilon_{ij} = \frac{\partial u_i}{\partial x^j} + \frac{\partial u_j}{\partial x^i}$.
- 4. L'étude des milieux déformables fait apparaître un autre tenseur, le tenseur des contraintes σ , qui décrit la distribution des forces selon les directions : σ_{ij} est égal à la i-ème composante de la force exercée sur un petit élément de volume du corps, s'appliquant sur l'élément de surface orthogonale à la direction \hat{n} , $F_i = \sigma_{ij}\hat{n}_j$, voir fig. 4 ; on démontre en étudiant l'équilibre interne du milieu que σ est aussi un tenseur symétrique. Pour de petites déformations, on attend une relation linéaire entre les tenseurs σ et ε . Ces deux tenseurs sont en effet reliés par la loi de Hooke

$$\sigma = \underline{H} \cdot \varepsilon$$

impliquant le tenseur d'élasticité \underline{H} de rang 4, ou si on préfère, $\sigma_{ij} = \sum_{k\ell} H_{ijk\ell} \varepsilon_{k\ell}$, qui est le "contracté" de \underline{H} et ε .

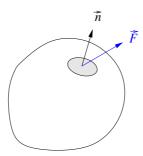


Fig. 4: Tenseur des contraintes. Le vecteur \hat{n} est normal à la surface et dirigé vers l'extérieur de l'élément de volume considéré.

5. Dans l'effet piézoélectrique, une déformation d'un solide fait apparaître une polarisation électrique dans ce solide. (Cet effet était utilisé dans les têtes de lecture des tourne-disques.) À nouveau dans un régime linéaire de petites déformations, le champ \vec{D} (de "déplacement électrique") est fonction linéaire du tenseur de déformation, la relation impliquant un tenseur

$$\vec{D} = \underline{\gamma} \cdot \varepsilon$$

ou encore, en composantes

$$D^i = \sum_{j,k} \gamma^{ijk} \varepsilon_{jk}$$

avec un nouveau tenseur, le tenseur piézoélectrique $\underline{\gamma}$, cette fois un tenseur de rang 3.

*

Chapitre 2. Déterminants

On a déjà rencontré le déterminant d'une matrice (ou d'un tableau) 2×2

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

On observe que cette expression a des propriétés de linéarité par rapport à chacune de ses colonnes (ou de ses lignes), et d'antisymétrie dans l'échange de ces colonnes (ou lignes). Ce sont ces propriétés que nous allons généraliser pour donner une définition générale d'un déterminant.

1. Rappels sur les permutations de p objets

On considère p objets numérotés de 1 à p: ce peut être p boules, p vecteurs, p fonctions, etc. On s'intéresse à l'ensemble \mathcal{S}_p de leurs permutations, c'est-à-dire à toutes les manières de les ordonner. Il y a p choix possibles pour le premier, (p-1) choix possibles pour le deuxième, etc, 1 choix pour le dernier, donc au total "p factoriel" soit $p! = 1.2. \cdots (p-1).p$ permutations des p objets. Autrement dit l'ensemble \mathcal{S}_p a p! éléments. Si σ est une permutation de \mathcal{S}_p , il est conventionnel de noter $\sigma = \begin{pmatrix} 1 & 2 & \cdots & p \\ \sigma(1) & \sigma(2) & \cdots & \sigma(p) \end{pmatrix}$ en plaçant en vis à vis les (indices des) objets initiaux et ceux de leurs images par la permutation ; on peut aussi noter plus simplement $\sigma = (\sigma(1) & \sigma(2) & \cdots & \sigma(p))$. Par exemple, si p = 3, $(1\,2\,3)$ est la permutation identité, $(2\,3\,1)$ est la permutation circulaire ou cyclique qui fait correspondre 2 à 1, etc, 1 à 3. Cette notation permet de construire aisément la composition $\tau \circ \sigma$ de deux permutations σ et τ de \mathcal{S}_p . (À nouveau l'ordre importe, et la notation $\tau \circ \sigma$ signifie qu'on effectue d'abord σ , puis τ). Il suffit d'écrire les images par τ des images par σ de la permutation initiale. Soit

$$\begin{pmatrix} 1 & 2 & \cdots & p \\ \sigma(1) & \sigma(2) & \cdots & \sigma(p) \\ \tau(\sigma(1)) & \tau(\sigma(2)) & \cdots & \tau(\sigma(p)) \end{pmatrix}$$

Par exemple, toujours avec p=3, si $\sigma=(2\,3\,1)$ et $\tau=(3\,2\,1)=\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$, on calcule $\tau\circ\sigma=(2\,1\,3)$. Vérifier que $\sigma\circ\tau=(1\,3\,2)\neq\tau\circ\sigma$, autrement dit, en général, le produit

de deux permutations n'est pas commutatif. Toute permutation σ a un inverse, qui est la permutation σ^{-1} telle que $\sigma^{-1} \circ \sigma = 1$ la permutation identité. On l'obtient aisément en échangeant les deux lignes de la permutation σ et en réordonnant la première (et son image dans la seconde) par ordre croissant. Exemple, $\sigma = (2\ 3\ 1) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, $\sigma^{-1} = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (3\ 1\ 2)$. On en conclut qu'on a aussi $\sigma \circ \sigma^{-1} = 1$. Dans la suite on abrégera $\tau \circ \sigma$ en $\tau \sigma$.

D'un point de vue mathématique, les permutations forment donc un *groupe*. Il est aisé de vérifier l'associativité du produit.

Parmi les permutations, les plus simples consistent à échanger deux objets, voisins ou non dans l'ordre initial. On parle de transposition pour un tel échange. Par exemple $\tau = (3\,2\,1)$ est la transposition de 1 et 3. Un théorème important qui va nous être très utile dans la suite et que nous admettrons est le suivant

Théorème 1 : Toute permutation σ de S_p peut s'écrire comme un produit de transpositions. Cela peut être fait de multiples façons, mais le nombre entier de transpositions N_{tr} est pair ou impair indépendamment de la manière dont on procède.

Que toute permutation σ puisse être obtenue par produit de transpositions est clair : partant de la permutation initiale, on amène 1 à sa position $\sigma(1)$ par une transposition, puis 2 à sa position, etc. Il est moins évident que la parité de ce nombre de transpositions est fixée.

Esquisse de la preuve. On montre d'abord que toute décomposition de la permutation identité 1 en transpositions en compte toujours un nombre pair. Puis pour une permutation quelconque σ écrite de deux façons comme produit de transpositions $\sigma = t_1 \cdots t_r = t'_1 \cdots t'_s$, on a donc $1 = t_1 \cdots t_r t'_s \cdots t'_1$, (car $t_i t_i = 1$), donc r + s est pair et r et s sont de même parité. c.q.f.d.

Définition : On appelle signature de la permutation σ et on note ϵ_{σ} le nombre $(-1)^{N_{tr}}$. Par le théorème précédent, ce nombre est bien défini, indépendant de la décomposition de σ en transpositions. La signature du produit $\tau\sigma$ est le produit des signatures de σ et τ

$$\epsilon_{\tau\sigma} = \epsilon_{\sigma}\epsilon_{\tau} \tag{1.1}$$

(en donner une preuve simple) et en particulier

$$\epsilon_{\sigma^{-1}} = \epsilon_{\sigma} \tag{1.2}$$

puisque leur produit vaut 1, selon (1.1).

On parle de permutation paire, resp. impaire, pour une permutation de signature +1, resp. -1.

Exemple. $\sigma = (2\,3\,1) = (2\,1\,3)(1\,3\,2)$ mais aussi $= (1\,3\,2)(2\,1\,3)(1\,3\,2)(2\,1\,3), \epsilon_{\sigma} = +1, \sigma$ est paire.

2. Formes multilinéaires. Formes antisymétriques. Fonction déterminant

On a rencontré au chapitre précédent des applications linéaires d'un espace vectoriel E dans un espace E'. Si l'espace E' est l'espace des nombres réels (ou complexes, selon le cas), on parle de "forme linéaire". Définissons maintenant une forme multilinéaire $F(X_1, X_2, \dots, X_p)$ de p vecteurs de E. Parler d'une "forme" signifie que l'application est de E dans l'ensemble des nombres réels (ou complexes) : F prend des valeurs réelles (ou complexes). Dire qu'elle est multilinéaire signifie qu'elle est linéaire dans chacun de ses p "arguments" X_1, X_2, \dots, X_p , autrement dit que

$$\forall \lambda, \mu \in \mathbb{R} , \forall q = 1, \dots, p \quad F(X_1, \dots, \lambda X_q' + \mu X_q'', \dots, X_p)$$

$$= \lambda F(X_1, X_2, \dots, X_q', \dots X_p) + \mu F(X_1, X_2, \dots, X_q'', \dots X_p) .$$

$$(2.1)$$

Exemple : si x_i^j est la j-ième composante du vecteur X_i dans une base donnée, et pour toute permutation $\sigma \in \mathcal{S}_p$, $F(X_1, \dots, X_p) = x_1^{\sigma(1)} x_2^{\sigma(2)} \dots x_p^{\sigma(p)}$ est une forme multilinéaire.

> Formes antisymétriques.

Définition : Une forme $F(X_1, X_2, \dots, X_p)$ de p vecteurs de l'espace E est (complètement) antisymétrique si elle change de signe pour toute transposition de deux indices i et j

$$F(X_1, X_2, \dots, X_i, \dots, X_j, \dots, X_p) = -F(X_1, X_2, \dots, X_j, \dots, X_i, \dots, X_p)$$
 (2.2)

Exemple : la composante sur l'axe des z (par exemple) d'un produit vectoriel de deux vecteurs \vec{V} et \vec{W} de l'espace \mathbb{R}^3 s'écrit $(\vec{V} \wedge \vec{W})_z = V_x W_y - V_y W_x$. Elle est bilinéaire et antisymétrique en \vec{V} et en \vec{W} .

Une conséquence immédiate de l'antisymétrie est que si ${\cal F}$ a deux arguments identiques, elle s'annule

$$F(X_1, X_2, \dots, X_i, \dots, X_i, \dots, X_p) = -F(X_1, X_2, \dots, X_i, \dots, X_i, \dots, X_p) = 0.$$
 (2.3)

L'expression

$$\det A = \det(\mathbf{A}^1, \cdots \mathbf{A}^p) = \sum_{\sigma \in \mathcal{S}_p} \epsilon_{\sigma} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{p\sigma(p)}$$
(2.4)

définit le déterminant d'une matrice A carrée $p \times p$, ou encore le déterminant d'un système de p vecteurs \mathbf{A}^j de \mathbb{R}^p . C'est donc une somme "alternée" (signe ϵ_{σ}) sur toutes les permutations σ de l'ensemble des permutations \mathcal{S}_p . Cette expression a bien les deux propriétés

précédentes: elle est multilinéaire dans les composantes des $\mathbf{A}^j = \{a_{ij}\}$; elle est antisymétrique car si τ est la transposition de i et j,

$$\det(\mathbf{A}^{1}, \cdots, \mathbf{A}^{j}), \cdots, \mathbf{A}^{i}, \cdots, \mathbf{A}^{p}) = \sum_{\sigma \in \mathcal{S}_{p}} \epsilon_{\sigma} a_{1 \sigma(1)} \cdots a_{i \sigma(j)} \cdots a_{j \sigma(i)} \cdots a_{n \sigma(p)}$$

$$= \sum_{\sigma \in \mathcal{S}_{p}} \epsilon_{\sigma} a_{1 \sigma\tau(1)} a_{2 \sigma\tau(2)} \cdots a_{n \sigma\tau(p)}$$

$$= -\sum_{\sigma \in \mathcal{S}_{p}} \epsilon_{\sigma\tau} a_{1 \sigma\tau(1)} a_{2 \sigma\tau(2)} \cdots a_{n \sigma\tau(p)}$$

$$= -\sum_{\sigma' \in \mathcal{S}_{p}} \epsilon_{\sigma'} a_{1 \sigma'(1)} a_{2 \sigma'(2)} \cdots a_{n \sigma'(p)}$$

$$= -\det(\mathbf{A}^{1}, \cdots, \mathbf{A}^{i}, \cdots, \mathbf{A}^{j}, \cdots, \mathbf{A}^{p}) \quad (2.5)$$

où à la troisième ligne on a utilisé le fait que selon (1.1), $\epsilon_{\sigma\tau} = -\epsilon_{\sigma}$ puisque τ est une transposition et donc $\epsilon_{\tau} = -1$, et à la quatrième, que sommer sur tous les σ équivaut à sommer sur tous les $\sigma' = \sigma\tau$.

Le fait qu'il soit équivalent de sommer sur les σ ou les $\sigma\tau$, quelle que soit la permutation τ fixée, signifie que quand σ parcourt tout l'ensemble \mathcal{S}_p , $\sigma\tau$ le parcourt aussi, c'est-à-dire "atteint" chaque élément de \mathcal{S}_p une fois et une seule. En effet, si $\sigma \neq \sigma'$, on a aussi $\sigma\tau \neq \sigma'\tau$, sans quoi, si on avait $\sigma\tau = \sigma'\tau$, en multipliant par la droite par l'inverse de τ , on aurait $\sigma = \sigma'$, ce qui est contradictoire. Les p! permutations $\sigma\tau$ sont donc toutes distinctes, donc atteignent bien toutes les permutations de \mathcal{S}_p , cqfd.

L'équation (2.5) a bien établi que det A est une forme antisymétrique de $\mathbf{A}^1, \dots, \mathbf{A}^p$, c'est-à-dire qu'il change de signe par transposition de deux \mathbf{A}^i et \mathbf{A}^j . L'action d'une permutation quelconque σ des \mathbf{A}^i sur le déterminant s'obtient alors grâce à la décomposition de σ en transpositions

$$\det(\mathbf{A}^{\sigma(1)}, \dots, \mathbf{A}^{\sigma(p)}) = \epsilon_{\sigma} \det(\mathbf{A}^{1}, \dots, \mathbf{A}^{p})$$
(2.6)

puisque chaque transposition change son signe et que $\epsilon_{\sigma} = (-1)^{N_{tr}}$.

3. Propriétés du déterminant

Proposition 1 : Tout déterminant possédant deux colonnes égales ou proportionnelles est nul.

Cela découle de l'antisymétrie, cf la propriété (2.3).

Proposition 2 : det A = 0 si les p vecteurs colonnes \mathbf{A}^j sont linéairement dépendants.

Preuve. Supposons les p vecteurs colonnes \mathbf{A}^j de A linéairement dépendants. On peut toujours supposer que c'est le dernier, \mathbf{A}^p , qui s'exprime comme combinaison linéaire des précédents

$$\mathbf{A}^p = \sum_{j=1}^{p-1} \lambda_j \mathbf{A}^j$$
 .

Grâce à sa multilinéarité le déterminant s'exprime alors comme

$$\det A = \det(\mathbf{A}^1, \mathbf{A}^2, \cdots, \mathbf{A}^p) = \sum_{j=1}^{p-1} \lambda_j \det(\mathbf{A}^1, \mathbf{A}^2, \cdots, \mathbf{A}^{p-1}, \mathbf{A}^j)$$

mais en raison de l'antisymétrie, chaque terme de la somme s'annule puisqu'il contient deux colonnes identiques \mathbf{A}^{j} , cf Proposition 1, donc det A = 0, cqfd.

Corollaire 1 : On ne change pas la valeur d'un déterminant en ajoutant à une colonne donnée une combinaison linéaire des autres colonnes.

Supposons qu'on ajoute à la colonne j une combinaison linéaire des autres colonnes :

$$\det(\mathbf{A}^1, \dots, \mathbf{A}^j + \sum_{j' \neq j} \lambda_{j'} \mathbf{A}^{j'}, \dots, \mathbf{A}^p) = \det A + \sum_{j' \neq j} \lambda_{j'} \det(\mathbf{A}^1, \dots, \mathbf{A}^{j'}, \dots, \mathbf{A}^p) = \det A$$

d'après la Proposition 1.

Proposition 3: $\det A^T = \det A$

Preuve:

$$\det A^{T} = \sum_{\sigma \in \mathcal{S}_{p}} \epsilon_{\sigma} \prod_{i=1}^{p} (A^{T})_{i \sigma(i)} = \sum_{\sigma \in \mathcal{S}_{p}} \epsilon_{\sigma} \prod_{i} a_{\sigma(i) i}$$

$$= \sum_{\sigma \in \mathcal{S}_{p}} \epsilon_{\sigma} \prod_{i} a_{i \sigma^{-1}(i)} = \sum_{\sigma \in \mathcal{S}_{p}} \epsilon_{\sigma^{-1}} \prod_{i} a_{i \sigma^{-1}(i)}$$

$$= \sum_{\sigma' \in \mathcal{S}_{p}} \epsilon_{\sigma'} \prod_{i} a_{i \sigma'(i)} = \det A.$$

où on a utilisé (1.2) ainsi que le fait que sommer sur tous les σ équivaut à sommer sur tous les σ^{-1} , par un argument similaire à celui utilisé plus haut pour la sommation sur $\sigma\tau$.

Puisque la transposition échange les vecteurs colonnes et les vecteurs lignes, il découle de la Proposition 3 que partout où nous avons raisonné sur les vecteurs colonnes, nous aurions pu le faire sur les vecteurs lignes. Les deux propositions 1 et 2 et leur corollaire ont donc une version équivalente, portant sur les lignes:

Proposition 1': Tout déterminant possédant deux lignes égales ou proportionnelles est nul.

Proposition 2': det A = 0 si les p vecteurs lignes \mathbf{A}_i sont linéairement dépendants.

Corollaire 1': On ne change pas la valeur d'un déterminant en ajoutant à une ligne donnée une combinaison linéaire des autres lignes.

Proposition 4: det(A.B) = det A det B

Preuve : Les vecteurs colonnes de C = A.B ont pour composantes $(\mathbf{C}^j)_i = (C)_{ij} = (A.B)_{ij} = \sum_k (\mathbf{A}^k)_i b_{kj}$, ce sont donc des combinaisons multilinéaires des composantes des vecteurs colonnes de A, et on peut écrire

$$\det(A.B) = \det C = \det(\mathbf{C}^1, \dots, \mathbf{C}^p) = \sum_{k_1, k_2, \dots, k_p} \det(\mathbf{A}^{k_1} \mathbf{A}^{k_2} \dots \mathbf{A}^{k_p}) b_{k_1 1} b_{k_2 2} \dots b_{k_p p}$$

mais l'antisymétrie de $\det(\mathbf{A}^{k_1}\mathbf{A}^{k_2}\cdots\mathbf{A}^{k_p})$ nous dit que seules les sommes sur des k_i tous distincts contribuent, c'est-à-dire qu'on somme sur toutes les permutations σ de \mathcal{S}_p , $k_i = \sigma(i)$

$$\det(A.B) = \sum_{\sigma \in \mathcal{S}_p} b_{\sigma(1)} \, _1 b_{\sigma(2)} \, _2 \cdots b_{\sigma(p)} \, _p \det(\mathbf{A}^{\sigma(1)} \mathbf{A}^{\sigma(2)} \cdots \mathbf{A}^{\sigma(p)})$$

$$= \left(\sum_{\sigma \in \mathcal{S}_p} \epsilon_{\sigma} b_{\sigma(1)} \, _1 b_{\sigma(2)} \, _2 \cdots b_{\sigma(p)} \, _p \right) \det(\mathbf{A}^1 \mathbf{A}^2 \cdots \mathbf{A}^p) \quad \text{par } (2.6)$$

$$= \left(\sum_{\sigma \in \mathcal{S}_p} \epsilon_{\sigma} b_{1 \sigma^{-1}(1)} b_{2 \sigma^{-1}(2)} \cdots b_{p \sigma^{-1}(p)} \right) \det(\mathbf{A}^1 \mathbf{A}^2 \cdots \mathbf{A}^p)$$

$$= \det B \det A$$

où on a utilisé une fois encore l'antisymétrie du déterminant pour remettre au prix d'un signe ϵ_{σ} les colonnes de A dans l'ordre standard, et où on a alors fait apparaître le déterminant de B. (Exercice : justifier la dernière égalité.)

On a alors deux corollaires

Corollaire 2 : Si la matrice A est inversible, det A^{-1} det A = 1.

Corollaire 3 : Deux matrices semblables (au sens du Chap.1 (4.4)) ont même déterminant. qui découlent immédiatement de la Proposition 4.

