Université Paris-Saclay

Mathématiques : Mise à niveau I333a

Notes de cours pour la 3e année de licence E3A

Responsable de l'UE :
Thomas Colas
Institut d'Astrophysique Spatiale
Université Paris-Saclay
thomas.colas@ias.u-psud.fr

Auteure principale :
Corinne Donzaud

Compléments :
Thomas Colas

2020-2021

Table des matières

Première partie

Algèbre linéaire

Chapitre 1

Calcul matriciel

Dans ce chapitre l'ensemble $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Un élément de $\mathbb{K}$ est appelé scalaire.

1.1 Matrices

1.1.1 Définitions

Etant donnés deux entiers n et p strictement positifs, une matrice à n lignes et p colonnes est un tableau rectangulaire de scalaires (réels ou complexes). L'indice de ligne i va de 1 à n, l'indice de colonne j va de 1 à p.

$$A = (a_{i,j}) = \begin{pmatrix} a_{11} & ... & a_{1j} & ... & a_{1p} \\ ... & & & & \\ ... & & & & \\ a_{i1} & ... & a_{ij} & ... & a_{ip} \\ ... & & & & \\ ... & & & & \\ a_{n1} & ... & a_{nj} & ... & a_{np} \end{pmatrix} \tag{1.1}$$

Les entiers n et p sont les dimensions de la matrice ; $a_{ij} \in \mathbb{K}$ est son coefficient d'ordre (i, j).

L'ensemble des matrices à n lignes et p colonnes et à coefficients dans $\mathbb{K}$ est noté $\mathcal{M}_{np}(\mathbb{K})$.

La matrice est dite nulle si tous ses coefficients sont nuls.

Toute matrice réduite à une ligne $\begin{pmatrix} a_1 & a_2 & ... & a_p \end{pmatrix}$ est appelée matrice-ligne. Elle peut être identifiée avec le $p-uplet$ $(a_1, a_2, ..., a_p)$ de $\mathbb{K}^p$.

Toute matrice réduite à une colonne $\begin{pmatrix} a_1 \\ a_2 \\ ... \\ a_n \end{pmatrix}$ est appelée matrice-colonne. Elle peut être identifiée avec le vecteur dont les coordonnées, dans une certaine base, sont $a_1, a_2, ..., a_n$.

Matrices carrées

Une matrice carré d'ordre n est telle que $n = p$. On note l'ensemble de ces matrices $\mathcal{M}_n(\mathbb{K})$.

Une matrice est dite diagonale si tous les coefficients hors de la diagonale sont nuls, la diagonale étant définie par les coefficients a_{ii}, $i = 1, .., n$.

La matrice identité d'ordre n est la matrice diagonale dont tous les coefficients a_{ii} valent 1. Cette matrice est notée I_n, elle est définie, en utilisant la notation de Kronecker, par $a_{ij} = \delta_{ij}$, $i = 1, .., n$, $j = 1, ..., n$.

Une matrice est dite triangulaire supérieure (inférieure) si tous les coefficients en-dessous (en-dessus) de la diagonale sont nuls.

Une matrice $A = (a_{ij})$ de $\mathcal{M}_{np}(\mathbb{K})$ est symétrique si $\forall (i,j) \in \mathbb{K}^2, a_{ij} = a_{ji}$. Elle est antisymétrique si $a_{ij} = -a_{ji}$. Dans ce dernier cas, les coefficients de la diagonale sont nuls.

1.1.2 Opérations élémentaires

On peut ajouter deux matrices de même dimension et multiplier une matrice par un scalaire : le résultat est une matrice de même dimension [1].

Addition de deux matrices

Si $A = (a_{ij})$ et $B = (b_{ij})$ sont deux matrices de $\mathcal{M}_{np}(\mathbb{K})$, leur somme $A + B$ est la matrice $(a_{ij} + b_{ij})$. Par exemple :

$$\begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 1 & -1 \end{pmatrix} + \begin{pmatrix} -3 & 1 \\ 5 & -3 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 7 & 0 \\ 1 & 1 \end{pmatrix}$$

Multiplication d'une matrice par un scalaire

Si $A = (a_{ij})$ est une matrice de $\mathcal{M}_{np}(\mathbb{K})$ et λ est un nombre de l'ensemble $\mathbb{K}$ (réel ou complexe), le produit λA est la matrice (λa_{ij}). Par exemple :

$$-2 \begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} -2 & -2 \\ -4 & -6 \\ -2 & 2 \end{pmatrix}$$

1.1.3 Produit matriciel

Définition 1 *Soient $A = (a_{ij})$ une matrice de $\mathcal{M}_{np}(\mathbb{K})$ et $B = (b_{ij})$ une matrice de $\mathcal{M}_{pq}(\mathbb{K})$. On définit la matrice $C = AB$, élément de $\mathcal{M}_{nq}(\mathbb{K})$ avec des coeffcients c_{ij} tels que pour tout i de $1, ..., n$ et pour tout j de $1, ..., q$, $c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$.*

La multiplication entre matrices est définie en multipliant terme à terme les éléments d'une ligne de A par les éléments d'une colonne de B. Il est donc possible d'effectuer des produits de matrices de taille différente pour peu que le nombre de colonnes de la première coïncide avec le nombre de lignes de la seconde. Pour effectuer le produit, il est conseillé de placer B au-dessus et à droite de A.