En effet, $\det A^{-1} \det A = \det(A^{-1}.A) = \det \mathbb{I} = 1$ et $\det V^{-1}.A.V = \det V^{-1} \det A \det V = \det A$.

Théorème fondamental : det $A \neq 0$ si et seulement si A est inversible.

Preuve. On a vu plus haut (Th. 3, § 3.5 du chap 1) qu'une matrice A est non inversible si et seulement si ses vecteurs colonnes sont linéairement dépendants, mais alors det A = 0 selon la Proposition 2. Donc (en inversant les propositions) si det $A \neq 0$, A est non singulière (inversible). Réciproquement si A est inversible (ou "régulière"), le Corollaire précédent nous dit que det A. det $A^{-1} = 1$, ce qui indique que det $A \neq 0$. Le théorème est démontré.

* Deux erreurs grossières à ne pas faire!

Attention, il est **faux** d'écrire $\det(\lambda A) \stackrel{?}{=} \lambda \det A$. Le déterminant étant une forme *multilinéaire* de ses vecteurs colonnes (ou lignes) qui sont ici tous multipliés par un même facteur λ , la formule correcte pour un déterminant $p \times p$ est

$$\det(\lambda A) = \lambda^p \det A \ . \tag{3.1}$$

Par ailleurs, alors que le déterminant d'un produit est le produit des déterminants, le déterminant d'une somme de matrices n'admet pas de formule simple : en général, $\det(A+B) \neq \det A + \det B$!

4. Méthodes de calcul

4.1. Calcul direct

La méthode de calcul direct par l'expression (2.4) peut s'utiliser pour les petites valeurs de p.

Exemples. On a rappelé au début de ce chapitre la valeur du déterminant d'une matrice 2×2 . Donnons maintenant celle d'un déterminant 3×3

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$
$$-a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33}$$

où les termes successifs correspondent aux 3 permutations paires (123), (231) et (312), puis aux trois impaires, (132), (321) et (213). Il peut être bon de se rappeler les 6 termes et leur signe sous forme graphique, cf Figure 4.

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Fig. 5: Les 6 termes du déterminant 3×3

4.2. Combinaisons linéaires des colonnes ou des lignes

Par les Corollaires des Propositions 2 et 2' ci-dessus, on ne change pas le déterminant en ajoutant à une ligne (resp. une colonne) une combinaison linéaire des autres lignes (resp. colonnes). Dans certains cas, cela permet d'obtenir (presque) sans calcul la valeur du déterminant.

Exemples:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 5 & 9 \\ 0 & 3 & 9 & 19 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 3 & 10 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1$$

On a successivement retranché la première ligne aux suivantes, puis un multiple par 2 ou 3 de la seconde aux suivantes, puis 3 fois la troisième à la dernière. Le calcul s'achève alors facilement puisque seule la permutation identité contribue et donne le produit des termes diagonaux, soit 1.

Autres exemples:

$$\begin{vmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} = \begin{vmatrix} a+b+c & a+b+c & a+b+c \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix}$$

$$= (a+b+c) \begin{vmatrix} 1 & 1 & 1 \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} = (a+b+c) \begin{vmatrix} 1 & 0 & 0 \\ 2b & -a-b-c & 0 \\ 2c & 0 & -c-a-b \end{vmatrix}$$

$$= (a+b+c)^3$$

où à la première ligne de l'équation on a ajouté les deux dernières lignes de la matrice à sa première ligne, puis à la seconde ligne de l'équation, on a retranché la première colonne des deux suivantes. On a finalement développé le déterminant selon la première ligne (cf ci-dessous § 4.3).

Ou encore

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 5 & 4 & 3 & 2 \\ 2 & 3 & 4 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -5 & -10 & -15 \\ 0 & -6 & -12 & -18 \\ 0 & -1 & -2 & -3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -5 & -10 & -15 \\ 0 & -6 & -12 & -18 \\ 0 & 0 & 0 & 0 \end{vmatrix} = 0 ,$$

pourquoi?

4.3. Développement par rapport à une ligne ou à une colonne. Mineurs

Le déterminant étant une fonction multilinéaire de ses vecteurs lignes ou de ses vecteurs colonnes, c'est-à-dire une fonction *linéaire* de chacun de ses vecteurs lignes ou colonnes, on peut le développer selon les composantes d'une ligne ou d'une colonne. Par exemple le développement par rapport à la *i*-ème ligne est

$$\det A = \sum_{j=1}^{p} a_{ij} A^{ij} \qquad i \text{ fix\'e}$$

$$(4.1)$$

où le cofacteur A^{ij} (attention aux notations!) s'exprime lui-même comme un déterminant, mais d'une matrice $(p-1)\times(p-1)$. On a donc gagné en complexité. La règle est la suivante

$$A^{ij} = (-1)^{i+j} \Delta^{ij} = (-1)^{i+j} \det(\tilde{A}^{(ij)})$$
(4.2)

où $\tilde{A}^{(ij)}$ est la matrice obtenue en ôtant dans A la i-ième ligne et la j-ième colonne et $\Delta^{ij} = \det(\tilde{A}^{(ij)})$ est le $mineur\ d'ordre\ (i,j)$. En combinant ces deux dernières formules, on obtient l'importante relation

$$\det A = \sum_{j=1}^{p} (-1)^{i+j} a_{ij} \Delta^{ij} \qquad i \text{ fix\'e}$$

$$(4.3)$$

ou aussi, en développant par rapport à la j-ième colonne

$$\det A = \sum_{i=1}^{p} (-1)^{i+j} a_{ij} \Delta^{ij} \qquad j \text{ fixé} . \qquad (4.3)'$$

La preuve de (4.3) ou (4.3)' découle de l'application de la définition (2.4) du déterminant. Le cofacteur A^{ij} de a_{ij} provient des permutations σ où $\sigma(i)=j$, et dans la somme sur les permutations σ' des p-1 autres éléments, on vérifie que $\epsilon_{\sigma'}=\epsilon_{\sigma}(-1)^{i+j}$.

Fig. 6: Développement d'un déterminant 4×4 : (a) selon la première ligne, (b) selon la seconde colonne. Dans chaque cas, la ligne et la colonne détruites sont barrées par une ligne en tirets.

Deux exemples de tels développements sont représentés de façon symbolique sur la figure 6, pour une matrice 4×4 . De façon plus explicite, ils se lisent

$$\det A = a_{11}\Delta^{11} - a_{12}\Delta^{12} + a_{13}\Delta^{13} - a_{14}\Delta^{14}$$
$$= -a_{12}\Delta^{12} + a_{22}\Delta^{22} - a_{32}\Delta^{32} + a_{42}\Delta^{42}.$$

On a évidemment tout intérêt à utiliser ce développement par rapport à une ligne (ou à une colonne) qui contient de nombreux zéros.

Exemples

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 0 & 0 \end{vmatrix} = +7 \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = 7(2 \times 6 - 3 \times 5) = -21$$

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 0 \\ 6 & 7 & 0 \end{vmatrix} = +3 \begin{vmatrix} 4 & 5 \\ 6 & 7 \end{vmatrix} = 3(4 \times 7 - 5 \times 6) = -6$$

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 0 & 6 & 7 \\ 8 & 0 & 9 & 0 \\ 10 & 0 & 11 & 0 \end{vmatrix} = -2 \begin{vmatrix} 5 & 6 & 7 \\ 8 & 9 & 0 \\ 10 & 11 & 0 \end{vmatrix} = -2 \times 7 \begin{vmatrix} 8 & 9 \\ 10 & 11 \end{vmatrix} = 2 \times 7 \times 2 = 28$$

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 0 & 6 & 0 \\ 7 & 8 & 9 & 10 \\ 11 & 0 & 12 & 0 \end{vmatrix} = -4 \begin{vmatrix} 5 & 0 & 6 \\ 7 & 8 & 9 \\ 11 & 0 & 12 \end{vmatrix} - 10 \begin{vmatrix} 1 & 2 & 3 \\ 5 & 0 & 6 \\ 11 & 0 & 12 \end{vmatrix} = -4 \times 8 \begin{vmatrix} 5 & 6 \\ 11 & 12 \end{vmatrix} + 10 \times 2 \begin{vmatrix} 5 & 6 \\ 11 & 12 \end{vmatrix} = 72$$

Il y a évidemment des façons variées de calculer ces déterminants.

4.4. Méthode du pivot de Gauss

La méthode du pivot de Gauss est une conséquence directe des propriétés du \S 4.2. Elle consiste à faire apparaître n-1 zéros sur une colonne (ou une ligne) d'un déterminant d'ordre n par combinaisons linéaires de lignes (ou de colonnes). Le déterminant est alors égal au produit du seul élément restant dans cette colonne (ou ligne), le pivot, par son cofacteur. On passe ainsi d'un déterminant d'ordre n à un déterminant d'ordre n-1.

En pratique, on cherche une ligne ou une colonne contenant des nombres simples (pas trop grands, pour simplifier les calculs), de préférence des zéros, et si possible un 1. Supposons par exemple que le déterminant a un élément $a_{ij} = 1$. On ne modifie pas le déterminant en retranchant à la k-ième ligne a_{kj} fois la i-ème, pour tout $k \neq i$: le nouveau déterminant a une j-ième colonne faite de 0 sauf à la i-ème ligne. On peut alors développer selon cette colonne, se ramener à un déterminant d'ordre n-1, puis itérer l'algorithme.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & & \ddots & & \\ a_{i1} & a_{i2} & \cdots & a_{ij} = 1 & \cdots & a_{in} \\ \vdots & & & \ddots & & \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} - a_{1j}a_{i1} & a_{12} - a_{1j}a_{i2} & \cdots & 0 & \cdots & a_{1n} - a_{1j}a_{in} \\ a_{21} - a_{2j}a_{i1} & a_{22} - a_{2j}a_{i2} & \cdots & 0 & \cdots & a_{2n} - a_{2j}a_{in} \\ \vdots & & & \ddots & \vdots & & \\ a_{n1} & a_{n2} & \cdots & 1 & \cdots & a_{in} \\ \vdots & & & \vdots & \ddots & \\ a_{n1} - a_{nj}a_{i1} & a_{n2} - a_{nj}a_{i2} & \cdots & 0 & \cdots & a_{nn} - a_{nj}a_{in} \end{vmatrix}$$

$$= (-1)^{i+j} \begin{vmatrix} a_{11} - a_{1j}a_{i1} & a_{12} - a_{1j}a_{i2} & \cdots & a_{1n} - a_{1j}a_{in} \\ a_{21} - a_{2j}a_{i1} & a_{22} - a_{2j}a_{i2} & \cdots & a_{2n} - a_{2j}a_{in} \\ \vdots & & \ddots & \vdots \\ a_{n1} - a_{nj}a_{i1} & a_{n2} - a_{nj}a_{i2} & \cdots & a_{nn} - a_{nj}a_{in} \end{vmatrix}$$

où la *i*-ème ligne et la *j*-ème colonne ont été ôtées pour fabriquer le dernier déterminant.

Si le déterminant de départ ne possède aucun élément égal à 1, on choisit un élément non nul a_{ij} , on écrit la j-ième colonne comme $\mathbf{A}^j = a_{ij} \frac{\mathbf{A}^j}{a_{ij}}$, on factorise a_{ij} hors du déterminant, et on est alors ramené au cas précédent, avec une j-ième colonne ayant un 1 à la i-ème ligne.

Bien sûr, partout dans ce qui précède, les mots "ligne" et "colonne" peuvent être intervertis.

Exemple: En utilisant cet algorithme, calculer le déterminant

$$D = \begin{vmatrix} 2 & -1 & 3 & 5 \\ 3 & 1 & -1 & 5 \\ 5 & -2 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{vmatrix} .$$

(Réponse : 60)

4.5. Calcul de l'inverse d'une matrice

Dans la formule (4.1), remplaçons au membre de droite a_{ij} par les éléments d'une autre ligne $a_{i'j}$ de A. L'expression obtenue est, toujours selon (4.1), le développement d'un déterminant dans lequel la ligne i' apparaît deux fois, et qui est donc nul (cf Proposition 2'). On a donc $\sum_{j=1}^{p} a_{i'j} A^{ij} = 0$, et cette relation et (4.1) peuvent être mises sous une forme unique

$$\sum_{j=1}^{p} a_{i'j} A^{ij} = \delta_{ii'} \det A . {(4.5)}$$

Si A est inversible, det $A \neq 0$ et on peut récrire cette identité sous la forme

$$\sum_{i=1}^{p} a_{i'j} \frac{1}{\det A} A^{ij} = \delta_{ii'}$$

qu'on compare à celle satisfaite par la matrice inverse A^{-1}

$$\sum_{j=1}^{p} a_{i'j} (A^{-1})_{ji} = \delta_{ii'} .$$

L'identité (4.5) nous fournit donc une formule compacte pour la matrice inverse

$$A^{-1} = \frac{1}{\det A} \left(\operatorname{Cof} A \right)^{T} , \qquad (4.6)$$

où la comatrice Cof A est la matrice d'élément i,j égal au cofacteur A^{ij} , et $(\text{Cof }A)^T$ est sa transposée.

Exemple. Calculons par cette formule la matrice inverse de la matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 3 & 6 & 10 \end{pmatrix}$. On vérifie que det A = -2. On calcule alors la comatrice $\operatorname{Cof} A = \begin{pmatrix} 0 & -5 & 3 \\ -4 & 7 & -3 \\ 2 & -3 & 1 \end{pmatrix}$, et la matrice inverse est donc $A^{-1} = -\frac{1}{2}(\operatorname{Cof} A)^T = \begin{pmatrix} 0 & 2 & -1 \\ \frac{5}{2} & -\frac{7}{2} & \frac{3}{2} \\ -\frac{3}{2} & \frac{3}{2} & -\frac{1}{2} \end{pmatrix}.$

5. Applications des déterminants

5.1. Critère d'indépendance linéaire

Une application très fréquente et très utile des déterminants découle des propositions 2 et 2' et du Théorème fondamental du \S 3 :

Proposition 5 : Dans un espace de dimension n, un système de n vecteurs est linéairement indépendant ssi le déterminant de leurs n composantes dans une base arbitraire est non nul. Dans un espace de dimension n, un système de p < n vecteurs est linéairement indépendant ssi l'un des déterminants $p \times p$ formés avec p de leurs composantes dans une base arbitraire est non nul.

Dire que cela constitue une condition nécessaire et suffisante signifie qu'inversement, la nullité du déterminant $n \times n$ dans le premier cas, de tous les déterminants $p \times p$ dans le second, assure que les vecteurs sont linéairement dépendants.

Exemples : Montrer que les vecteurs de composantes (1,1,3), (2,2,3) et (0,0,1) sont linéairement dépendants dans \mathbb{R}^3 . Qu'en est-il des vecteurs (1,1,3) et (2,2,3) ?

5.2. Équation d'une droite de \mathbb{R}^2 , d'un plan de \mathbb{R}^3

Comme application du critère précédent, cherchons l'équation d'une droite du plan \mathbb{R}^2 , ou d'un plan de l'espace \mathbb{R}^3 .

Soient M_1 et M_2 deux points du plan, de coordonnées (x_1, y_1) et (x_2, y_2) dans un repère donné (O, \vec{i}, \vec{j}) . Autrement dit, les vecteurs $\overrightarrow{OM_1}$ et $\overrightarrow{OM_2}$ s'écrivent respectivement

$$\overrightarrow{OM_1} = x_1 \overrightarrow{i} + y_1 \overrightarrow{j} , \qquad \overrightarrow{OM_2} = x_2 \overrightarrow{i} + y_2 \overrightarrow{j} .$$

Un point M du plan, de coordonnées (x, y), appartient à la droite M_1M_2 ssi les vecteurs $\overline{M_1M}$ et $\overline{M_1M_2}$ sont colinéaires, c'est-à-dire linéairement dépendants, donc selon le critère de la Proposition 5,

$$\det(\overrightarrow{M_1M}, \overrightarrow{M_1M_2}) = \begin{vmatrix} x - x_1 & x_2 - x_1 \\ y - y_1 & y_2 - y_1 \end{vmatrix} = 0$$
 (5.1)

soit $(x-x_1)(y_2-y_1)-(y-y_1)(x_2-x_1)=0$ ou encore, si $(x_2-x_1)(y_2-y_1)\neq 0$,

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1}$$

forme familière de l'équation d'une droite passant par les deux points M_1 et M_2 . Si $(x_2 - x_1)(y_2 - y_1) = 0$, par exemple $y_2 = y_1$ (mais $x_2 - x_1 \neq 0$!), l'équation se réduit à $y = y_1$, droite parallèle à l'axe des x.

Le raisonnement se transpose facilement à trois dimensions : soient M_1 , M_2 et M_3 trois points de l'espace, de coordonnées (x_i, y_i, z_i) , i = 1, 2, 3, dans un repère $\vec{i}, \vec{j}, \vec{k}$. Un point M de coordonnées (x, y, z) appartient au plan passant par M_1 , M_2 , M_3 ssi les trois vecteurs $\overrightarrow{M_1M}$, $\overrightarrow{M_1M_2}$ et $\overrightarrow{M_1M_3}$ sont coplanaires, c'est-à-dire linéairement dépendants, donc selon le critère de la Proposition 5,

$$\det(\overrightarrow{M_1M}, \overrightarrow{M_1M_2}, \overrightarrow{M_1M_3}) = \begin{vmatrix} x - x_1 & x_2 - x_1 & x_3 - x_1 \\ y - y_1 & y_2 - y_1 & y_3 - y_1 \\ z - z_1 & z_2 - z_1 & z_3 - z_1 \end{vmatrix} = 0 , \qquad (5.2)$$

soit, en développant selon la première colonne et en modifiant légèrement l'ordre des lignes

$$(x-x_1) \begin{vmatrix} y_2 - y_1 & y_3 - y_1 \\ z_2 - z_1 & z_3 - z_1 \end{vmatrix} + (y-y_1) \begin{vmatrix} z_2 - z_1 & z_3 - z_1 \\ x_2 - x_1 & x_3 - x_1 \end{vmatrix} + (z-z_1) \begin{vmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{vmatrix} = 0$$

qui est l'équation du plan passant par les trois points M_1 , M_2 et M_3 . Bien noter la structure de cette équation : le coefficient de $(x-x_1)$ est le mineur associé dans le déterminant (5.2),

les deux autres termes s'en déduisant par permutation cyclique $x \mapsto y \mapsto z \mapsto x$. On vérifie aisément, par exemple en soustrayant la deuxième colonne de (5.1) ou de (5.2) à toutes les autres, que cette équation est indépendante du choix du point M_1 plutôt que M_2 ou M_3 , comme il se doit.

Exemple. Soient trois points O, origine des coordonnées, P_1 de coordonnées x_1, y_1 et $P_2: x_2, y_2$. Le plan passant par O, P_1 , P_2 , (on dit aussi "sous-tendu" par les vecteurs $\overrightarrow{OP_1}$ et $\overrightarrow{OP_2}$), a pour équation

$$\begin{vmatrix} x & x_1 & x_2 \\ y & y_1 & y_2 \\ z & z_1 & z_2 \end{vmatrix} = 0$$
, soit $x(y_1z_2 - y_2z_1)$ + perm. cycl. = 0. Ainsi, dans l'exemple du paragraphe précédent,

l'équation du plan sous-tendu par les vecteurs (1,1,3) et (2,2,3), est -3(x-y)=0. C'est donc un plan passant par l'axe des z, d'équation x=y, c'est le plan bissecteur des vecteurs \vec{i} et \vec{j} .

5.3. Wronskien

Les déterminants sont utiles aussi pour discuter l'(in)dépendance linéaire de fonctions. Le wronskien est le déterminant de n fonctions et de leurs n-1 premières dérivées. Une condition suffisante de l'indépendance de ces fonctions est la non-nullité de cette fonction wronskien, voir TD.

5.4. Interprétation géométrique du déterminant. Jacobien

Anticipant un peu sur la discussion du Chap. 5, considérons l'espace euclidien \mathbb{R}^n et supposons qu'on y a construit un repère (une base) $\vec{e_i}$ orthonormé, c'est-à-dire constitué de vecteurs deux à deux orthogonaux et normés, ce qu'on résume dans une formule unique

$$\vec{e}_i.\vec{e}_j = \delta_{ij} . ag{5.3}$$

Soit $(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_p)$ un système de p vecteurs avec $p \leq n$. On note x_{ij} leurs composantes dans la base \vec{e}_i

$$\vec{X}_i = \sum_j x_{ij} \vec{e}_j \ .$$

On définit alors le parallélépipè de généralisé, ou parallélotope, $\mathcal P$ construit sur ces vecteurs, comme l'ensemble des points de $\mathbb R^n$

$$\mathcal{P}: \qquad \vec{\xi} = \sum_{i=1}^{p} \xi_i \vec{X}_i, \qquad 0 \le \xi_i \le 1 \ .$$
 (5.4)

On démontre (et nous admettrons) que le volume de ce parallélotope est donné par

$$vol(\mathcal{P}) = |\det(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_p)| = |\det(x_{ij})|. \tag{5.5}$$

- Il est nul si les vecteurs $\vec{X}_1, \vec{X}_2, \cdots, \vec{X}_p$ sont linéairement dépendants. Cela correspond bien à l'image qu'on se fait de cette situation, où (au moins) un des vecteurs \vec{X}_i est combinaison linéaire des autres, et appartient donc au sous-espace qu'engendrent ces derniers : le parallélotope est aplati, donc de volume nul ;
- si on dilate toutes les longueurs des \vec{X}_i par un même facteur réel λ , selon la formule (3.1), le volume est multiplié par λ^p , conformément à l'analyse dimensionnelle.

Si p=n, le même déterminant (5.5), mais sans valeur absolue, définit le volume algébrique du polytope \mathcal{P} à n dimensions. Son signe est positif ou négatif selon que le système des n vecteurs $\vec{X}_1, \vec{X}_2, \cdots, \vec{X}_n$ forme un repère orienté positivement ou négativement (par rapport au repère initial $\{\vec{e}_i\}$).

Dans le même ordre d'idées, considérons un changement de variables $\vec{x} \mapsto \vec{x}'$ dans une intégrale à n dimensions $I = \int_V \mathrm{d}^n x F(\vec{x})$. On démontre que l'élément de volume infinitésimal $\mathrm{d}^n x$ est relié à celui $\mathrm{d}^n x'$ dans les nouvelles coordonnées par

$$d^{n}x = \left| \det \left(\frac{\partial x_{i}}{\partial x'_{i}} \right) \right| d^{n}x' =: |J| d^{n}x'.$$

On appelle jacobien le déterminant des gradients du changement de variables. La valeur absolue |J| du jacobien est donc le facteur à insérer dans un changement de variables.

Chapitre 3. Systèmes linéaires d'équations algébriques de 1er degré

On rencontre très souvent en mathématiques ou en physique le problème de résoudre un système d'équations linéaires (du premier degré) couplées. Nous allons voir que les méthodes matricielles et déterminantales sont très efficaces pour traiter ces questions.

1. Position du problème

1.1. Système d'équations considéré comme un problème vectoriel

Considérons le système de p équations à n inconnues x_1, x_2, \dots, x_n :

$$(S) \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ & \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pn}x_n &= b_p \end{cases}$$

$$(1.1)$$

avec des coefficients a_{ij} et b_j (réels) donnés et dont on cherche les solutions x_j .

Considérons les n vecteurs colonnes à p composantes

$$V_{1} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, V_{2} = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, V_{n} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix}$$
(1.2)

et soit

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} .$$

Le système (S) de l'équation (1.1) se récrit comme une équation vectorielle

$$x_1V_1 + x_2V_2 + \cdots + x_nV_n = B$$
.

Pour que le système (1.1) soit "possible", c'est-à-dire admette des solutions en x, il faut et il suffit que B appartienne à l'espace vectoriel V, sous-espace de \mathbb{R}^p , engendré par les n vecteurs V_1, V_2, \dots, V_n .

Donc deux cas se présentent

- (1) B n'appartient pas à ce sous-espace \mathcal{V} : le système est "impossible", il n'admet pas de solution en x:
- (2) B appartient à \mathcal{V} . Soit r le rang du système de vecteurs V_1, V_2, \dots, V_n . Rappelons que ce rang obéit à deux inégalités : $r \leq p$ et $r \leq n$, cf équ. (2.3) du Chap.1 (dans lequel les notations p et n sont échangées). Quitte à les renuméroter, on peut toujours supposer que les r premiers V_1, V_2, \dots, V_r sont indépendants. Ils engendrent l'espace \mathcal{V} et en particulier V_{r+1}, \dots, V_n en sont des combinaisons linéaires. Alors quels que soient $x_{r+1}, x_{r+2}, \dots, x_n$, le vecteur

$$B' = B - x_{r+1}V_{r+1} - \cdots + x_nV_n$$

appartient à \mathcal{V} , et on peut trouver r nombres réels x_1, x_2, \cdots, x_r tels que

$$x_1V_1 + x_2V_2 + \cdots + x_rV_r = B - x_{r+1}V_{r+1} - \cdots + x_nV_n$$
.

On en conclut que

Théorème 1 : Si le rang du système de vecteurs V_1, V_2, \dots, V_n est r et si $B \in \mathcal{V}$, le système (1.1) admet des solutions dépendant de n-r paramètres.

Exemple : soit le système

$$\begin{cases} x_1 + x_2 &= 1\\ x_1 + x_3 &= 1\\ x_2 - x_3 &= 0 \end{cases}$$

Les trois vecteurs $V_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $V_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ et $V_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ sont linéairement dépendants et forment un système de rang 2 puisque $V_1 = V_2 + V_3$. Le vecteur $B = V_1$ appartient bien sûr au sous-espace \mathcal{V} . La solution dépend de 3-2=1 paramètre arbitraire, par exemple x_3 , soit

$$x_1 = 1 - x_3$$
 $x_2 = x_3$.

1.2. Systèmes d'équations homogènes

On appelle système homogène un système de la forme (1.1) dans lequel le second membre s'annule, $B = \{b_i\} = 0$

$$\begin{cases}
 a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= 0 \\
 a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= 0 \\
 &\vdots \\
 a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= 0
\end{cases}$$
(1.3)

Dans ce cas, il est clair que l'on est toujours dans le cas (2) de la discussion précédente : le vecteur 0 appartient toujours à l'espace \mathcal{V} engendré par les vecteurs V_1, V_2, \dots, V_n .

On note aussi que les solutions $\{x_j\}$ forment elles-mêmes un sous-espace vectoriel de \mathbb{R}^n . En effet si $x^{(1)}, x^{(2)}, \dots x^{(q)}$ sont q solutions de (1.3), où chaque $x^{(k)}$ est un vecteur de \mathbb{R}^n à n composantes $x_j^{(k)}$, toute combinaison linéaire $\sum_{k=1}^q \lambda_k x^{(k)}$ est aussi solution.

Si le rang des n vecteurs V_1, V_2, \cdots, V_n est n, c'est-à-dire s'ils sont linéairement indépendants, la seule façon d'écrire 0 comme combinaison linéaire $\sum_j x_j V_j$, donc la seule solution du système, est la solution triviale $x_j = 0, \ \forall j = 1, \cdots, n$.

Si le rang des V_j est r < n, donc que les vecteurs V_j sont linéairement dépendants, on a comme au paragraphe précédent une solution dépendant de n-r paramètres indépendants. En effet, supposant à nouveau les r premiers V_1, V_2, \dots, V_r indépendants, pour tout choix de x_{r+1}, \dots, x_n , on peut trouver x_1, \dots, x_r tels que

$$x_1V_1 + x_2V_2 + \dots + x_rV_r = -x_{r+1}V_{r+1} - \dots + x_nV_n$$

c'est-à-dire trouver une solution de (1.3).

Bien entendu cette discussion n'est qu'un cas particulier de celle du \S 1.1 quand le second membre B=0. Inversement, étant données deux solutions du système (1.1), leur différence satisfait le système (1.3).

1.3. Interprétation matricielle

Considérons le système S d'un autre point de vue, comme représentant une application linéaire d'un espace dans un autre. E désigne l'espace vectoriel \mathbb{R}^n , doté d'une base e_i , F l'espace \mathbb{R}^p , avec une base f_i . La matrice

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{pmatrix}$$

peut être considérée comme la matrice d'une application linéaire \mathcal{A} de E dans F. Résoudre le système (1.1) équivaut à trouver les vecteurs \vec{X} représentés dans la base \vec{e}_i par le vecteur colonne

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in E = \mathbb{R}^n$$

dont l'image par \mathcal{A} est le vecteur $\vec{B} \in F$:

$$\mathcal{A}(\vec{X}) = \vec{B}$$
 ou encore $AX = B$.

À nouveau, deux cas se présentent

- (1) $\vec{B} \notin \mathcal{A}(E)$, c'est-à-dire l'image de E par \mathcal{A} ne contient pas \vec{B} : le système est "impossible";
- (2) $\vec{B} \in \mathcal{A}(E)$. Il existe au moins un vecteur \vec{X}_0 tel que $\mathcal{A}(\vec{X}_0) = \vec{B}$, ou encore $AX_0 = B$. Alors tout autre X solution de (1.1) est tel que (par soustraction des deux équations satisfaites par X et X_0)

$$\mathcal{A}(\vec{X} - \vec{X}_0) = 0$$
 ou encore $A(X - X_0) = 0$.

On obtient donc toutes les solutions à partir de l'une d'entre elles \vec{X}_0 en lui ajoutant un vecteur quelconque \vec{Y} du "noyau", c'est-à-dire satisfaisant $\mathcal{A}(\vec{Y})=0$. Si l'application \mathcal{A} a pour rang r, c'est-à-dire si la matrice A a le rang r, ce noyau a pour dimension n-r (cf Théorème du § 3.2 au chap. 1). La solution dépend donc de n-r paramètres arbitraires.

On a bien retrouvé les conclusions des deux sections précédentes.

2. Rang d'un système, déterminant principal

Une question importante est donc de déterminer le rang r d'un système de vecteurs V_1, V_2, \dots, V_n , ou le rang de la matrice A (dont ces vecteurs sont les vecteurs-colonnes). On se rappelle la Proposition 5 du § 5 du chapitre 2, d'où découle la méthode suivante. À partir de la matrice $A = (a_{ij})$, on forme des déterminants $q \times q$ en ne gardant que les éléments communs à q lignes et q colonnes et on cherche la plus grande valeur de q telle qu'un tel déterminant soit non nul. En pratique, on commence par donner à q la valeur la plus grande possible $(q = \inf(n, p))$, puis si tous ces déterminants sont nuls, on passe à l'entier immédiatement inférieur, etc. On s'arrête dès qu'on a trouvé un déterminant non nul, on appelle un tel déterminant d déterminant d principal de la matrice (ou du système). Le rang r cherché est la dimension de ce déterminant principal. Noter que le déterminant principal n'est en général pas unique.

Exemple. Soit la matrice

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 3 & 4 \\ 2 & 2 & 4 & 5 \end{pmatrix}$$

Le rang de cette matrice 3×4 est au plus 3. Les trois vecteurs lignes sont évidemment linéairement dépendants, puisque la troisième ligne est la somme des deux précédentes (et donc tout sous-déterminant de taille 3×3 est nul). Le rang est donc au plus 2. Calculons les sous-déterminants de taille 2×2 . Le premier en haut à gauche, $\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}$, est nul mais le suivant

$$\begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} = 2$$

est non nul, c'est un déterminant principal de A et le rang de A est bien 2.

Les inconnues x_i correspondant aux colonnes du déterminant principal sont appelées inconnues principales. Dans l'exemple précédent, ce seraient x_1 et x_3 , mais on peut toujours, quitte à renuméroter les variables et à permuter les colonnes de A, supposer que ce sont les r premières x_1, x_2, \dots, x_r . De même on appelle équations principales les équations correspondant aux lignes du déterminant principal, et on peut supposer que ce sont les r premières du système.

3. Discussion et résolution. Systèmes de Cramer

3.1. $p = r \le n$

* Si n = r = p, le système est dit de Cramer. La matrice A des coefficients est régulière (rang = dimension) donc inversible. Le système admet une solution **unique** qu'on écrit sous la forme

$$X = A^{-1}B .$$

Mais on se rappelle que $A^{-1} = \frac{1}{\det A} \mathrm{Cof}\,A^T$, d'où on tire les formules de Cramer

$$x_{j} = \frac{b_{1}A^{1j} + b_{2}A^{2j} + \dots + b_{n}A^{nj}}{\det A} .$$
(3.1)

En utilisant la formule (4.1) du chap. 2 (ou plutôt son analogue pour un développement par rapport à la j-ème colonne), on voit que le numérateur n'est autre que le déterminant de la matrice A où on a substitué la colonne des b_i à la colonne des a_{ij} , c'est-à-dire le vecteur-colonne B au vecteur-colonne V_j (cf. équ. (1.2))

$$x_j = \frac{\det(V_1, V_2, \cdots, \overset{\downarrow^j}{B}, \cdots, V_n)}{\det(V_1, V_2, \cdots, V_n)} . \tag{3.2}$$

* Si p = r < n, on a $r = \dim \mathcal{A}(E) = \dim \mathbb{R}^p = p$ qui nous assure que tout $\vec{B} \in \mathcal{V} = \mathcal{A}(E)$ et le système est toujours possible. Par ailleurs r < n, on a plus d'inconnues que d'équations. On choisit alors arbitrairement les n - r inconnues non principales et on résout le système de Cramer sur les r inconnues principales. La solution a donc une indétermination d'ordre n - r.

Exemples 1. Soit le système $\begin{cases} 2x_1 + 3x_2 = 1 \\ x_1 + 2x_2 = 2 \end{cases}$. Il a n = p = 2. Son rang est 2 car le déterminant $\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} = 1$. C'est donc un système de Cramer, la solution est donnée par (3.1) ou (3.2).

$$x_1 = \frac{\begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix}} = -4, \quad x_2 = \frac{\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix}} = 3.$$

On peut aussi l'écrire sous forme matricielle

$$AX = B$$
, $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

donc

$$X = A^{-1}B = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -4 \\ 3 \end{pmatrix} .$$

2. Soit maintenant le système $\begin{cases} 2x_1+3x_2+x_3&=1\\ x_1+2x_2+x_3&=2 \end{cases}. \text{ On a } n=3,\ p=r=2,\ \text{et en}$ récrivant $\begin{cases} 2x_1+3x_2&=1-x_3\\ x_1+2x_2&=2-x_3 \end{cases} \text{ on se ramène à un système de Cramer en } x_1 \text{ et } x_2 \text{ ; la solution dépend de } n-r=1 \text{ paramètre } x_3:\ x_1=-4+x_3,\ x_2=3-x_3.$

3.2. r < p. Déterminants caractéristiques

Dans ce cas, le nombre d'équations est supérieur au rang de la matrice. Supposons comme plus haut que les inconnues principales et les équations principales sont les r premières. Le système formé par ces r équations principales dans les r inconnues principales est un système de Cramer, avec les n-r inconnues non principales formant des paramètres aux seconds membres. Pour que le système (S) de (1.1) ait une solution il faut et il suffit que chacune des équations non principales soit encore satisfaite par la solution qu'on vient de

trouver pour les équations principales. Autrement dit, il faut et il suffit que pour tout q vérifiant $r < q \le p$, le système

$$(S_q) \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \vdots & & & \\ a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rn}x_n &= b_r \\ a_{q1}x_1 + a_{q2}x_2 + \dots + a_{qn}x_n &= b_q \end{cases}$$

$$(3.3)$$

soit possible, ou encore que le vecteur $\vec{b} = (b_1, b_2, \dots, b_r, b_q)$ appartienne au sous-espace engendré par les $\vec{v}_i = (a_{1i}, a_{2,i}, \dots, a_{ri}, a_{qi}), i = 1, \dots, r$. Il faut et il suffit donc, selon le critère de la Proposition 5 du § 5, chap. 2, que les p-r déterminants caractéristiques

$$d_{q} = \begin{vmatrix} a_{11} & \cdots & a_{1r} & b_{1} \\ \vdots & \cdots & \vdots & \vdots \\ a_{r1} & \cdots & a_{rr} & b_{r} \\ a_{a1} & \cdots & a_{ar} & b_{a} \end{vmatrix}$$

s'annulent pour tous les $q, r < q \le p$.

Théorème de Rouché-Fontené. Si un des p-r déterminants caractéristiques d_q ne s'annule pas, le système est impossible : pas de solution.

S'ils s'annulent tous, les r équations principales forment un système de Cramer à r équations et r inconnues principales, il y a indétermination d'ordre n-r.

3.3. Système homogène

Comme on l'a vu au § 1.2, pour qu'on ait une solution autre que la solution triviale $X = (x_j) = 0$, il faut et il suffit que le rang du système (de la matrice) soit strictement inférieur au nombre des inconnues, r < n.

Exemple. Soit (avec des notations un peu différentes) le système de trois équations à trois inconnues

$$\begin{cases} ax + by + cz &= 0\\ a'x + b'y + c'z &= 0\\ a''x + b''y + c''z &= 0 \end{cases},$$
(3.4)

ou encore $x\vec{V}_1 + y\vec{V}_2 + z\vec{V}_3 = 0$ avec $\vec{V}_1 = \begin{pmatrix} a \\ a' \\ a'' \end{pmatrix}$, etc. Si $D = \begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix} \neq 0$, le système est de Cramer et n'a que la solution triviale x = y = z = 0. Si D = 0 et si deux des vecteurs \vec{V}_i sont non colinéaires, par exemple \vec{V}_1 et \vec{V}_2 , le système est de rang 2. Par

exemple, si $ab' - ba' \neq 0$, le système des deux premières équations avec les termes en z au second membre est de Cramer, et on calcule donc x et y en termes de z

$$x = \frac{bc' - cb'}{ab' - ba'}z , \quad y = \frac{ca' - ac'}{ab' - ba'}z .$$

Si les trois vecteurs sont colinéaires mais non nuls, le rang est 1, il y a indétermination d'ordre 2, deux des inconnues sont des paramètres arbitraires, par exemple y et z, et alors x = -(by + cz)/a.

4. Un exemple détaillé

Soit le système

$$\begin{cases} \lambda x + y + z + t &= 1\\ x + \lambda y + z + t &= \mu\\ x + y + \lambda z + t &= \mu^{2}\\ x + y + z + \lambda t &= \mu^{3} \end{cases}$$
(4.1)

qui a donc p = n = 4. Le déterminant D de la matrice des coefficients se calcule aisément par combinaison des lignes et colonnes

$$D = \begin{vmatrix} \lambda & 1 & 1 & 1 \\ 1 & \lambda & 1 & 1 \\ 1 & 1 & \lambda & 1 \\ 1 & 1 & 1 & \lambda \end{vmatrix} = (\lambda + 3) \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & \lambda & 1 & 1 \\ 1 & 1 & 1 & \lambda \end{vmatrix} = (\lambda + 3) \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & \lambda - 1 & 0 & 0 \\ 1 & 0 & \lambda - 1 & 0 \\ 1 & 0 & 0 & \lambda - 1 \end{vmatrix} = (\lambda + 3)(\lambda - 1)^3.$$

Il faut donc distinguer selon les valeurs de λ

(a) Si
$$\lambda \neq 1$$
 et $\lambda \neq -3$

Le système est de Cramer et admet donc une solution unique. On peut utiliser la formule générale (3.1) mais il est sans doute préférable de combiner les équations de (4.1) : en les ajoutant toutes les quatre, on a $x+y+z+t=\frac{1+\mu+\mu^2+\mu^3}{\lambda+3}$ dont on retranche alors chacune pour en tirer

$$x = \frac{(\lambda+2) - \mu - \mu^2 - \mu^3}{(\lambda+3)(\lambda-1)} \quad , \quad y = \frac{\mu(\lambda+2) - 1 - \mu^2 - \mu^3}{(\lambda+3)(\lambda-1)}$$
$$z = \frac{\mu^2(\lambda+2) - 1 - \mu - \mu^3}{(\lambda+3)(\lambda-1)} \quad , \quad t = \frac{\mu^3(\lambda+2) - 1 - \mu - \mu^2}{(\lambda+3)(\lambda-1)}$$

(b) $\lambda = -3$, la matrice des coefficients est de rang r = 3 et on peut prendre par exemple $\begin{vmatrix} -3 & 1 & 1 \\ 1 & -3 & 1 \\ 1 & 1 & -3 \end{vmatrix} = -16$ comme déterminant principal, donc x, y, z comme inconnues $\begin{vmatrix} 1 & 1 & -3 \\ 1 & 1 & -3 \end{vmatrix}$

principales. Le déterminant caractéristique est obtenu en bordant le précédent en

$$d = \begin{vmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & \mu \\ 1 & 1 & -3 & \mu^2 \\ 1 & 1 & 1 & \mu^3 \end{vmatrix} = (1 + \mu + \mu^2 + \mu^3) \begin{vmatrix} -3 & 1 & 1 \\ 1 & -3 & 1 \\ 1 & 1 & -3 \end{vmatrix} = -16(1 + \mu + \mu^2 + \mu^3)$$

donc $d = -16(1 + \mu)(1 + \mu^2)$ ne s'annule que si $\mu = -1$ (on travaille ici sur les réels!).