Posons par exemple $A = \begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 1 & -1 \end{pmatrix}$ et $\mathrm{B} = \begin{pmatrix} 0 & 1 & -1 & -2 \\ -3 & -2 & 0 & 1 \end{pmatrix}$

$$A * B = \begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 1 & -1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & -1 & -2 \\ -3 & -2 & 0 & 1 \end{pmatrix} \\ \begin{pmatrix} -3 & -1 & -1 & -1 \\ -9 & -4 & -2 & -1 \\ 3 & 3 & -1 & -3 \end{pmatrix}$$

Proposition 1 *Soient $(\lambda, \mu) \in \mathbb{K}^2$:*

- *Si $A \in \mathcal{M}_{np}(\mathbb{K})$, $B \in \mathcal{M}_{pq}(\mathbb{K})$ et $C \in \mathcal{M}_{pq}(\mathbb{K})$, alors $A(\lambda B + \mu C) = \lambda AB + \mu AC$ (linéarité à droite).*
- *A et $B \in \mathcal{M}_{np}(\mathbb{K})$ et $C \in \mathcal{M}_{pq}(\mathbb{K})$, alors $(\lambda A + \mu B)C = \lambda AC + \mu BC$ (linéarité à gauche).*
- *$A \in \mathcal{M}_{np}(\mathbb{K})$, $B \in \mathcal{M}_{pq}(\mathbb{K})$ et $C \in \mathcal{M}_{qr}(\mathbb{K})$, alors $A(BC) = (AB)C$ (associativité).*

1. Cela résulte de la structure d'espace vectoriel de $\mathcal{M}_{np}(\mathbb{K})$.

Les produits AB et BA ne sont simultanément possibles que si $A \in \mathcal{M}_{np}(\mathbb{K})$ et $B \in \mathcal{M}_{pn}(\mathbb{K})$. Alors AB est carrée d'ordre n, tandis que BA est carrée d'ordre p.

Si A et B sont toutes les deux carrées d'ordre n, alors AB et BA sont carrées d'ordre n, mais on a en général $AB \neq BA$: le produit matriciel n'est pas commutatif. Dans le cas contraire, on dit que A et B commutent. Une façon d'exprimer cette propriété de non-commutativité (utilisée en particulier en mécanique quantique) consiste à introduire le commutateur $[A, B] \equiv AB - BA$. La relation de non-commutation s'écrit donc simplement sous la forme $[A, B] \neq 0$.

1.1.4 Transposition

Définition 2 *Soit $A = (a_{ij})$ une matrice de $\mathcal{M}_{np}(\mathbb{K})$. On appelle transposée de A et on note ${}^TA$ la matrice de $\mathcal{M}_{pn}(\mathbb{K})$ dont le coefficient d'ordre (i, j) est a_{ji}.*

Par exemple,

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 1 & -1 \end{pmatrix}, \quad {}^TA = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & -1 \end{pmatrix}$$

$$B = \begin{pmatrix} 0 & 1 & -1 & -2 \\ -3 & -2 & 0 & 1 \end{pmatrix}, \quad {}^TB = \begin{pmatrix} 0 & -3 \\ 1 & -2 \\ -1 & 0 \\ -2 & 1 \end{pmatrix}$$

Proposition 2
$\forall A \in M_{np}(\mathbb{K}) \quad {}^T({}^TA) = A$
$\forall (A, B) \in M_{np}(\mathbb{K})^2, \forall (\lambda, \mu) \in \mathbb{K}^2 \quad {}^T(\lambda A + \mu B) = \lambda\, {}^TA + \mu\, {}^TB$
$\forall (A, B) \in M_{np}(\mathbb{K})^2, {}^T(AB) = {}^TB\, {}^TA$ *(attention à l'ordre !)*