- (i) $\mu \neq -1$, le système est impossible : pas de solution ;
- (ii) $\mu = -1$, indétermination d'ordre n r = 1, $x = z = t \frac{1}{2}$, y = t.
- (c) $\lambda = 1$: le rang r = 1, il y a une seule équation principale (disons la première) et une seule inconnue principale, disons x. Le système admet une solution ssi les trois déterminants caractéristiques

$$d_2 = \begin{vmatrix} 1 & 1 \\ 1 & \mu \end{vmatrix}, \quad d_3 = \begin{vmatrix} 1 & 1 \\ 1 & \mu^2 \end{vmatrix}, \quad d_4 = \begin{vmatrix} 1 & 1 \\ 1 & \mu^3 \end{vmatrix}$$

sont nuls, ce qui n'est vrai que si $\mu = \mu^2 = \mu^3 = 1$. (Plus simplement ici, on peut revenir au système (4.1) où on remplace λ par 1 : les 4 membres de gauche sont égaux, tandis que ceux de droite sont $1, \mu, \mu^2, \mu^3$. On retrouve que le système n'est possible que si $1 = \mu = \mu^2 = \mu^3$.) Donc

- (i) $\mu \neq 1$, système impossible, pas de solution ;
- (ii) $\mu = 1$, le système est de rang r = 1, une seule équation x + y + z + t = 1, indétermination d'ordre n-r = 3, par exemple y, z, t arbitraires et x = 1-y-z-t.

5. Applications mécaniques. Équilibre statique de solides indéformables

Considérons un solide indéformable, soumis à différentes forces statiques $\vec{F}_{\text{ext},i}$ s'appliquant en des points M_i : son poids \vec{P} , les réactions \vec{R}_i de différents corps avec lesquels il est en contact, etc. On sait que l'équilibre statique est conditionné par deux conditions vectorielles

$$\sum_{i} \vec{F}_{\text{ext},i} = 0 \qquad \sum_{i} \overrightarrow{OM_i} \wedge \vec{F}_{\text{ext},i} = 0 , \qquad (5.1)$$

pour un point O arbitraire. Ces 6 conditions (sur les composantes) constituent donc un système linéaire homogène dans les $\vec{F}_{\mathrm{ext},i}$ qui contraint les valeurs possibles des forces. En général, certaines des forces $\vec{F}_{\mathrm{ext},i}$ sont connues, poids, traction par un ressort etc, et le système est inhomogène dans les autres forces (réactions des supports, etc) qui sont les inconnues du problème. Dans de nombreux problèmes, en particulier impliquant des forces de frottement, ce système est indéterminé : certaines composantes des forces de réaction demeurent non déterminées.

On en verra un exemple en TD avec le problème de la stabilité d'une échelle.

6. Applications électriques. Circuits et lois de Kirchhoff

Soit un circuit électrique composé de résistances et de générateurs de tensions constantes données. Il s'agit de déterminer les intensités circulant dans toutes les branches du circuit. Le problème se ramène à un système d'équations linéaires couplées, auquel nous pouvons appliquer les techniques que nous venons d'étudier.

On obtient le système de la manière suivante : On fait un choix arbitraire d'orientation de chaque branche α du réseau et on lui attribue une variable d'intensité i_{α} : elle est à valeur algébrique, et sera positive ou négative selon que le courant circule dans le sens de l'orientation choisie ou en sens contraire ; l'ensemble des intensités i_{α} constitue l'ensemble des inconnues du problème. On écrit alors les lois de Kirchhoff :

- la loi des nœuds dit qu'à chaque nœud (jonction de plusieurs branches), la somme des intensités algébriques arrivant à ce nœud est nulle ; (ou encore que la somme des intensités entrantes est égale à la somme des sortantes) ;
- la loi des mailles dit que pour chaque circuit élémentaire fermé (ou "maille"), la somme des différences de potentiel le long des branches de la maille s'annule. Ces différences de potentiel sont pour chaque branche la somme de la chute ohmique $R_{\alpha}i_{\alpha}$ comptée algébriquement et de l'éventuelle tension (elle aussi comptée algébriquement) créée par un générateur.

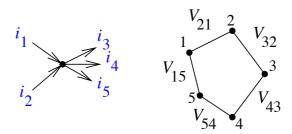


Fig. 7: Lois de Kirchhoff : loi des nœuds $i_1 + i_2 = i_3 + i_4 + i_5$, loi des mailles $V_{21} + V_{32} + \cdots + V_{51} = 0$.

Le système linéaire résultant a autant d'inconnues que de branches dans le circuit, ce qui peut être assez considérable (par exemple, 6 intensités dans le circuit de la figure 8). Il est souvent préférable de réduire le nombre de ces variables en utilisant les relations de nœuds. C'est ce qui a été fait sur la figure 8, où les 4 relations aux quatre nœuds ont permis de récrire le problème en termes de 3 intensités i_1 , i_2 , i_3 . Noter qu'on a choisi ces variables pour être des "intensités de maille", ce qui signifie que l'intensité circulant dans une branche donnée se lit comme somme à coefficients ± 1 , selon l'orientation, des

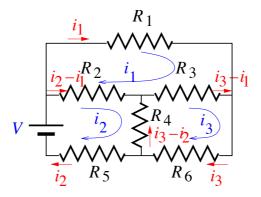


Fig. 8: Circuit à trois mailles

intensités des mailles auxquelles appartient cette branche. Les trois lois de mailles donnent alors le système

$$\begin{cases} R_{1}i_{1} + R_{3}(i_{1} - i_{3}) + R_{2}(i_{1} - i_{2}) &= 0 \\ R_{2}(i_{2} - i_{1}) + R_{4}(i_{2} - i_{3}) + R_{5}i_{2} &= V \\ R_{3}(i_{3} - i_{1}) + R_{5}i_{3} + R_{4}(i_{3} - i_{2}) &= 0 \end{cases}$$
soit
$$\begin{pmatrix} R_{1} + R_{2} + R_{3} & -R_{2} & -R_{3} \\ -R_{2} & R_{2} + R_{4} + R_{5} & -R_{4} \\ -R_{3} & -R_{4} & R_{3} + R_{4} + R_{6} \end{pmatrix} \begin{pmatrix} i_{1} \\ i_{2} \\ i_{3} \end{pmatrix} = \begin{pmatrix} 0 \\ V \\ 0 \end{pmatrix}$$

En présence de sources externes (ici le potentiel V), le système n'est pas homogène. Le déterminant de la matrice A n'est pas nul, le système est de Cramer et admet une solution unique.

La méthode s'étend aussi à des circuits comportant des condensateurs et des inductances, soumis à un courant de fréquence ω . Les calculs s'effectuent maintenant en complexes, ce qui ne présente aucune difficulté nouvelle pour les calculs de déterminants et les résolutions de systèmes d'équations linéaires.

Chapitre 4. Valeurs propres, vecteurs propres. Diagonalisation.

Étant donnée une matrice carrée A, on cherche à la mettre sous une forme semblable (au sens du chap. 1, (4.4)) particulièrement simple, à savoir une forme diagonale. Autrement dit on cherche une base dans laquelle seuls les éléments diagonaux de la matrice sont non nuls. On verra que cela n'est pas toujours possible, mais que les valeurs susceptibles d'apparaître sur la diagonale, les *valeurs propres*, peuvent être caractérisées assez simplement.

Les implications physiques de cette opération de "diagonalisation", dans des problèmes impliquant des oscillateurs mécaniques ou électriques couplés, sont importantes et seront discutées ensuite. Mais il existe bien d'autres problèmes physiques où ces concepts sont importants. Signalons ainsi qu'en Physique Quantique, où les quantités observables sont représentées par des opérateurs linéaires agissant sur l'espace vectoriel des états (ou des "fonctions d'onde"), les valeurs propres de ces opérateurs constituent les valeurs susceptibles d'être observées dans une expérience...

1. Vecteurs et valeurs propres

1.1. Définitions de base

Considérons une matrice carrée diagonale $n \times n$, c'est-à-dire de la forme

$$a_{ij} = \lambda_i \delta_{ij}$$

ce qu'on écrira encore

$$A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$
.

Si $\vec{e_i}$, $i=1,\cdots,n$ désignent les vecteurs de la base où A a cette forme, on voit que

$$Ae_i = \lambda_i e_i$$

où e_i est la matrice colonne représentant \vec{e}_i , soit $(e_i)_j = \delta_{ij}$.

Cela nous mène à la définition suivante, pour une matrice A carrée quelconque **Définition :** Soit A une matrice carrée, soit X un vecteur (matrice colonne) non nul tel que

$$AX = \lambda X , \qquad (1.1)$$

avec λ un nombre réel (ou complexe, voir plus bas). On dit que X est vecteur propre de A pour la valeur propre λ .

On vient de voir que si la matrice A est diagonale, alors chaque vecteur de base est un vecteur propre pour la valeur propre donnée par le terme correspondant de la diagonale de A.

Réciproquement, supposons que l'on ait trouvé n vecteurs propres linéairement indépendants X_i (une hypothèse pas innocente, comme on va le voir). Alors ces X_i peuvent être choisis comme nouvelle base de l'espace vectoriel, et dans cette base, A est diagonale. Une telle matrice est dite diagonalisable. Autrement dit, si la matrice A est diagonalisable, il existe une matrice V telle que $V^{-1}AV$ soit une matrice diagonale de valeurs propres,

$$V^{-1}AV = \Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) := \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \iff A = V\Lambda V^{-1}.$$

$$(1.2)$$

Mais toute matrice n'est pas diagonalisable. Ainsi comme on va le voir plus bas, la matrice "triangulaire supérieure" $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ n'est pas diagonalisable.

1.2. Valeurs propres d'une matrice singulière

Supposons que la matrice A est singulière. Cela signifie que ses n vecteurs colonnes \mathbf{A}^j ne sont pas indépendants (cf chap. 1, Théorème du § 4.6), donc qu'il existe n nombres non tous nuls x_j tels que $\sum_j x_j \mathbf{A}^j = 0$, soit encore

$$\forall i = 1, \dots, n \qquad \sum_{j} a_{ij} x_j = 0 \ . \tag{1.3}$$

Cela exprime que le vecteur X de composantes x_j est vecteur propre de A pour la valeur propre nulle. La réciproque est évidente : la condition (1.3) exprime la dépendance linéaire des colonnes de A, donc le fait qu'elle est singulière.

Proposition 1 : Une matrice est singulière (non inversible) ssi elle admet la valeur propre 0.

1.3. Sous-espace propre.

Soient X et Y deux vecteurs propres de A de même valeur propre : $AX = \lambda X$, $AY = \lambda Y$. Il est clair que toute combinaison linéaire de X et Y est aussi vecteur propre pour la valeur propre λ : $A(\alpha X + \beta Y) = \lambda(\alpha X + \beta Y)$. Les vecteurs propres de A pour une valeur propre donnée λ forment donc un sous-espace vectoriel, appelé espace propre de la valeur propre λ .

Proposition 2 : Deux vecteurs propres pour deux valeurs propres $\lambda \neq \mu$ sont nécessairement indépendants.

Preuve. Soient X un vecteur propre pour la valeur propre λ et Y un vecteur propre pour $\mu \neq \lambda$. Supposons X et Y linéairement dépendants. On aurait aX + bY = 0 pour deux nombres a et b non tous deux nuls : supposons par exemple b non nul. Appliquant A à cette relation, on aurait $0 = A(aX + bY) = a\lambda X + b\mu Y = 0$, soit une autre relation linéaire entre X et Y. En combinant ces deux relations (par exemple, λ fois la première moins la seconde), on aurait $b(\lambda - \mu)Y = 0$, avec $b \neq 0$ et $Y \neq 0$, ce qui implique $\lambda = \mu$ contrairement à l'hypothèse. La proposition est démontrée.

Plus généralement on démontre (par récurrence) que q vecteurs propres correspondant à q valeurs propres distinctes sont nécessairement linéairement indépendants.

Corollaire 1 : Si une matrice carrée $n \times n$ possède n valeurs propres distinctes, cette matrice est diagonalisable.

En effet elle possède alors n vecteurs propres indépendants (Prop. 2), on peut les choisir comme base, et la matrice est donc diagonalisable.

Exercice. Démontrer la Proposition suivante

Proposition 3 : Une matrice est diagonalisable ssi la somme des dimensions de ses espaces propres est égale à n.

1.4. Polynôme caractéristique

Soit λ une valeur propre de A. Nous récrivons la condition (1.1) sous la forme

$$(A - \lambda \mathbf{I})X = 0. (1.4)$$

Comme on l'a vu à la proposition 1 ci-dessus, l'existence d'un vecteur X satisfaisant (1.4) est la condition nécessaire et suffisante pour que $A - \lambda \mathbb{I}$ soit singulière. Mais se rappelant le Théorème fondamental du chapitre 2 (§ 3), cela est équivalent à $\det(A - \lambda \mathbb{I}) = 0$.

Pour une matrice carrée $n \times n$, l'expression

$$P(z) = \det(A - z\mathbb{I}) \tag{1.5}$$

est un polynôme de la variable z, de degré n en raison de la multilinéarité du déterminant.

Définition : Ce polynôme est appelé polynôme caractéristique de A.

On vient de voir que toute valeur propre est une racine du polynôme caractéristique. Réciproquement (grâce au fait que toutes les propositions impliquées sont des conditions nécessaires et suffisantes), toute racine de P(z) est une valeur propre de A.

Théorème 1 : Les valeurs propres sont les racines du polynôme caractéristique.

Selon que l'on travaille sur \mathbb{R} , ensemble des nombres réels, ou sur \mathbb{C} , les choses sont un peu différentes. Sur \mathbb{C} le polynôme caractéristique a exactement n racines, distinctes ou non ("théorème fondamental de l'algèbre"). On peut donc écrire

$$P(z) = \det(A - z\mathbb{I}) = \prod_{i=1}^{n} (\lambda_i - z)$$
(1.6)

avec le coefficient de z^n égal à $(-1)^n$ (pourquoi ?). En revanche il peut arriver que le polynôme caractéristique n'ait pas de racine sur \mathbb{R} , auquel cas la matrice A n'a pas de valeur propre réelle, ou qu'il n'ait que n' < n racines réelles.

Corollaire 2 : Une matrice réelle $n \times n$ a au plus n valeurs propres réelles, distinctes ou non.

Exemples

1. Soit
$$A = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$$
. Le polynôme caractéristique s'écrit

$$P(z) = \begin{vmatrix} 2-z & 1\\ 1 & -1-z \end{vmatrix} = (2-z)(-1-z) - 1 = z^2 - z - 3$$

et a deux racines distinctes $\frac{1}{2}(1 \pm \sqrt{13})$. La matrice est donc diagonalisable. On va voir au paragraphe suivant comment déterminer ses vecteurs propres.

2. Soit $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Son polynôme caractéristique est

$$\begin{vmatrix} 1-z & 1 \\ 0 & 1-z \end{vmatrix} = (1-z)^2 ,$$

et la seule valeur propre possible est $\lambda = 1$, la racine (double) de P(z). On montrera plus bas que la matrice n'est pas diagonalisable, n'ayant qu'un seul vecteur propre indépendant

pour la valeur propre 1. Cette situation doit être comparée avec celle de la matrice $A = \mathbb{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, qui a 1 comme valeur propre double mais qui est évidemment diagonalisable puisque déjà diagonale!

3. Soit $A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$ la matrice d'une rotation d'angle α dans le plan. Le polynôme caractéristique se calcule aisément

$$P(z) = z^2 - 2z\cos\alpha + 1$$

dont les racines sont complexes $z_{1,2} = \exp \pm i\alpha$. La matrice a deux valeurs propres distinctes, elle est diagonalisable à condition de "passer dans les complexes"; ses vecteurs propres sont alors eux-mêmes à composantes complexes, comme on verra ci-dessous. Sur \mathbb{R} , en revanche, la matrice n'est pas diagonalisable (pour $\alpha \neq 0, \pi$).

> Trace et déterminant en termes des valeurs propres

On constate dans les exemples précédents et on démontre aisément en général que

Proposition 4 : La somme des racines du polynôme caractéristique d'une matrice A est égale à sa trace, leur produit à son déterminant

$$\sum_{i=1}^{n} \lambda_i = \operatorname{tr} A \qquad \text{et} \qquad \prod_{i=1}^{n} \lambda_i = \det A \ . \tag{1.7}$$

En effet, en faisant z=0 dans (1.6), on a $P(0)=\det A=\prod_i\lambda_i$. La première propriété découle de la multilinéarité du déterminant : il n'est pas difficile d'identifier le terme en z^{n-1} dans le développement du déterminant comme $(-1)^{n-1}\sum_i a_{ii}$ et dans l'expression $\prod_i (\lambda_i - z)$ comme $(-1)^{n-1}\sum_i \lambda_i$.

Noter que pour des matrices 2×2 , on peut donc écrire l'équation caractéristique sous la forme

$$P(z) = z^2 - (\operatorname{tr} A) z + \det A = 0$$
, (1.8)

dont les coefficients se calculent aisément au vu de A.

Noter enfin que des matrices semblables ont même polynôme caractéristique, puisque si $B = W^{-1}AW$, alors $P_B(z) = \det(B-z\mathbb{I}) = \det(W^{-1}(A-z\mathbb{I})W) = \det(A-z\mathbb{I}) = P_A(z)$.

2. Diagonalisation d'une matrice

2.1. Détermination des vecteurs propres

Supposons qu'on connaisse une valeur propre λ de la matrice A, soit par recherche des racines de son polynôme caractéristique, soit par une autre méthode. Que peut-on dire alors de l'espace propre pour cette valeur propre ? Un vecteur propre X pour la valeur propre λ satisfait

$$(A - \lambda \mathbf{I})X = 0$$
.

On est ramené à un système linéaire homogène du type étudié au chapitre 3, dont on cherche les solutions X non triviales (non nulles), puisqu'un vecteur propre X est par définition non nul. On va donc s'intéresser au noyau de $A - \lambda \mathbb{I}$ et y chercher un ensemble maximal de vecteurs indépendants, donc une base du sous-espace propre de valeur propre λ .

Noter que certaines valeurs propres pouvant être complexes, on peut être amené à rechercher les vecteurs propres complexes correspondants, et donc à étendre la discussion du chapitre 3 au cas complexe, ce qui ne présente aucune difficulté nouvelle.

2.2. Diagonalisation. Changement de base.

Supposons maintenant qu'on a trouvé n vecteurs propres indépendants X_i de la matrice A. Formons la matrice V dont les X_i sont les vecteurs-colonnes. On a

$$AV = A\underbrace{(X_1 X_2 \cdots X_n)}_{V} = (\lambda_1 X_1 \ \lambda_2 X_2 \ \cdots \ \lambda_n X_n) = (X_1 X_2 \cdots X_n) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

donc $AV = V\Lambda$ où $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ est la matrice diagonale des valeurs propres, ce qu'on peut encore récrire

$$AV = V\Lambda \quad \Longleftrightarrow \quad A = V\Lambda V^{-1} \quad \Longleftrightarrow \quad \Lambda = V^{-1}AV \; .$$

La matrice V est donc la matrice qui diagonalise la matrice A.

On vient donc de démontrer la proposition

Proposition 5 : Si la matrice A admet n vecteurs propres indépendants X_i , elle est diagonalisable. La matrice V de diagonalisation, telle que $A = V\Lambda V^{-1}$, est la matrice dont les vecteurs-colonnes sont les vecteurs propres de A.