Illustration avec les exemples des matrices A et B définies ci-dessus :

$$AB = \begin{pmatrix} -3 & -1 & -1 & -1 \\ -9 & -4 & -2 & -1 \\ 3 & 3 & -1 & -3 \end{pmatrix}, \quad {}^T(AB) = \begin{pmatrix} -3 & -9 & 3 \\ -1 & -4 & 3 \\ -1 & -2 & -1 \\ -1 & -1 & -3 \end{pmatrix}$$

$$\begin{matrix} \\ {}^TB\, {}^TA = \end{matrix} \begin{pmatrix} 0 & -3 \\ 1 & -2 \\ -1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & -1 \end{pmatrix} = \begin{pmatrix} -3 & -9 & 3 \\ -1 & -4 & 3 \\ -1 & -2 & -1 \\ -1 & -1 & -3 \end{pmatrix}$$

Proposition 3

- *A est symétrique $\iff {}^TA = A$.*
- *A est antisymétrique $\iff {}^TA = -A$.*
- *$A\, {}^TA$ est une matrice carrée symétrique.*

En effet, ${}^T(A\, {}^TA) = {}^T({}^TA)\, {}^TA = A\, {}^TA$
Exemples :

$$A\, {}^TA = \begin{pmatrix} 2 & 5 & 0 \\ 5 & 13 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \quad {}^TAA = \begin{pmatrix} 6 & 6 \\ 6 & 11 \end{pmatrix}$$

1.2 Trace, déterminant et rang

Ces trois grandeurs sont des scalaires ($\in \mathbb{K}$). Trace et déterminant ne concernent que les matrices carrées.

1.2.1 Trace

Définition 3 *Soit A une matrice carrée d'ordre n, à coefficients dans $\mathbb{K}$, de terme général a_{ij}. On appelle trace de A, et on note $tr(A)$, la somme des coefficients diagonaux de A. Autrement dit,*

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$$

Proposition 4

Pour toutes matrices A et B de $\mathcal{M}_{np}(\mathbb{K})$ et pour tous scalaires $(\lambda, \mu) \in \mathbb{K}^2$:

$$\mathrm{tr}(\lambda A + \mu B) = \lambda \mathrm{tr}(A) + \mu \mathrm{tr}(B)$$

Proposition 5

Soient A une matrice de $\mathcal{M}_{np}(\mathbb{K})$ et B une matrice de $\mathcal{M}_{pn}(\mathbb{K})$. La matrice AB est carrée d'ordre n, celle de BA carrée d'ordre p. AB et BA ont la même trace :

$$\mathrm{tr}(AB) = \mathrm{tr}(BA)$$

En effet, soient c_{ij} le terme général de AB, et d_{ij} le terme général de BA, $\mathrm{tr}(AB) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}(\sum_{k=1}^{p} a_{ik}b_{ki})$ et $\mathrm{tr}(BA) = \sum_{i=1}^{p} d_{ii} = \sum_{i=1}^{p}(\sum_{k=1}^{n} b_{ik}a_{ki}) = \sum_{k=1}^{p}(\sum_{i=1}^{n} b_{ki}a_{ik})$.

Proposition 6 *Soient A, B et $C \in \mathcal{M}_n(\mathbb{K})$*

$$\mathrm{tr}(ABC) = \mathrm{tr}(CAB) = \mathrm{tr}(BCA)$$

mais attention $\mathrm{tr}(ABC) \neq \mathrm{tr}(CBA)$!

1.2.2 Déterminant

Définition

Nous nous contenterons d'une approche utilitaire.

Déterminant d'une matrice 2×2 :

Soit $A = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}$, on appelle déterminant de A le scalaire (réel ou complexe)

$$\det A \equiv \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1y_2 - y_1x_2$$

Déterminant d'une matrice 3×3 :

Soit $B = \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix}$, on appelle déterminant de B le scalaire

$$\begin{aligned} \mathrm{detB} &\equiv \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2 - z_1y_2x_3 - z_2y_3x_1 - z_3y_1x_2 \\ &= x_1 \begin{vmatrix} y_2 & y_3 \\ z_2 & z_3 \end{vmatrix} - y_1 \begin{vmatrix} x_2 & x_3 \\ z_2 & z_3 \end{vmatrix} + z_1 \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} \text{ dvpt suivant une colonne} \\ &= x_1 \begin{vmatrix} y_2 & y_3 \\ z_2 & z_3 \end{vmatrix} - x_2 \begin{vmatrix} y_1 & y_3 \\ z_1 & z_3 \end{vmatrix} + x_3 \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} \text{ dvpt suivant une ligne} \end{aligned}$$

Généralisation à une matrice n×n :

Les dernières factorisations montrent qu'un déterminant d'ordre 3 peut être d´eveloppé le long d'une colonne ou d'une ligne en faisant apparaître des déterminants d'ordre 2. Cette procédure est générale. On peut développer un déterminant d'ordre n selon une ligne ou une colonne en pondérant chaque déterminant d'ordre $(n-1)$ obtenu en supprimant la ligne et la colonne correspondante. Donnons deux définitions avant de donner la forme générale de la factorisation :