2.3. Exemples

Reprenons les exemples précédents :

1er exemple. $A = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$, valeurs propres $\lambda_{\pm} = \frac{1}{2}(1 \pm \sqrt{13})$. Pour λ_{+} , on écrit $A - \lambda_{+} \mathbb{I} = \begin{pmatrix} \frac{3}{2} - \frac{1}{2}\sqrt{13} & 1 \\ 1 & -\frac{3}{2} - \frac{1}{2}\sqrt{13} \end{pmatrix}$ qui est bien singulière (déterminant nul) et dont le noyau est engendré par les $X = \begin{pmatrix} \xi \\ \eta \end{pmatrix}$ tels que $(A - \lambda_{+} \mathbb{I})X = 0$, soit

$$\frac{1}{2}(3-\sqrt{13})\xi + \eta = 0 \ .$$

On prend par exemple $\xi = 1$ et $\eta = -\frac{1}{2}(3 - \sqrt{13})$. Un vecteur propre de A pour la valeur propre λ_+ est donc $X_+ = \begin{pmatrix} 1 \\ -\frac{1}{2}(3 - \sqrt{13}) \end{pmatrix}$ (ou tout vecteur qui lui est proportionnel). Un vecteur propre pour la valeur propre λ_{-} s'obtient ici simplement en changeant partout le signe de $\sqrt{13}$ dans l'expression de X_{+} (le vérifier). Finalement

$$X_{+} = \begin{pmatrix} 1 \\ -\frac{1}{2}(3 - \sqrt{13}) \end{pmatrix} \qquad X_{-} = \begin{pmatrix} 1 \\ -\frac{1}{2}(3 + \sqrt{13}) \end{pmatrix}.$$

Ces deux vecteurs propres peuvent être choisis comme nouvelle base, ce qui diagonalise la matrice

$$A = V \begin{pmatrix} \frac{1+\sqrt{13}}{2} & 0\\ 0 & \frac{1-\sqrt{13}}{2} \end{pmatrix} V^{-1} , \quad V = \begin{pmatrix} 1 & 1\\ \frac{-3+\sqrt{13}}{2} & \frac{-3-\sqrt{13}}{2} \end{pmatrix} , \quad V^{-1} = \frac{1}{\sqrt{13}} \begin{pmatrix} \frac{3+\sqrt{13}}{2} & 1\\ \frac{-3+\sqrt{13}}{2} & -1 \end{pmatrix}$$
$$V^{-1}AV = \begin{pmatrix} \lambda_{+} & 0\\ 0 & \lambda_{-} \end{pmatrix} .$$

2ème exemple. $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, valeur propre $\lambda = 1$ double. Le noyau de $(A - \mathbb{I}) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ est engendré par les vecteurs $X = \begin{pmatrix} \xi \\ \eta \end{pmatrix}$ tels que $\eta = 0, \xi$ arbitraire, ce qui donne un espace propre de dimension 1, engendré par le vecteur $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. La matrice A n'a qu'un seul vecteur

propre indépendant, elle n'est pas diagonalisable. '
3ème exemple. $A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$, valeurs propres complexes mais distinctes $\lambda_{\pm} = -\sin \alpha$ $\exp \pm i\alpha$. L'espace propre de la valeur propre λ_+ , c'est-à-dire le noyau de $A - \lambda_+ \mathbb{I} = \begin{pmatrix} -i\sin\alpha & -\sin\alpha \\ \sin\alpha & -i\sin\alpha \end{pmatrix} = \sin\alpha \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix}$ est engendré par les vecteurs $X_+ = \begin{pmatrix} \xi \\ \eta \end{pmatrix}$ tels

que $\xi - i\eta = 0$, par exemple $X_+ = \begin{pmatrix} i \\ 1 \end{pmatrix}$. Le vecteur complexe conjugué $X_- = \begin{pmatrix} -i \\ 1 \end{pmatrix}$ est vecteur propre pour λ_{-} . La matrice de diagonalisation est donc

$$V = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} , \quad V^{-1} = \frac{1}{2} \begin{pmatrix} -i & 1 \\ i & 1 \end{pmatrix} , \quad V^{-1}AV = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} .$$

(Exercice: vérifier le calcul de $V^{-1}AV$.) On voit que les calculs en nombres complexes n'offrent pas de difficulté supplémentaire.

2.4. "Triangularisation" d'une matrice. Théorème de Cayley-Hamilton

Même si la matrice A n'est pas diagonalisable, on démontre par récurrence sur sa dimension n que l'on peut toujours la "triangulariser" dans \mathbb{C} , c'est-à-dire trouver une matrice W à coefficients réels ou complexes telle que $T = W^{-1}AW$ soit une matrice triangulaire supérieure, dont tous les éléments au dessous de la diagonale sont nuls : $T_{ij} = 0$ si i > j. Dans ce cas encore, les valeurs apparaissant sur la diagonale sont les valeurs propres, avec leur multiplicité, pourquoi?

On verra en TD des exemples concrets de "triangularisation" d'une matrice non diagonalisable.

▷ Polynôme caractéristique et polynôme minimal

Appelons z_1, \dots, z_p les racines distinctes dans $\mathbb C$ du polynôme caractéristique P(z), et m_r , $r=1, \dots p$ la multiplicité de la racine z_r . On a $\sum_{r=1}^p m_r = n$, l'ensemble des valeurs propres (dans $\mathbb C$) de A est $\{\lambda_1, \lambda_2, \dots, \lambda_n\} = \{\underbrace{z_1, \dots, z_1}_{m_1}, \dots, \underbrace{z_p, \dots, z_p}_{m_p}\}$, et $P(z) = \prod_{r=1}^p (z_r - z)^{m_r}$. Si n_r est la dimension de l'espace propre pour la valeur propre $\lambda = z_r$, on a $n_r \leq m_r$.

Théorème de Cayley-Hamilton: La matrice A satisfait son équation caractéristique, autrement dit P(A) = 0.

Il faut noter que la matrice A commutant avec elle-même, avec ses puissances A^k et avec la matrice identité, la factorisation d'un polynôme comme P s'applique aussi quand on l'évalue en remplaçant z par A et on peut écrire $P(A) = (A - z_1 \mathbb{I})^{m_1} \cdots (A - z_p \mathbb{I})^{m_p}$ en écrivant les facteurs dans un ordre arbitraire.

Fig. 9: Le calcul de P(A) pour A triangulaire, schématisé ici pour p=3 valeurs propres distinctes, de multiplicités m_1, m_2, m_3 . Les matrices T sont triangulaires supérieures, $(T_1 - z_1 \mathbb{I})$ est strictement triangulaire, donc $(T_1 - z_1 \mathbb{I})^{m_1} = 0$ etc.

Pour démontrer le théorème, observons d'abord que si T est une matrice $m \times m$ triangulaire "strictement supérieure", c'est-à-dire telle que $T_{ij}=0$ si $i \geq j$, donc avec des zéros sur sa diagonale, alors $T^m=0$ (le vérifier pour m=2,3). Écrivons alors la matrice A sous forme triangulaire. Le r-ième facteur $(A-z_r\mathbb{I})$ de P(A) est lui-même une matrice triangulaire, avec dans le r-ième bloc le long de la diagonale, une matrice triangulaire strictement supérieure. En s'appuyant sur le schéma de la figure 9, vérifier que P(A)=0. Exemples et applications du théorème de Cayley-Hamilton.

– Considérons un espace E de dimension n muni d'une base e_1, \dots, e_n et l'opération A de projection dans le sous-espace E_1 de dimension q engendré par e_1, \dots, e_q "parallélement" au sous-espace E_2 engendré par e_{q+1}, \dots, e_n : cela généralise ce que nous avons fait au chap. 1, § 3.1. Les valeurs propres et espaces propres de A se trouvent sans aucun calcul: tout vecteur de l'espace E_1 est invariant, donc est vecteur propre de valeur propre 1, tout vecteur de l'espace E_2 est envoyé par A sur 0, donc est vecteur propre de valeur propre 0. La matrice A est diagonalisable (la somme des dimensions des sous-espaces propres E_1 et E_2 est n), son polynôme caractéristique est $(z-1)^q z^{n-q}$, et donc selon le théorème $(A-\mathbb{I})^q A^{n-q}=0$. En fait toute projection satisfait une équation beaucoup plus simple, $A^2=A$, comme on s'en convainc géométriquement: en itérant la projection (en calculant A^2), on ne modifie rien au résultat de la première projection! Or si $A^2=A$, on a aussi ($\mathbb{I}-A$)² = ($\mathbb{I}-A$) (le vérifier), donc ($\mathbb{I}-A$)⁴ $A^{n-q}=(\mathbb{I}-A)A=0$. − Soit A une matrice 2×2 . Elle satisfait donc son équation caractéristique (1.8), $A^2-(\operatorname{tr}A)A+(\det A)\mathbb{I}=0$. Si on veut calculer la trace de A^2 , il suffit de prendre la trace de cette expression pour obtenir $\operatorname{tr}A^2=(\operatorname{tr}A)^2-2$ det A. Comment calculer alors $\operatorname{tr}A^3$?

On peut enfin démontrer, et nous admettrons, la proposition suivante.

Proposition 6 : La matrice A est diagonalisable ssi elle satisfait $(A-z_1\mathbb{I})\cdots(A-z_p\mathbb{I})=0$, dite équation minimale.

Dans un sens, la proposition est évidente : si la matrice est diagonalisable, en se mettant dans la base où elle est diagonale, $(A-z_1\mathbb{I})\cdots(A-z_p\mathbb{I})$ est un produit de matrices diagonales, et la r-ième a m_r zéros dans son r-ième bloc. Ce produit est donc la matrice nulle. La preuve de la réciproque est plus délicate.

Exemple : la projection que nous venons de considérer est diagonalisable et satisfait ($\mathbb{I} - A$)A = 0. Autre exemple : quelles sont l'équation caractéristique et l'équation minimale satisfaites par une réflexion orthogonale dans un plan de \mathbb{R}^3 ? Une matrice R telle que $R^2 = \mathbb{I}$ est dite *involutive*.

3. Conséquences et applications de la diagonalisation

3.1. Matrices commutantes

Proposition 7 : Deux matrices carrées A et B de même dimension n, A ayant n valeurs propres distinctes, commutent ssi elles ont une base de vecteurs propres communs.

Preuve. La condition est nécessaire : si $AX = \lambda X$, on déduit $ABX = BAX = \lambda BX$. Le vecteur BX, s'il est non nul, est donc vecteur propre de A de valeur propre λ , et par l'hypothèse que les valeurs propres de A sont distinctes, est donc proportionnel à X, donc $BX = \mu X$, et X est aussi vecteur propre de B. Si BX = 0, X est vecteur propre de B de valeur propre nulle. Tout vecteur propre de l'une est vecteur propre de l'autre.

La condition est aussi suffisante. En effet les vecteurs propres communs X_i forment une base et sur chaque vecteur de cette base, $AX_i = \lambda_i X_i$, $BX_i = \mu_i X_i \implies (AB - BA)X_i = 0$, donc aussi pour tout vecteur X, A et B commutent, cqfd.

Applications. Si A a n valeurs propres distinctes, trouver les matrices carrées B qui commutent avec A. Réponse : si on diagonalise $A = V\Lambda V^{-1}$ en $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, ce sont toutes les matrices B de la forme $B = V\operatorname{diag}(\mu_1, \dots, \mu_n)V^{-1}$ avec μ_i quelconques. Pour le physicien, cette proposition prend tout son sens en Mécanique Quantique, puisque deux observables représentées par deux matrices A et B peuvent être mesurées simultanément ssi A et B commutent...

3.2. Puissances et exponentielle d'une matrice

domine sur $(-1)^p$, et on peut donc approxim

Soit A une matrice $n \times n$ dont nous désirons calculer la p-ième puissance, A^p , et en particulier, étudier le comportement asymptotique de A^p quand p tend vers l'infini. Si la matrice A est diagonalisable, $A = V\Lambda V^{-1}$, avec $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, on calcule aisément $A^p = V\Lambda^p V^{-1}$ et le calcul de Λ^p est trivial

$$A^p = V\Lambda^p V^{-1}, \qquad \Lambda^p = \operatorname{diag}\left(\lambda_1^p, \cdots, \lambda_n^p\right).$$

$$Exemple \ A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \operatorname{donc}$$

$$A^p = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 3^p & 0 \\ 0 & (-1)^p \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}. \quad \text{Asymptotiquement, le terme } 3^p$$

$$A^{p} \approx \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 3^{p} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{3^{p}}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} .$$

En général, on voit qu'asymptotiquement la puissance p-ième d'une matrice est dominée par sa (ou ses) plus grande(s) valeur(s) propre(s) (en valeur absolue ou en module). C'est d'ailleurs une méthode possible pour déterminer numériquement cette ou ces valeurs propre(s).

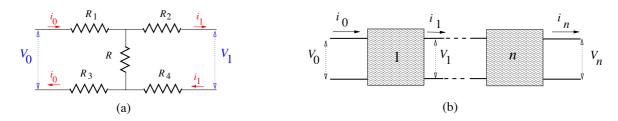


Fig. 10: (a) : quadripôle élémentaire. (b) chaîne de quadripôles en cascade.

Application : étude de la transmission d'une chaîne de quadripôles. Considérons une ligne de transmission électrique constituée d'une chaîne de quadripôles tels ceux étudiés en exercice au TD 1 et au TP1. On suppose que p quadripôles identiques sont montés "en cascade", voir figure 10. On désire relier le signal (V_n, i_n) au bout de la ligne à celui d'entrée (V_0, i_0) . Il est approprié d'utiliser la "matrice de transfert" T, cf ex. IV du TD 1, $\begin{pmatrix} V_j \\ i_j \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} V_{j-1} \\ i_{j-1} \end{pmatrix}$, de la diagonaliser, $T = W\Lambda W^{-1}$ et de calculer $T^n = W\Lambda^n W^{-1}$. On peut écrire $\begin{pmatrix} V_n \\ i_n \end{pmatrix} = T^n \begin{pmatrix} V_0 \\ i_0 \end{pmatrix}$. Pour le quadripôle de la figure , on a

$$T = \begin{pmatrix} 1 + \frac{R_{24}}{R} & -R_{1234} - \frac{R_{13}R_{24}}{R} \\ -\frac{1}{R} & 1 + \frac{R_{13}}{R} \end{pmatrix}$$

avec $R_{12} = R_1 + R_2$, etc. On vérifie immédiatement que det T = 1 et trT > 2, ce qui implique que les deux valeurs propres sont de la forme $\lambda > 1$ et $1/\lambda$.

Comme $\lambda^n \to \infty$ quand $n \to \infty$, la matrice T^n n'a pas de limite. Comment cela est-il compatible avec notre intuition qui nous dit que le signal doit s'atténuer le long de la ligne, à cause de la dissipation ohmique? Il faut bien voir que l'on ne peut pas fixer arbitrairement V_0 et i_0 et déduire des équations précédentes les valeurs de V_n et i_n . Typiquement on va fixer la tension d'entrée V_0 , imposer qu'à l'extrémité droite de la ligne, la tension (par exemple) est fixée, et on détermine alors les intensités i_0 et i_n . Si on écrit la matrice inverse de diagonalisation comme $W^{-1} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, la formule de diagonalisation de T^n :

 $W^{-1}T^n = \begin{pmatrix} \lambda^n & 0 \\ 0 & \lambda^{-n} \end{pmatrix} W^{-1}$ appliquée à $\begin{pmatrix} V_0 \\ i_0 \end{pmatrix}$ implique la paire de relations

$$\begin{cases} \alpha V_n + \beta i_n = \lambda^n (\alpha V_0 + \beta i_0) \\ \gamma V_n + \delta i_n = \lambda^{-n} (\gamma V_0 + \delta i_0) \end{cases}$$
(3.1)

Supposons par exemple qu'on ait branché un ampèremètre de résistance négligeable à la sortie (à droite). La tension V_n aux bornes de cet ampèremètre est nulle, et on y lit l'intensité i_n . La deuxième équation nous dit que $i_n = \delta^{-1} \lambda^{-n} (\gamma V_0 + \delta i_0)$, qui tend vers zéro (comme λ^{-n}) quand $n \to \infty$, il y a bien atténuation. La première équation nous apprend alors que $(\alpha V_0 + \beta i_0) = \lambda^{-n} \beta i_n$ qui tend vers zéro plus vite encore (comme λ^{-2n}). Cette relation détermine donc la valeur asymptotique de i_0 en fonction de $V_0: i_0 \approx -\alpha V_0/\beta$. Pour le quadripôle ci-dessus, on trouve une résistance effective du réseau égale à $\lim V_0/i_0 = -\beta/\alpha = \frac{1}{2}(R_{13} - R_{24} + \sqrt{R_{1234}}\sqrt{4R + R_{1234}})$. Bien entendu, si on avait branché un voltmètre de résistance infinie à droite plutôt que l'ampèremètre, on aurait $i_n = 0$, $V_n \sim \lambda^{-n}$ et la même relation asymptotique entre V_0 et i_0 .

▷ Le calcul des puissances d'une matrice via sa diagonalisation s'étend au calcul de l'exponentielle d'une matrice. Avec les mêmes notations que précédemment, on définit

$$\exp A = \sum_{p=0}^{\infty} \frac{A^p}{p!} = \sum_{p=0}^{\infty} \frac{(V\Lambda V^{-1})^p}{p!} = V \sum_{p=0}^{\infty} \frac{\Lambda^p}{p!} V^{-1} = V \operatorname{diag}(e^{\lambda_i}) V^{-1} , \qquad (3.2)$$

puisque $\Lambda^p = \text{diag}(\lambda_1^p, \dots, \lambda_n^p)$. On verra au paragraphe suivant que l'exponentielle d'une matrice se rencontre couramment dans la résolution des équations différentielles.

Toutes les séries ci-dessus convergent. Il faut d'abord définir une norme sur les matrices qui généralise la valeur absolue sur les réels ou le module sur les complexes. Par exemple $\parallel A-B \parallel = \sum_{i,j=1}^n |A_{ij}-B_{ij}|$. La convergence signifie que $\parallel \exp A - \sum_{p=0}^N \frac{A^p}{p!} \parallel$ peut être rendu aussi petit que l'on veut à condition de prendre N assez grand, et cela découle de la convergence de l'exponentielle usuelle des réels ou complexes.

4. Applications aux systèmes linéaires d'équations différentielles. Oscillateurs couplés

4.1. Systèmes de 2 équations différentielles linéaires

Les systèmes d'équations différentielles linéaires couplées peuvent aussi bénéficier d'un traitement par des méthodes matricielles.

Considérons par exemple le système linéaire de deux équations du premier ordre à coefficients constants a, b, c, d, que doivent satisfaire les fonctions inconnues $y_1(x)$ et $y_2(x)$

$$y_1' = ay_1 + by_2 y_2' = cy_1 + dy_2$$
(4.1)

et complété par la donnée de **deux** conditions initiales, par exemple de $y_1(0)$ et $y_2(0)$. Comme on l'a vu dans le cours LP 206, ce système peut être traité de plusieurs façons :

1) En dérivant la première équation par rapport à x et en y reportant l'expression de y'_2 donnée par la seconde, (c'est-à-dire en éliminant y'_2 entre ces deux équations), on obtient la paire d'équations

$$y_1'' = ay_1' + b(cy_1 + dy_2) y_1' = ay_1 + by_2$$
(4.2)

Entre ces deux équations, on peut cette fois éliminer y_2 , ce qui conduit à

$$y_1'' = (a+d)y_1' + (bc-ad)y_1. (4.3)$$

On traite alors cette équation différentielle linéaire du second ordre à coefficients constants par la méthode familière : la solution générale en $y_1(x)$ (et de même en $y_2(x)$) est une combinaison linéaire de fonctions exponentielles $e^{\lambda_1 x}$ et $e^{\lambda_2 x}$, où λ_1 et λ_2 sont solutions (supposées distinctes) de l'équation

$$\lambda^{2} - (a+d)\lambda + (ad - bc) = 0.$$
 (4.4)

On a donc $y_1(x) = A_1 e^{\lambda_1 x} + B_1 e^{\lambda_2 x}$, d'où l'on tire $y_2(x)$ par la première équation (4.1), et les constantes A_1 et B_1 sont alors déterminées grâce aux conditions initiales.

En résumé, le système de deux équations linéaires du premier ordre à coefficients constants (4.1) est équivalent à une équation linéaire du second ordre.

2) Une autre méthode consiste à effectuer des combinaisons linéaires des deux équations de (4.1) avec des coefficients ξ et η constants (indépendants de x) quelconques : on trouve que

$$(\xi y_1 + \eta y_2)' = (a\xi + c\eta)y_1 + (b\xi + d\eta)y_2. \tag{4.5}$$

Supposons qu'on sache trouver ξ et η tels que

$$a\xi + c\eta = \lambda\xi b\xi + d\eta = \lambda\eta$$
(4.6)

pour un certain nombre λ . Alors, en reportant (4.6) dans (4.5), on voit que l'équation (4.5) est une équation différentielle linéaire du premier ordre particulièrement simple pour la fonction $y := \xi y_1 + \eta y_2$

$$y' = \lambda y , \qquad (4.7)$$

ce qui s'intègre immédiatement en $y=Ce^{\lambda x}$. Le système homogène (4.6) n'admet de solution non nulle en (ξ,η) que si son déterminant s'annule

$$(d - \lambda)(a - \lambda) - bc = 0 , \qquad (4.8)$$

équation du second degré en λ qui admet elle-même deux solutions, réelles ou complexes, distinctes ou confondues. Supposons les distinctes, il existe donc deux valeurs λ_1 et λ_2 de λ , donc aussi deux paires (ξ_1, η_1) et (ξ_2, η_2) (à un facteur global près) remplissant les conditions ci-dessus, c'est-à-dire conduisant à une équation du type (4.7) pour la combinaison linéaire correspondante. Chacune de ces deux combinaisons linéaires est appelée mode propre du système initial (4.1). Par les combinaisons algébriques que nous avons effectuées, nous avons donc réduit le système initial à la solution de deux équations du premier ordre découplées. Les fonctions $y_1(x)$ et $y_2(x)$ se déterminent finalement en résolvant le système $\xi_1 y_1 + \eta_1 y_2 = C_1 e^{\lambda_1 x}$, $\xi_2 y_1 + \eta_2 y_2 = C_2 e^{\lambda_2 x}$. Les constantes C_1 et C_2 sont fixées par les conditions initiales.

On note que les deux approches ont en commun de faire jouer un rôle central à l'équation "caractéristique" (4.4)-(4.8).

3) Récrivons le système linéaire (4.1) sous une forme matricielle

$$\frac{d}{dx} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \tag{4.9}$$

ou en définissant

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} ,$$

$$\frac{d}{dx}Y = AY . \tag{4.10}$$

Supposons maintenant que la matrice A puisse se diagonaliser par un changement de base $A = V\Lambda V^{-1}, \ V^{-1} = \begin{pmatrix} \xi_1 & \eta_1 \\ \xi_2 & \eta_2 \end{pmatrix}$ une matrice 2×2 indépendante de $x, \ \Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$

la matrice diagonale où λ_1 et λ_2 sont les valeurs propres de A. Définissant la matrice-colonne $\tilde{Y} = V^{-1}Y = \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix} = \begin{pmatrix} \xi_1 y_1 + \eta_1 y_2 \\ \xi_2 y_1 + \eta_2 y_2 \end{pmatrix}$, on a en multipliant les deux membres de l'équation matricielle (4.10) par V^{-1}

$$\frac{d}{dx}\tilde{Y} = \Lambda \tilde{Y}$$

$$\begin{pmatrix} \tilde{y}_1' \\ \tilde{y}_2' \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix}$$
(4.11)

c'est-à-dire deux équations découplées

$$\tilde{y}_1' = \lambda_1 \tilde{y}_1 \quad \text{et} \quad \tilde{y}_2' = \lambda_2 \tilde{y}_2 .$$
 (4.12)

La démarche reproduit celle suivie au point 2). Les valeurs propres λ_1 et λ_2 sont les racines de l'équation (4.8) ; la diagonalisation de la matrice, c'est-à-dire l'existence d'une matrice V, est assurée par l'hypothèse que ces deux racines sont distinctes ; et les combinaisons \tilde{y}_1 et \tilde{y}_2 sont les deux "modes propres" définis en 2).

L'avantage de cette méthode matricielle est sa puissance et sa généralité. Elle s'étend sans difficulté à des systèmes d'équations différentielles de dimension et/ou d'ordre plus élevés.

4.2. Systèmes de n équations

Considérons un système de n équations différentielles du premier ordre

$$\frac{d}{dx}Y(x) = AY(x) \quad \text{ou encore} \quad \frac{d}{dx} \begin{pmatrix} y_1(x) \\ \vdots \\ y_n(x) \end{pmatrix} = A \begin{pmatrix} y_1(x) \\ \vdots \\ y_n(x) \end{pmatrix}$$
(4.13)

avec A une matrice de coefficients constants. Cette équation est complétée par n conditions initiales (ou "au bord"), par exemple $y_i(0) = y_{i0}$ donnés pour $i = 1, \dots, n$.

 \triangleright Supposons la matrice A diagonalisable, et soient $\lambda_1, \dots, \lambda_n$ ses valeurs propres, et V la matrice de ses vecteurs propres. On écrit comme plus haut $\Lambda = \operatorname{diag}(\lambda_i) = V^{-1}AV$. Définissant alors

$$\tilde{Y}(x) = V^{-1}Y(x),$$

on a (après multiplication de (4.13) par V^{-1})

$$\frac{d}{dx}\tilde{Y}(x) = V^{-1}\frac{d}{dx}Y(x) = V^{-1}AY = V^{-1}AVV^{-1}Y = \Lambda \tilde{Y} , \qquad (4.14)$$

complété par les conditions au bord $\tilde{Y}(0) = V^{-1}Y(0)$. Mais ce nouveau système est aisé à intégrer, puisqu'il est diagonal :

$$\frac{d}{dx}\tilde{Y}(x) = \Lambda \tilde{Y}(x) \Longleftrightarrow \frac{d}{dx}\tilde{y}_i(x) = \lambda_i \tilde{y}_i(x) \tag{4.15}$$

dont la solution est

$$\tilde{y}_i(x) = e^{\lambda_i x} \tilde{y}_i(0) \iff \tilde{Y}(x) = \exp(\Lambda x) \tilde{Y}(0)$$

ou en revenant aux notations de départ

$$Y(x) = V\tilde{Y}(x) = V \operatorname{diag}(e^{\lambda_i x}) V^{-1} Y(0)$$
 (4.16)

En comparant avec (3.2), on voit que ceci n'est autre que

$$Y(x) = \exp(Ax)Y(0) . \tag{4.17}$$

En pratique, la diagonalisation (4.16) permet d'obtenir une résolution du système (4.13) plus explicite que la forme (4.17), cf exercices de TD. Elle permet aussi de mieux comprendre la physique en jeu, comme on va l'illustrer sur l'étude d'oscillateurs couplés.

4.3. Oscillateurs couplés

La méthode précédente s'applique aussi à des équations différentielles couplées d'ordre plus élevé, telles celles qui régissent des oscillateurs harmoniques couplés, voir cours de LP 206. Nous allons en reprendre la discussion avec les méthodes de l'algèbre linéaire.

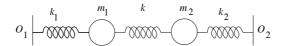


Fig. 11: 2 oscillateurs couplés

▷ Oscillations longitudinales de deux oscillateurs couplés

On considère le système constitué de deux masses ponctuelles m_1 et m_2 reliées par des ressorts de raideurs k_1 et k_2 fixés en O_1 et O_2 et couplées par un ressort de raideur k, voir Fig. 11. x_1 et x_2 représentent les écarts de m_1 et m_2 par rapport à leurs positions d'équilibre, comptés positivement vers la droite. On calcule alors les forces auxquelles sont soumises chacune des deux masses m_1 et m_2 et on écrit leurs équations du mouvement sous la forme

$$m_1 \ddot{x}_1 = -k_1 x_1 + k(x_2 - x_1)$$

$$m_2 \ddot{x}_2 = -k_2 x_2 + k(x_1 - x_2)$$
(4.18)

Dans la suite on supposera, pour simplifier les calculs, que les masses sont identiques, $m = m_1 = m_2$, ainsi que les constantes de raideur $k_1 = k_2 = k'$; k_1 et k_2 jouent alors des rôles identiques. On peut donc récrire (4.18) sous la forme:

$$\ddot{x}_1 + (\Omega_0^2 + \omega_0^2)x_1 - \omega_0^2 x_2 = 0$$
$$\ddot{x}_2 + (\Omega_0^2 + \omega_0^2)x_2 - \omega_0^2 x_1 = 0$$

avec $\omega_0^2 = k/m$ et $\Omega_0^2 = k'/m$.

On peut récrire ce système sous forme matricielle

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \tag{4.19}$$

avec
$$A = \begin{pmatrix} (\Omega_0^2 + \omega_0^2) & -\omega_0^2 \\ -\omega_0^2 & (\Omega_0^2 + \omega_0^2) \end{pmatrix}$$
.

Cherchons des solutions de la forme $\binom{a}{b}e^{i\omega t}$, étant entendu que comme il est usuel dans ce type de problème, l'introduction d'exponentielles complexes sert uniquement à simplifier les calculs intermédiaires ; la solution physique s'obtient à la fin en imposant des conditions initiales telles qu'elle est bien une combinaison réelle des solutions complexes obtenues. L'équation (4.19) se récrit alors comme

$$(A - \omega^2 \mathbb{I}) \begin{pmatrix} a \\ b \end{pmatrix} = 0$$

c'est-à-dire comme une équation aux valeurs et vecteurs propres.

Dans le cas présent de masses et de coefficients de raideur égaux, il est facile de voir que ces vecteurs propres sont $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, autrement dit que les modes propres du système (4.19) sont $\xi_{\pm} = x_1 \pm x_2$ et satisfont des équations simples, découplées,

$$\ddot{\xi}_{+} + \Omega_{0}^{2} \xi_{+} = 0$$

$$\ddot{\xi}_{-} + (\Omega_{0}^{2} + 2\omega_{0}^{2}) \xi_{-} = 0$$

où on note $\omega_+ = \Omega_0$ et $\omega_- = \sqrt{\Omega_0^2 + 2\omega_0^2}$ les deux fréquences propres du système. La solution générale pour $x_1(t)$ et $x_2(t)$ en découle. Supposant par exemple que le système au temps t=0 est tel que seule la coordonnée x_1 est non nulle, $x_1(0)=a$, tandis que $x_2(0)=0$ et que les vitesses initiales sont nulles $\dot{x}_1(0)=\dot{x}_2(0)=0$, l'expression de $x_1(t)$ et $x_2(t)$ est donnée par

$$x_1(t) = \frac{a}{2}(\cos\omega_+ t + \cos\omega_- t) = a\cos(\frac{\omega_+ + \omega_-}{2}t)\cos(\frac{\omega_- - \omega_+}{2}t)$$
$$x_2(t) = \frac{a}{2}(\cos\omega_+ t - \cos\omega_- t) = a\sin(\frac{\omega_+ + \omega_-}{2}t)\sin(\frac{\omega_- - \omega_+}{2}t).$$

La première expression sous forme de somme est adaptée à la discussion du couplage fort, voir ci-dessous, la seconde (produit cos cos ou sin sin) à celle du couplage faible, avec ses phénomènes de battement, etc.

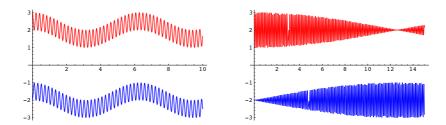


Fig. 12: On a porté verticalement $x_1 + \Delta$ (en rouge) et $x_2 - \Delta$ (en bleu), en prenant $\Delta = 2$. On a pris l'amplitude initiale a = 1. A gauche, couplage fort : $\omega_0 = 30$, $\Omega_0 = 1$; à droite, couplage faible, $\omega_0 = 5$, $\Omega_0 = 50$.

Les graphes de x_1 et de x_2 sont représentés sur la figure 12. Dans le cas d'un couplage fort, $\omega_0 \gg \Omega_0$, (ressort du milieu dur), $\omega_- \approx \omega_0 \sqrt{2} \gg \omega_+ = \Omega_0$, les deux oscillateurs oscillent ensemble, avec des oscillations rapides (de fréquence $\omega_0/2\pi$ élevée) et d'amplitude a/2 autour de leur mouvement d'oscillation lent (de fréquence $\Omega_0/2\pi$ basse). Dans le cas d'un couplage faible, $\omega_0 \ll \Omega_0$, $\omega_- \approx \omega_+ = \Omega_0$, $\omega_- - \omega_+ \approx \omega_0^2/\Omega_0 \ll \omega_0$, les oscillations du premier se transmettent peu à peu au deuxième, puis inversement. Les deux oscillateurs semblent être synchronisés à la fréquence $\frac{\omega_+ + \omega_-}{2}$, en quadrature de phase (déphasage de $\pi/2$), mais leur amplitude varie lentement (puisque $\omega_- - \omega_+ \ll \Omega_0$) comme $\cos(\frac{\omega_- - \omega_+}{2} t)$ ou $\sin(\frac{\omega_- - \omega_+}{2} t)$. Il y a donc des battements, eux aussi en quadrature. Il y a transfert d'énergie alternativement d'un oscillateur à l'autre.

▷ Oscillateurs électriques

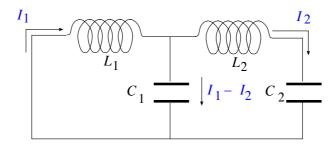


Fig. 13: 2 oscillateurs électriques couplés

On se rappelle que les tensions aux bornes et les courants traversant une bobine d'induction d'inductance L, resp. un condensateur de capacité C, satisfont

$$U_L = L \frac{\mathrm{d}I_L}{\mathrm{d}t} \qquad I_C = C \frac{\mathrm{d}U_C}{\mathrm{d}t} \ .$$

Dans le circuit de la figure 13, on a les relations

$$U_{L_1} = L_1 \frac{dI_1}{dt} \qquad I_1 - I_2 = C_1 \frac{dU_{C_1}}{dt}$$

$$U_{L_2} = L_2 \frac{dI_2}{dt} \qquad I_2 = C_2 \frac{dU_{C_2}}{dt}$$

$$U_{L_1} + U_{C_1} = 0 \qquad U_{L_2} + U_{C_2} - U_{C_1} = 0$$

et après élimination des I_i et des U_{L_i} , on trouve pour les $U_i \equiv U_{C_i}$ le système d'équations couplées suivant (le vérifier !)

$$L_1 \left(C_1 \frac{\mathrm{d}^2 U_1}{\mathrm{d}t^2} + C_2 \frac{\mathrm{d}^2 U_2}{\mathrm{d}t^2} \right) + U_1 = 0$$
$$L_2 C_2 \frac{\mathrm{d}^2 U_2}{\mathrm{d}t^2} + U_2 - U_1 = 0$$

Si $L_1 = L_2$ on peut reporter la seconde dans la première et obtenir

$$L_1 C_1 \frac{\mathrm{d}^2 U_1}{\mathrm{d}t^2} + 2U_1 - U_2 = 0$$
$$L_2 C_2 \frac{\mathrm{d}^2 U_2}{\mathrm{d}t^2} + U_2 - U_1 = 0$$

Montrer l'analogie avec les deux ressorts couplés étudiés plus haut. Étudier les modes propres de ce circuit. (Voir TD4).

On étudiera en TP ce type d'oscillateur, amorti par la présence de résistances.

Chapitre 5. Formes quadratiques et matrices symétriques.

1. Formes bilinéaires, formes quadratiques

1.1. Formes bilinéaires et quadratiques

On a déjà rencontré la notion de forme multilinéaire (Chap. 2). Sur un espace vectoriel E, on appelle forme bilinéaire réelle une application qui fait correspondre à toute paire de vecteurs $X, Y \in E$ un nombre réel f(X, Y), cette application étant linéaire en X et en Y, donc

$$f(\lambda_1 X_1 + \lambda_2 X_2, Y) = \lambda_1 f(X_1, Y) + \lambda_2 f(X_2, Y)$$

$$f(X, \mu_1 Y_1 + \mu_2 Y_2) = \mu_1 f(X, Y_1) + \mu_2 f(X, Y_2) . \tag{1.1}$$

La forme bilinéaire est dite symétrique si f(X,Y) = f(Y,X).

Exemples. Le produit scalaire $\vec{X}.\vec{Y}$ dans l'espace euclidien \mathbb{R}^n est une forme bilinéaire symétrique. La composante sur un axe donné du produit vectoriel $\vec{X} \wedge \vec{Y}$ dans l'espace \mathbb{R}^3 est une forme bilinéaire, mais pas symétrique (elle est en fait antisymétrique!). Si g et h sont deux fonctions d'une variable réelle, intégrables sur un intervalle (a,b), $f(g,h) = \int_a^b g(x)h(x) \, \mathrm{d}x$ est une forme bilinéaire symétrique en g et h.

Le premier exemple suggère la définition suivante : Étant donnée une forme bilinéaire symétrique f, on dit que X et Y sont orthogonaux pour f si f(X,Y)=0.

Étant donnée la forme bilinéaire f(X,Y), on lui associe une forme quadratique par

$$Q(X) = f(X, X) . (1.2)$$

Bien sûr, cette forme quadratique n'est pas linéaire : $Q(\lambda X) = \lambda^2 Q(X)$. Inversement pour toute forme quadratique Q, on peut construire une forme bilinéaire symétrique f telle que Q(X) = f(X, X) par l'opération de polarisation : on écrit simplement, grâce à la bilinéarité

$$f(X+Y,X+Y) = f(X,X) + f(X,Y) + f(Y,X) + f(Y,Y)$$
(1.3)

et si on fait l'hypothèse que f est symétrique, $f(X,Y) = \frac{1}{2}(f(X+Y,X+Y)-f(X,X)-f(Y,Y)) = \frac{1}{2}(Q(X+Y)-Q(X)-Q(Y))$.

Exemples. Reprenons les deux exemples ci-dessus. Au produit scalaire dans \mathbb{R}^n correspond la forme quadratique $\|\vec{X}\|^2 = \vec{X}.\vec{X}$ qui est la norme carrée (la longueur carrée) du vecteur \vec{X} . De même, $\int_a^b f^2(x) dx$ est une norme carrée pour les fonctions (de carré intégrable) sur (a, b).

Théorème de Pythagore. Soit f une forme bilinéaire symétrique, Q la forme quadratique associée, on a pour toute paire de vecteurs orthogonaux

$$\forall X, Y : f(X,Y) = 0 \Longrightarrow Q(X+Y) = Q(X) + Q(Y) , \qquad (1.4)$$

qui découle de (1.3).

1.2. Formes définies positives

On dit que la forme quadratique Q est définie positive si

$$\forall X \neq 0 \in E \qquad Q(X) > 0, \tag{1.5}$$

et donc Q(X)=0 si et seulement si X=0. La forme est semi-définie positive si l'inégalité n'est pas stricte : $\forall X \neq 0 \in E$ $Q(X) \geq 0$, elle est indéfinie si Q(X) peut prendre un signe ou l'autre selon la valeur de X. Par abus de langage on dit d'une forme bilinéaire qu'elle est définie positive, semi-définie positive, etc, si la forme quadratique associée l'est. Exemples. Le produit scalaire habituel dans l'espace euclidien \mathbb{R}^n est défini positif, $Q(\vec{X})$ définissant la norme carrée, c'est-à-dire la longueur carrée du vecteur \vec{X} . Au contraire, dans l'espace-temps de la Relativité restreinte (espace de Minkowski), la forme quadratique $Q(X) = c^2t^2 - x_1^2 - x_2^2 - x_3^2$ est indéfinie : les quadrivecteurs de "genre temps" ont une norme carrée positive, ceux de "genre espace" une norme carrée négative, ceux de "genre lumière" une norme nulle. Dans l'espace \mathbb{R}^2 , la forme quadratique $Q(X) = x_1x_2$ est indéfinie et la forme $Q'(X) = (x_1 - x_2)^2$ est semi-définie positive, pourquoi ?

Si la forme symétrique f est définie positive, pour toute paire X,Y de vecteurs non colinéaires et tout réel λ , le vecteur $\lambda X + Y$ n'est pas nul, donc $Q(\lambda X + Y) > 0$ est strictement positif. Or

$$Q(\lambda X + Y) = \lambda^2 Q(X) + 2\lambda f(X, Y) + Q(Y) .$$

est un trinôme du second degré en λ , et le fait qu'il est toujours strictement positif implique que son discriminant est négatif, donc

$$\Delta' = f(X, Y)^2 - Q(X)Q(Y) < 0$$

En revanche si X et Y sont colinéaires, il existe un λ_0 tel que $\lambda_0 X + Y = 0$, et alors $Q(\lambda X + Y) \ge 0$ s'annule en λ_0 mais ne change pas de signe, son discriminant est nul. On obtient ainsi l'inégalité de Schwarz

$$|f(X,Y)| \le (Q(X)Q(Y))^{\frac{1}{2}},$$
 (1.6)

avec égalité si et seulement si X et Y sont colinéaires.

Exemple: dans l'exemple précédent de l'espace euclidien \mathbb{R}^3 , cette inégalité nous dit que

$$|\vec{X}.\vec{Y}| \leq \parallel \vec{X} \parallel \parallel \vec{Y} \parallel$$

ou encore, si on se rappelle la formule de trigonométrie $\cos \theta = \frac{\vec{X}.\vec{Y}}{\|\vec{X}\| \|\vec{Y}\|}$, que $|\cos \theta| \leq 1$, avec égalité ssi $\theta = 0$ ou π donc \vec{X} et \vec{Y} colinéaires. Plus généralement, pour toute forme bilinéaire définie positive, l'inégalité de Schwarz (1.6) nous permet de définir (au signe près et à 2π près) l'angle θ entre deux vecteurs X et Y par $\cos \theta = f(X,Y)/(Q(X)Q(Y))^{\frac{1}{2}}$.

1.3. Représentations matricielles

Supposons que l'on a choisi une base e_i dans l'espace E. Dans cette base, on écrit les vecteurs $X = \sum_i x_i e_i$ et $Y = \sum_i y_i e_i$, donc la forme bilinéaire

$$f(X,Y) = \sum_{ij} x_i y_j f(e_i, e_j) = \sum_{ij} x_i b_{ij} y_j ,$$

où la matrice B de la forme bilinéaire (dans la base choisie e_i) est définie par

$$B = (b_{ij})$$
 $b_{ij} = f(e_i, e_j)$. (1.7)

Cette matrice est symétrique, $b_{ij} = b_{ji}$, si la forme bilinéaire l'est. Utilisant la même notation X et Y pour les matrices colonnes des composantes de X et Y, on voit que l'on peut écrire

$$f(X,Y) = X^T B Y .$$

Supposons maintenant que l'on effectue un changement de base $e_i \to e'_j = \sum_i e_i a_{ij}$ (cf Chap 1, (2.4)). Comme on l'a vu au chapitre 1, les composantes X et X' d'un vecteur donné dans l'ancienne et la nouvelle base sont reliées par X = AX' (Chap 1, (2.5)). Par conséquent la forme bilinéaire s'exprime maintenant selon $f(X,Y) = X^TBY = X'^TA^TBAY'$ donc à l'aide de la matrice $B' = A^TBA$ (et non pas selon $A^{-1}BA$ comme pour une application linéaire, comparer avec Chap 1, (4.4)!)

2. Réduction d'une forme quadratique

Dans toute cette section on supposera que les formes bilinéaires et les matrices associées sont symétriques.

2.1. Vecteurs orthogonaux, vecteurs orthonormés

Définition : Si f est une forme bilinéaire symétrique définie positive, on dit que des vecteurs X_1, \dots, X_k sont orthonormés (pour la forme f) si

$$f(X_i, X_j) = \delta_{ij}$$

autrement dit si ces vecteurs sont deux à deux orthogonaux : $f(X_i, X_j) = 0$ si $X_i \neq X_j$ et s'ils sont normés $Q(X_i) = 1$.

Lemme 1 : Si les vecteurs X_1, \dots, X_k sont orthonormés (pour la forme f), ils sont nécessairement linéairement indépendants.

La preuve (élémentaire!) est laissée en exercice.

2.2. Procédé d'orthonormalisation de Schmidt

Soit f une forme bilinéaire symétrique définie positive.

Théorème 1 : À partir de tout système de k vecteurs linéairement indépendants X_1, \dots, X_k , on peut construire un système de k vecteurs orthonormés $\tilde{X}_1, \dots, \tilde{X}_k$, combinaisons linéaires des X_1, \dots, X_k .

Preuve par récurrence sur k. Pour k=1, on dispose d'un vecteur X_1 non nul, donc de norme non nulle. Le vecteur $\tilde{X}_1=X_1/Q(X_1)^{\frac{1}{2}}$ est bien normé. Supposons alors la propriété vraie pour tout système de k-1 vecteurs, et considérons le système de k vecteurs linéairement indépendants X_1,\cdots,X_k . Le sous-système X_1,\cdots,X_{k-1} remplit la condition de récurrence, on peut donc construire un système de k-1 vecteurs orthonormés $\tilde{X}_1,\cdots,\tilde{X}_{k-1}$, combinaisons linéaires des X_1,\cdots,X_{k-1} . Le k-ième vecteur X_k est indépendant de X_1,\cdots,X_{k-1} donc aussi de $\tilde{X}_1,\cdots,\tilde{X}_{k-1}$. Cherchons un $Y=X_k+\sum_{i=1}^{k-1}\lambda_i\tilde{X}_i$ orthogonal à $\tilde{X}_1,\cdots,\tilde{X}_{k-1}$: en prenant le produit scalaire par f entre cet Y et les autres : $f(Y,\tilde{X}_i)=f(X_k,\tilde{X}_i)+\lambda_i$, on détermine $\lambda_i=-f(X_k,\tilde{X}_i)$. Finalement ce vecteur Y étant non nul (sans quoi X_k ne serait pas linéairement indépendant des $\tilde{X}_1,\cdots,\tilde{X}_{k-1}$), il suffit de le normer pour obtenir $\tilde{X}_k=Y/f(Y,Y)^{\frac{1}{2}}$ et terminer la preuve par récurrence.

Ce théorème a comme corollaire que l'on peut toujours trouver une base orthonormale dans l'espace vectoriel E.

Bien comprendre que ce théorème, sous l'hypothèse de l'existence d'une forme bilinéaire définie positive, nous ramène sur le terrain bien connu de la géométrie euclidienne. Dans la base orthonormée, la forme bilinéaire prend l'allure familière du produit scalaire

"en coordonnées rectangulaires", $f(X,Y) = \sum_i x_i y_i$, et la norme carrée $Q(X) = \sum_i x_i^2$. Un espace vectoriel doté d'une forme bilinéaire définie positive est appelé espace euclidien.

Exemple. Considérons l'espace E des polynômes de degré $\leq n$ dans la variable x et définissons la forme bilinéaire $f(p,q)=\int_{-1}^1 p(x)q(x)\,\mathrm{d}x$. Cette forme est évidemment symétrique et définie positive. À partir de la base naturelle $\{1,x,x^2,\cdots x^n\}$ de l'espace E, on peut, grâce au procédé d'orthonormalisation de Schmidt, construire une base orthonormée $p_k(x)$. Ce sont les polynômes $p_k(x)=(k+\frac{1}{2})^{\frac{1}{2}}P_k(x)$, avec P_k les "polynômes de Legendre" $P_0(x)=1$, $P_1(x)=x$, $P_2(x)=\frac{1}{2}(3x^2-1)$, etc. (Voir TP1.) Ces polynômes seront rencontrés par la suite dans le cours de mécanique quantique, où ils jouent un rôle important dans la description du moment angulaire.

2.3. Matrices orthogonales

Considérons un espace E doté d'une forme bilinéaire définie positive, donc euclidien. On notera dans la suite X.Y = f(X,Y) et $||X||^2 = Q(X)$. Soient e_i une base orthonormée, x_i , y_i les composantes de deux vecteurs X et Y dans cette base : $X = \sum_i x_i e_i$, $Y = \sum_i y_i e_i$.

Le produit scalaire et la norme carrée y prennent donc les expressions familières

$$X.Y = \sum_{i} x_{i} y_{i} \qquad ||X||^{2} = \sum_{i} x_{i}^{2} , \qquad (2.1)$$

et les composantes x_i , y_i s'expriment en termes de produits scalaires avec les vecteurs de base

$$x_i = e_i.X, \quad y_i = e_i.Y \ . \tag{2.2}$$

En termes des vecteurs colonnes des composantes de X et Y

$$X.Y = X^T Y = Y^T X$$
 $||X||^2 = X^T X$. (2.3)

Définition : On appelle matrice orthogonale toute matrice carrée $n \times n$ telle que

$$O^T O = I \iff O^{-1} = O^T \iff OO^T = I, \qquad (2.4)$$

En écrivant explicitement ces conditions d'orthonormalité, on obtient

$$\sum_{k} O_{ki} O_{kj} = \delta_{ij} \qquad \sum_{\ell} O_{i\ell} O_{j\ell} = \delta_{ij}$$
 (2.5)

qui expriment que les vecteurs colonnes d'une part, les vecteurs lignes de l'autre, de la matrice O sont orthonormés.

Exemples. Vérifier que les matrices $\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ et $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} \end{pmatrix}$ sont or-

thogonales. Plutôt que d'écrire le produit matriciel $O.O^T$, il suffit de calculer (mentalement

ou avec un crayon !) le produit scalaire de chaque colonne (ou ligne) avec elle-même et avec les autres.

Proposition 1 : Les transformations par des matrices orthogonales laissent invariant le produit scalaire.

En effet si X' = OX, Y' = OY, alors $X'^TY' = X^TO^TOY = X^TY$. En particulier, si X et Y sont deux vecteurs orthogonaux, leurs transformés par une matrice orthogonale X' = OX, Y' = OY le sont aussi : $X^TY = 0 \Longrightarrow X'^TY' = 0$.

Dans l'espace euclidien à 3 dimensions, les transformations définies par ces matrices sont telles que $\overrightarrow{OM} = X \mapsto \overrightarrow{OM'} = X' = OX$ avec |OM| = |OM'|, ce sont des isométries et on démontre que ce sont des rotations ou des réflexions par rapport à un plan, ou leurs composées (leurs produits). Le produit de deux matrices orthogonales est une matrice orthogonale (le vérifier !), ce qui signifie que la composée de deux isométries est une isométrie. La transformation $X \mapsto -X$ qui correspond à la matrice O = -I est le produit d'une rotation de π autour d'un axe quelconque Δ passant par l'origine O par la réflexion dans le plan orthogonal en O à Δ . Exercice : pour s'en convaincre, (1) faire la figure, (2) écrire les deux matrices qui effectuent ces transformations.

Proposition 2 : Tout changement de base orthonormée définit une matrice orthogonale. Réciproquement, toute matrice orthogonale transforme une base orthonormée en une autre base orthonormée.

Preuve : soient e_i et f_i deux bases orthonormées, $e_i.e_j = \delta_{ij}$, $f_i.f_j = \delta_{ij}$. Formons la matrice de changement de base, cf. equ. (2.4) du chapitre 1, $f_j = \sum_i e_i \Omega_{ij}$, Soient en utilisant les produits scalaires orthonormés : $\Omega_{ij} = e_i.f_j$. Mais selon l'observation faite en (2.2), Ω_{ij} représente la composante de e_i sur f_j (ou vice versa), et selon (2.5), l'orthonormalité de ces composantes est équivalente à l'orthogonalité de la matrice Ω , cqfd. La réciproque découle de la Proposition 1.

Remarque : en géométrie usuelle dans l'espace euclidien à 3 dimensions, cette proposition dit simplement qu'on passe d'un repère orthonormé à un autre par une rotation ou une "rotation-réflexion". Qu'en est-il de l'espace euclidien à deux dimensions ?

Proposition 3 : Tout changement de base par une matrice orthogonale transforme une matrice symétrique en matrice symétrique.

La preuve est immédiate : si $B = B^T$, pour toute matrice orthogonale O, $B' = O^T B O$ satisfait bien $B'^T = (O^T B O)^T = O^T B^T O = O^T B O = B'$ cqfd.

2.4. Diagonalisation d'une matrice symétrique

Supposons qu'on a dans E une autre forme bilinéaire symétrique f. (On ne la suppose pas définie positive.) Elle est définie par la donnée dans la base e_i (orthonormée pour le produit scalaire euclidien) de la matrice symétrique B, $B_{ij} = f(e_i, e_j)$.

On démontre alors l'important théorème

Théorème 2 : Toute matrice symétrique réelle B peut se diagonaliser par un changement de base orthogonal

$$\exists O \qquad B = O\Lambda O^T , \qquad \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \qquad \lambda_i \in \mathbb{R} .$$
 (2.6)

Autrement dit, toute matrice symétrique réelle B possède n valeurs propres réelles et un système de n vecteurs propres orthonormés.

Établissons d'abord le

Lemme 2 : Si B est une matrice symétrique réelle, (i) deux vecteurs propres de B correspondant à deux valeurs propres distinctes sont orthogonaux ; (ii) B n'a que des valeurs propres réelles.

Les deux propriétés du Lemme découlent du même argument :

(i) Soient λ et μ deux valeurs propres distinctes de B, de vecteurs propres respectifs X et Y. On a donc $BX = \lambda X$, $BY = \mu Y$. Calculons de deux façons Y^TBX

$$Y^T B X = \lambda Y^T X$$

= $(Y^T B X)^T = X^T B^T Y = X^T B Y = \mu X^T Y = \mu Y^T X$,

(où on a utilisé le fait que le nombre Y^TBX , considéré comme un matrice 1×1 , est égal à son transposé, et il en est de même de Y^TX). Il en découle que $(\lambda - \mu)Y^TX = 0$, donc puisque $\lambda \neq \mu$, $Y^TX = 0$, les vecteurs propres X et Y sont orthogonaux (pour le produit scalaire euclidien (2.1)).

(ii) Soit λ une valeur propre de B pour le vecteur propre X. Il se pourrait que λ et X soient complexes. Écrivant $BX = \lambda X$ et sa conjuguée (B est réelle) $BX^* = \lambda^* X^*$, on voit qu'on est dans les conditions du point (i) précédent, mais cette fois X et X^* ne peuvent être orthogonaux, car $XX^* = \sum_{i=1}^n |x_i|^2$, somme des modules carrés des composantes de X, qui ne peut s'annuler que si X = 0. On en conclut que $\lambda \neq \lambda^*$ est impossible, donc que λ est réelle.

Preuve du Théorème. Démonstration par récurrence sur la dimension n de la matrice (ou de l'espace vectoriel). Pour n=1, le théorème est évident : dans l'espace de dimension 1, le vecteur de base peut être normé. Supposons le théorème vrai pour tout espace de dimension $\leq n-1$. Dans l'espace E de dimension n, doté d'une base e_i orthonormée pour le produit scalaire euclidien (2.1), B a au moins un vecteur propre X_1 de valeur propre λ_1 , que nous supposerons normé. Par le lemme, λ_1 et X_1 sont réels, et $X_1 = \sum_i x_i e_i$. Supposons par exemple que $x_1 \neq 0$. Le système de vecteurs $\{X_1, e_2, \cdots, e_n\}$ est une base, et par le procédé d'orthonormalisation de Schmidt, on peut construire une base orthonormée $X_1, \tilde{e}_2, \cdots \tilde{e}_n$. Soit F le sous-espace engendré par les vecteurs $\tilde{e}_2, \cdots \tilde{e}_n$. C'est l'espace orthogonal au vecteur propre X_1 . Montrons que B laisse le sous-espace F invariant. En effet pour tout $j=2,\cdots,n$, on a $X_1.B\tilde{e}_j=\tilde{e}_j.BX_1=\lambda_1\tilde{e}_j.X_1=0$, $B\tilde{e}_j$ est orthogonal à X_1 donc appartient au sous-espace F. B est représentée dans F par une matrice symétrique (Proposition 3 du § 2.3). On peut maintenant appliquer à F, espace de dimension n-1, l'hypothèse de récurrence : il possède une base orthonormée de vecteurs propres $X_2, \cdots X_n$ de B. Par la Proposition 2 du § 2.3, on passe de la base initiale e_i à cette nouvelle base X_1, X_2, \cdots, X_n par un changement de base orthogonal, cqfd.

2.5. Réduction d'une forme quadratique

Soient Q une forme quadratique, B la matrice réelle symétrique qui la représente dans une base,

$$Q(X) = \sum_{ij} B_{ij} x_i x_j .$$

Selon le théorème 2 du § précédent, on peut trouver un changement de base par une matrice orthogonale qui diagonalise la matrice $B: B = O\Lambda O^T$, où O a pour vecteurs-colonnes les vecteurs propres (orthonormés) de B, cf chapitre 4, § 2.2. Récrivant $Q(X) = X^T B X = X^T O\Lambda O^T X$, on voit que le changement de coordonnées X = OX', soit $x_i = \sum_j O_{ij} x'_j$, diagonalise la forme quadratique $Q: Q = X'^T \Lambda X'$, c'est-à-dire l'exprime comme somme de carrés avec comme coefficients les valeurs propres de B.

Théorème 3 : Pour toute forme quadratique $Q(X) = \sum_{ij} B_{ij} x_i x_j$, il existe un changement de coordonnées orthogonal $x'_i = \sum_j O_{ji} x_j$ qui la diagonalise

$$Q(X) = \sum_{i=1}^{n} \lambda_i x_i^{'2} = \sum_{i=1}^{n} \lambda_i (\sum_j O_{ji} x_j)^2 .$$
 (2.7)

On appelle cette expression la $forme \ r\'eduite$ de Q.

Corollaire 1 : Une forme quadratique est définie positive (resp. semi-définie positive) ssi les valeurs propres de sa matrice sont strictement positives (resp. positives ou nulles). Elle est indéfinie si sa matrice possède des valeurs propres des deux signes.

Exemples. Au vu de ce corollaire, il n'est pas difficile de voir que dans l'espace de dimension 3, la forme $Q_1(X) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_2x_3$ est définie positive, que $Q_2(X) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_2x_3 - x_1x_3$ est semi-définie positive, et que $Q_3(X) = x_1^2 - 2x_2x_3$ est indéfinie. Vérifier que les valeurs propres de leurs matrices sont respectivement $B_1: \{1 \pm \frac{\sqrt{2}}{2}, 1\}, B_2: \{\frac{3}{2}, \frac{3}{2}, 0\}$ et $A_3: \{1, 1, -1\}$ et que leurs formes réduites s'écrivent effectivement

$$\begin{split} Q_1(X) &= \frac{2+\sqrt{2}}{8}(x_1-\sqrt{2}x_2+x_3)^2 + \frac{1}{2}(x_1-x_3)^2 + \frac{2-\sqrt{2}}{8}(x_1+\sqrt{2}x_2+x_3)^2 \\ Q_2(X) &= \frac{3}{2}\left(\frac{x_1-x_3}{\sqrt{2}}\right)^2 + \frac{3}{2}\left(\frac{-x_1+2x_2-x_3}{\sqrt{6}}\right)^2 \\ Q_3(X) &= x_1^2 + \left(\frac{x_2-x_3}{\sqrt{2}}\right)^2 - \left(\frac{x_2+x_3}{\sqrt{2}}\right)^2 \;. \end{split}$$

2.6. Diagonalisation simultanée de deux matrices symétriques commutantes

Le théorème 2 du paragraphe § 2.4 possède encore un autre corollaire intéressant.

Corollaire 2 : Si A et B sont deux matrices réelles symétriques qui commutent, AB = BA, on peut les diagonaliser simultanément par un même changement de base orthogonal.

Noter que le théorème 2 en est un cas particulier, quand A = I, la matrice identité. Noter aussi que ce résultat prolonge (pour des matrices symétriques) celui obtenu à la Prop. 7 du Chap. 4. Ici on n'a pas besoin de supposer les valeurs propres de l'une des matrices distinctes.

Avant de prouver ce Corollaire, démontrons le

Lemme 3: Si A et B commutent, un espace propre de A est invariant par B.

En effet si X est un vecteur propre de A, $AX = \lambda X$, et $ABX = BAX = \lambda BX$. Ou bien BX = 0, ou bien il est vecteur propre de A. Dans les deux cas, il appartient à l'espace propre.

Preuve du corollaire : on diagonalise d'abord A selon le théorème. B laisse invariant tout sousespace propre F de A, c'est une matrice symétrique dans F, on peut donc l'y diagonaliser toujours selon le théorème par un nouveau changement de base orthogonal (qui n'affecte pas la diagonalité de A). En procédant ainsi dans chaque sous-espace propre de A, on construit une base orthonormale dans laquelle les deux matrices sont diagonales, cqfd.

3. Extension aux formes sesquilinéaires et matrices hermitiennes

Toute la discussion qui précède se généralise à des formes à valeurs complexes dites formes sesquilinéaires et aux matrices "hermitiennes" qui leur sont associées. Cette situation est importante pour le physicien en particulier dans les applications à la physique quantique. Nous nous bornerons ici à de brèves indications.

Si E désigne maintenant un espace vectoriel sur les nombres complexes (c'est-à-dire les combinaisons linéaires de vecteurs peuvent se faire avec des nombres complexes), on définit une forme sesquilinéaire f(X,Y) comme une forme (une application dans les nombres complexes) qui est linéaire dans son second argument Y, mais "antilinéaire" dans le premier X. Ce qui remplace (1.1) est donc

$$f(\lambda_1 X_1 + \lambda_2 X_2, Y) = \lambda_1^* f(X_1, Y) + \lambda_2^* f(X_2, Y)$$

$$f(X, \mu_1 Y_1 + \mu_2 Y_2) = \mu_1 f(X, Y_1) + \mu_2 f(X, Y_2) ,$$
 (3.1)

où λ^* est le complexe conjugué de μ . La forme est dite hermitienne si $f(Y,X) = f(X,Y)^*$.

Exemples. (1) Considérons l'espace \mathbb{C}^n des vecteurs X à n composantes x_i complexes. La forme $\sum_i x_i^* y_i$ est une forme sesquilinéaire hermitienne. (2) Soit à nouveau g et h des fonctions d'une variable réelle $x \in (a,b)$, mais cette fois à valeurs complexes (par exemple e^{ix}). L'expression $\int_a^b g(x)^* h(x) dx$ est une forme sesquilinéaire hermitienne.

On définit comme précédemment la forme Q(X) = f(X, X) associée à une forme sesquilinéaire, qui peut être selon les cas définie positive, semi-définie positive, indéfinie, etc. On a à nouveau une inégalité de Schwarz, qui s'exprime de la même façon. Dans une base, une forme hermitienne s'exprime à l'aide d'une matrice $A = (a_{ij})$, $a_{ij} = f(e_i, e_j)$, et la matrice est hermitienne ce qui signifie que

$$A = A^{T*} \Longleftrightarrow a_{ij} = a_{ji}^* . {3.2}$$

Si $X = \sum_i x_i e_i$ et $Y = \sum_j y_j e_j$, $f(X,Y) = \sum_{ij} x_i^* y_j f(e_i,e_j) = \sum_{ij} x_i^* y_j a_{ij}$. On définit en général la conjugaison hermitique d'une matrice B par

$$B^{\dagger} = B^{T*} . \tag{3.3}$$

Les matrices hermitiennes sont donc les matrices égales à leur conjuguée hermitique. On démontre que leurs valeurs propres sont toutes réelles, c'est l'analogue du Lemme 2 du § 2.4.

Par ailleurs on définit les matrices unitaires. Ce sont les matrices carrées U à éléments complexes qui satisfont

$$U^{\dagger}U = I \Longleftrightarrow U^{-1} = U^{\dagger} . \tag{3.4}$$

On démontre alors le théorème

Théorème 4: Toute matrice hermitienne peut se diagonaliser par un changement de base unitaire

$$A = U\Lambda U^{\dagger}$$

qui implique le

Corollaire 3 : Toute forme sesquilinéaire peut se mettre sous forme diagonale par un changement de base unitaire,

$$f(X,Y) = X^{\dagger}AY = \tilde{X}^{\dagger}\Lambda\tilde{Y} = \sum_{i} \lambda_{i}\tilde{x}_{i}^{*}\tilde{y}_{i} \qquad \tilde{X} = U^{\dagger}X, \ \tilde{Y} = U^{\dagger}Y$$

et la forme est définie positive ssi les valeurs propres sont toutes (réelles) positives.

Finalement le corollaire de la fin du paragraphe précédent admet une extension : si A et B sont deux matrices hermitiennes qui commutent, elles peuvent être diagonalisées simultanément.

Cette dernière propriété joue un rôle fondamental en mécanique quantique : toute quantité observable y est représentée par un opérateur (une matrice, dans une base) hermitien(ne). Ses valeurs propres sont les valeurs que peut donner la mesure de cette observable. Le fait que deux "observables commutent" signifie qu'on peut les mesurer simultanément.

Récapitulons le parallèle entre formes bilinéaires et formes sesquilinéaires

forme bilinéaire symétrique f(X,Y) \leftrightarrow forme sesquilinéaire hermitienne f(X,Y)

forme quadratique Q(X) = f(X,X) \leftrightarrow norme carrée réelle de vecteurs complexes $Q(X) = f(\overline{X},X)$

 $\begin{array}{lll} \text{matrice symétrique} & A = A^T & \leftrightarrow \text{ matrice hermitienne} & A = A^\dagger \\ \text{matrice orthogonale} & O^{-1} = O^T & \leftrightarrow \text{ matrice unitaire} & U^{-1} = U^\dagger \end{array}$

diagonalisation par une matrice orthogonale \leftrightarrow diagonalisation par une matrice unitaire .

Toutes ces propriétés, étendues éventuellement à des espaces de dimension infinie, seront très utiles en mécanique quantique...

4. Applications physiques

4.1. Tenseur d'inertie

On connaît du cours de Mécanique les notions de moment cinétique \vec{J} et de moment d'inertie I d'un point matériel M de masse m en rotation de vitesse angulaire ω autour d'un axe \vec{u} . Si r est la distance du point M à l'axe \vec{u} , sa vitesse est $v = r\omega$, (ou mieux $\vec{v} = \omega \vec{u} \wedge \vec{r}$), son moment cinétique (par rapport à \vec{u}) est $J_u = mvr = mr^2\omega$, et son moment d'inertie (par rapport à \vec{u}) est $I_u = mr^2$, tel que

$$J_u = I_u \omega \,. \tag{4.1}$$

Ce moment d'inertie (et J_u) s'annule seulement si la masse m est sur l'axe \vec{u} . L'énergie cinétique de rotation est $T = \frac{1}{2}mv^2 = \frac{1}{2}I_u\omega^2 = \frac{1}{2}\frac{J_u^2}{I_u}$. Généralisons maintenant ces formules au cas d'un solide en rotation.

Soit un corps solide indéformable dont la distribution de masse est décrite par une densité $\rho(\vec{r})$ dans un certain volume V. Quand ce corps a un mouvement de rotation de vitesse angulaire ω autour d'une direction \vec{u} passant par un point O, le point M de coordonnées $\vec{r} = \overrightarrow{OM}$ a une vitesse $\vec{v} = \omega \vec{u} \wedge \vec{r}$ et le vecteur moment cinétique \vec{J} par rapport à O est la superposition des contributions $\overrightarrow{OM} \wedge \rho \vec{v}$ sommées sur tout le solide :

$$\vec{J} = \int d^3r \rho(\vec{r}) \vec{r} \wedge (\vec{\omega} \wedge \vec{r}). \tag{4.2}$$

C'est donc un vecteur dont les composantes sont des fonctions linéaires de celles de $\vec{\omega} = \omega \vec{u}$, ce qui généralise (4.1) et amène à définir le "tenseur d'inertie" I (ne pas confondre avec la matrice identité!)

$$\vec{J} = I.\vec{\omega} \,, \tag{4.3}$$

c'est-à-dire $J_i = \sum_j I_{ij}\omega_j$. En utilisant la formule du double produit vectoriel $\vec{o} \ \vec{a} \wedge (\vec{b} \wedge \vec{c}) = \vec{b}(\vec{a}.\vec{c}) - \vec{c}(\vec{a}.\vec{b})$, on calcule $J^i = \int d^3r \rho(\vec{r}) \sum_{j=1}^3 (r^2 \delta_{ij} - r^i r^j) \omega^j$, et le tenseur d'inertie I est décrit par la matrice symétrique 3×3

$$I_{ij} = \int d^3r \rho(\vec{r})(r^2 \delta_{ij} - r^i r^j) , \qquad (4.4)$$

ou encore,

$$I = \int dx \, dy \, dz \, \rho(x, y, z) \begin{pmatrix} y^2 + z^2 & -xy & -xz \\ -yx & x^2 + z^2 & -yz \\ -zx & -zy & x^2 + y^2 \end{pmatrix} = \begin{pmatrix} A & -F & -E \\ -F & B & -D \\ -E & -D & C \end{pmatrix} . \quad (4.5)$$

(Les notations A, B, \dots, F sont traditionnelles.)

Noter que le moment cinétique n'est en général pas colinéaire au vecteur rotation $\vec{\omega}$. Mais la matrice I_{ij} est symétrique, donc diagonalisable par changement de coordonnées orthogonal, (Théorème 2). Dans les nouveaux axes (OX, OY, OZ), $I = \text{diag}(I_X, I_Y, I_Z)$. Le long de ces "axes principaux d'inertie", le moment cinétique est colinéaire au vecteur rotation.

L'énergie cinétique de rotation se calcule aussi aisément. Comme $\vec{v}^2 = (\vec{\omega} \wedge \vec{r})^2 = \vec{\omega} \cdot (\vec{r} \wedge (\vec{\omega} \wedge \vec{r}))$

$$T = \int dx \, dy \, dz \, \frac{1}{2} \rho(x, y, z) \vec{v}^2 = \frac{1}{2} \vec{\omega} . \vec{J} = \frac{1}{2} \vec{\omega} . I . \vec{\omega} .$$

L'énergie cinétique de rotation est donc donnée par la forme quadratique Q associée à I. Cette forme est évidemment définie positive, en tant qu'énergie cinétique. Cela peut se voir aussi d'autres façons : les valeurs propres I_X, I_Y, I_Z sont les "moments d'inertie principaux" par rapport aux axes (OX, OY, OZ) et sont donc positifs. Enfin cette propriété résulte de l'inégalité de Schwarz. En effet, pour tout vecteur \vec{w} , $Q(\vec{w}) = \vec{w}.I.\vec{w} = \int d^3r \rho(\vec{r})(r^2w^2 - (\vec{r}.\vec{w})^2)$; or $r^2w^2 - (\vec{r}.\vec{w})^2 \geq 0$ avec égalité seulement si \vec{r} et \vec{w} sont colinéaires, (cf (1.6)), donc la somme $\int d^3r \rho(\vec{r})(r^2w^2 - (\vec{r}.\vec{w})^2)$ avec $\rho \geq 0$ est aussi ≥ 0 (et ne s'annulerait que pour une distribution linéaire de masses le long de la direction \vec{w}).

⁵ que l'on vérifie par le calcul direct : $\vec{b} \wedge \vec{c} = (b_y c_z - b_z c_y, \cdots), \ a \wedge (\vec{b} \wedge \vec{c}) = (a_y (b_x c_y - b_y c_x) - a_z (b_z c_x - b_x c_z), \cdots) = (b_x (a_x c_x + a_y c_y + a_z c_z) - c_x (a_x b_x + a_y b_y + a_z b_z), \cdots)$

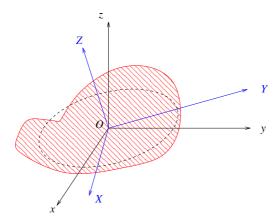


Fig. 14: Ellipsoïde d'inertie en O (ligne brisée) et axes principaux d'inertie.

Soient α, β, γ les coordonnées du vecteur \vec{w} . La forme quadratique s'écrit $Q(\vec{w}) = A\alpha^2 + B\beta^2 + C\gamma^2 - 2D\beta\gamma - 2E\gamma\alpha - 2F\alpha\beta$ et l'équation $Q(\vec{w}) = 1$ définit une surface dans l'espace à trois dimensions. C'est un ellipsoïde, dit *ellipsoïde d'inertie au point O*. Les axes principaux d'inertie sont les axes de cet ellipsoïde, voir figure 12.

Toutes ces notions sont importantes dans l'étude de la rotation du corps considéré et de la stabilité de ce mouvement de rotation.

Comment ce tenseur I est-il modifié si on change le point de référence $O \mapsto O'$? En particulier si on choisit le centre de gravité G? Se rappeler que par définition de G, si $\vec{r}' = \overline{GM}$, $\int d^3r' \rho(\vec{r}')\vec{r}' = 0$, tandis que $\int d^3r' \rho(\vec{r}') = M$, la masse du corps. Réponse : soit $\vec{g} = \overline{OG}$, alors $I_O = I_G + M(g^2\delta_{ij} - g_ig_j)$.

Soit I_G le tenseur d'inertie par rapport à son centre de gravité G dont la vitesse (instantanée) est \vec{V} . L'énergie cinétique totale s'écrit $T=\int d^3r\frac{1}{2}\rho(\vec{r})\vec{v}^2$ mais $\vec{v}(\vec{r})=\vec{V}+\vec{\omega}\wedge\vec{r}$ donc $\vec{v}^2=V^2+2\vec{V}.(\vec{\omega}\wedge\vec{r})+(\vec{\omega}\wedge\vec{r})^2$. Le terme "croisé" $\vec{V}.(\vec{\omega}\wedge\vec{r})$ s'annule par intégration sur \vec{r} , (puisqu'à nouveau, pour le centre de gravité, $\int d^3r\rho(\vec{r})\vec{r}=0$)), et $(\vec{\omega}\wedge\vec{r})^2=\vec{\omega}.(\vec{r}\wedge(\vec{\omega}\wedge\vec{r}))$, comme plus haut, d'où $\int d^3r\rho(\vec{r})(\vec{\omega}\wedge\vec{r})^2=\vec{\omega}.I_G.\vec{\omega}$. Finalement l'énergie est la somme de l'énergie cinétique de translation du centre de gravité et de l'énergie cinétique de rotation autour de G

$$T = \frac{1}{2}M\vec{V}^2 + \frac{1}{2}\vec{\omega}.I_G.\vec{\omega}. \tag{4.6}$$

4.2. Tenseur de conductivité

Dans un conducteur à trois dimensions (et non plus nécessairement filaire), la loi de proportionnalité des intensités aux différences de potentiel (loi d'Ohm), ou des densités de courant \vec{j} au champ électrique \vec{E} , prend une forme matricielle $j=\sigma E$, soit en composantes

$$\begin{pmatrix} j_1 \\ j_2 \\ j_3 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix} .$$

On admettra que le tenseur σ est symétrique $\sigma_{ij} = \sigma_{ji}$. Selon le théorème 2 de diagonalisation, il existe un repère orthonormé de coordonnées tel que σ soit diagonal. Dans ce repère, les nouvelles composantes de \vec{j} sont proportionnelles à celles de \vec{E} . On retrouve la loi d'Ohm habituelle, mais avec en général des conductivités différentes dans les différentes directions!

4.3. Stabilité des extrema d'un potentiel

Considérons un système mécanique à N degrés de liberté. Ce peut être un système de n particules massives en interaction dans l'espace \mathbb{R}^3 , chacune dotées de 3 coordonnées x, y, z, auquel cas le nombre de degrés de liberté est N=3n; ce peut être un solide indéformable dont la position dans l'espace \mathbb{R}^3 est repérée par les trois coordonnées de son centre de gravité G et par trois angles de rotation autour de G, d'où N=6, etc. On appellera $x_i, i=1,\cdots,N$ les coordonnées de ces N degrés de liberté.

Supposons la dynamique de ce système décrite par un potentiel $V(x_1, \dots, x_N)$. Cela implique que pour chaque degré de liberté, l'équation de la dynamique s'écrit $m_i \ddot{x}_i = -\frac{\partial V(x)}{\partial x_i}$. Les points stationnaires du système sont les extrema $x^{(0)}$ du potentiel (force nulle)

$$\left. \frac{\partial V(x)}{\partial x_i} \right|_{x^{(0)}} = 0 \ . \tag{4.7}$$

La question est de savoir lesquels de ces extrema correspondent à des positions d'équilibre stable. Pour cela on effectue un développement limité au deuxième ordre de V(x) au voisinage de l'un des $x^{(0)}$ en posant $x_i - x_i^{(0)} = \xi_i$. Selon la formule de Taylor à plusieurs variables

$$V(x) = V(x^{(0)}) + \frac{1}{2} \sum_{i=1}^{N} \xi_i^2 \frac{\partial^2 V(x)}{\partial x_i^2} \Big|_{x^{(0)}} + \sum_{1 \le i \le j \le N} \xi_i \xi_j \frac{\partial^2 V(x)}{\partial x_i \partial x_j} \Big|_{x^{(0)}} + \cdots$$
(4.8)

où le terme d'ordre 1 est absent en vertu de (4.7). Au terme constant $V(x^{(0)})$ près, qui ne joue aucun rôle dans la discussion, on voit que dans cette approximation de petits écarts ξ à l'extremum $x^{(0)}$, le potentiel est une forme quadratique dans les N variables ξ_i . Écrivons sa forme réduite (au sens du § 2.5)

$$V(x) = V(x^{(0)}) + \frac{1}{2} \sum_{i=1}^{N} \lambda_i \tilde{\xi}_i^2$$
(4.9)

avec λ_i les valeurs propres de B et $\tilde{\xi}_i$ les modes propres. La question du signe des λ_i va être cruciale :

– s'ils sont tous strictement positifs, c'est-à-dire si la forme quadratique est définie positive, (cf Théorème 3 et Corollaire), l'extremum $x^{(0)}$ est un minimum local: au voisinage de $x^{(0)}$, V(x) est strictement supérieur à $V(x^{(0)})$. Posons $B_{ij} = \frac{\partial^2 V(x)}{\partial x_i^2}\Big|_{x^{(0)}}$: les ξ_i satisfont $\ddot{\xi}_i + \sum_j B_{ij} \xi_j = 0$, soit des équations d'oscillateurs harmoniques couplés, comme

le montre mieux encore le passage aux modes propres $\tilde{\xi}_i$: $\dot{\tilde{\xi}}_i + \omega_i^2 \tilde{\xi}_i = 0$, où $\lambda_i = \omega_i^2$ est la valeur propre de B correspondant au mode $\tilde{\xi}_i$. Le système effectue donc de petites oscillations au voisinage de $x^{(0)}$, l'équilibre est **stable**;

– si en revanche la forme possède une ou plusieurs valeurs propres négatives ou nulles, le système est instable. Pour chaque $\tilde{\xi}_i$ de valeur propre $\lambda_i \leq 0$, l'équation de la dynamique est $\ddot{\tilde{\xi}}_i = \lambda_i \tilde{\xi}_i$, avec une force de rappel nulle ou négative : il n'y a plus d'oscillation, il y a même croissance de $\tilde{\xi}_i$ si $\lambda_i < 0$. L'extremum est un point d'équilibre **instable**.

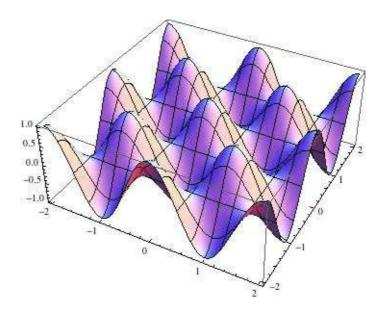


Fig. 15: Le potentiel $V(x,y) = \cos(\pi x)\cos(\pi y)$ avec ses extrema périodiques

Exemple. Ceci est illustré sur un système à deux degrés de liberté x et y, de potentiel $V(x,y)=\cos(\pi x)\cos(\pi y)$. Ce potentiel est représenté sur la figure 15. On lui donne le nom de "boîte à œufs" pour des raisons évidentes.... Il est aisé de voir sur la figure et le calcul confirme que le potentiel présente des extrema aux points x,y tous deux entiers ou tous deux demi-entiers. Seuls les points à x et y entiers de parité opposée (x=0,y=1 par exemple) sont des minima du potentiel, les points où x et y sont entiers de même parité sont des maxima, les points où ils sont tous deux demi-entiers sont des "cols", avec une valeur propre positive et une négative. Tout cela est bien en accord avec notre intuition : une petite bille lâchée dans ce potentiel oscille au fond d'un creux...

On verra en TD un autre exemple.