Définition 4 *Le déterminant mineur, noté $\mathcal{m}_{ij}$, d'un déterminant d'ordre n est le déterminant d'ordre $n-1$ obtenu en supprimant la $i^{ème}$ ligne et la $j^{ème}$ colonne de la matrice $A=(a_{ij})$.*

Définition 5 *Le cofacteur, noté $\mathcal{A}_{ij}$, est $\mathcal{A}_{ij}=(-1)^{i+j}\ \mathcal{m}_{ij}$.*

La procédure de factorisation s'écrit pour une matrice d'ordre n de la façon suivante :

Proposition 7

$$\det A=\sum_{j=1}^{n} a_{ij}\ \mathcal{A}_{ij} \text{ (dvpt selon la ligne } i)$$

$$\det A=\sum_{i=1}^{n} a_{ij}\mathcal{A}_{ij} \text{ (dvpt selon la colonne } j)$$

Exemple :

$$\begin{vmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & -2 & 1 \end{vmatrix} = 2\begin{vmatrix} 3 & -1 \\ -2 & 1 \end{vmatrix} - 0\begin{vmatrix} -1 & 1 \\ -2 & 1 \end{vmatrix} + 2\begin{vmatrix} -1 & 1 \\ 3 & -1 \end{vmatrix} = 2\times 1+2\times -2=-2$$

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

Voici une autre méthode pour calculer le déterminant, qui utilise les pivots d'une matrice échelonnée (voir la méthode de Gauss au paragraphe 2.3) :

Proposition 8 *Le déterminant est égal au produit des pivots non nuls, multiplié par $(-1)^r$ où r est le nombre de changements de lignes effectué.*

Cette méthode est celle qui demande le moins d'opérations ($2n^2/3$ pour un déterminant d'ordre n alors que la méthode des cofacteurs en demande $n!$. Soit pour une matrice d'ordre 25, 10^4 opérations à comparer à 1.5×10^{25} !). C'est donc cette dernière méthode qui est utilisée par les ordinateurs.

Propriétés

Proposition 9 *Soit A et $B\in\mathcal{M}_n(\mathbb{K})$,* $\det(AB)=\det A\times\det B$

mais attention, $\det(A+B)\neq\det A+\det B$

La proposition 7 qui nous sert de définition du déterminant montre que les mots ligne et colonne peuvent être remplacés l'un par l'autre dans l'énoncé des propriétés du déterminant.

Proposition 10 $\det A=\det{}^T A$

Proposition 11

- *Si une ligne (ou une colonne) ne comprend que des éléments nuls, alors le déterminant est nul.*
- *Si tous les éléments d'une ligne (ou d'une colonne) sont multipliés par la même constante (non nulle), alors le déterminant est multiplié par cette constante.*
- *Si l'on ajoute à une ligne (ou à une colonne) une combinaison linéaire des autres lignes (ou colonnes), alors le déterminant n'est pas modifié.*
- *Si l'on échange deux lignes (ou deux colonnes), alors le déterminant est multiplié par (-1).*
- *Pour une matrice triangulaire supérieure (ou inférieure), le déterminant est le produit des éléments diagonaux.*

Une conséquence importante de ces propriétés est la suivante : si 2 lignes (ou 2 colonnes) sont proportionnelles, alors le déterminant est nul. Plus généralement, si un vecteur colonne s'exprime comme une combinaison linéaire d'autres vecteurs colonnes, alors le déterminant est nul. La réciproque est vraie, ce qui conduit à la propriété très utile[2] :

Proposition 12 *Soit A une matrice d'ordre n*
$\det A \neq 0 \iff$ *les vecteurs colonnes de A sont linéairement indépendants*

1.2.3 Rang

Définition 6 *Le rang d'une matrice, noté* rangA, *est le nombre de vecteurs colonnes de cette matrice linéairement indépendants.*

Proposition 13 *Soit une matrice A d'ordre n,* $\text{rang} A = n \iff \det A \neq 0$

La connaissance de ce nombre joue un rôle fondamental dans la résolution des systèmes d'équations linéaires.

La proposition 20 fournit une méthode de calcul. On en donne dès à présent une autre :

Proposition 14 *Le rang d'une matrice A est la dimension de la plus grande sous-matrice carrée de A dont le déterminant ne s'annule pas.*

Par exemple la matrice rectangulaire 3×2 définie par

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & 0 \end{pmatrix}$$

a 3 sous-matrices 2×2 :

$$\begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix} \text{ et } \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

Le déterminant des 2 dernières est nul mais le déterminant de la première vaut -6, donc $\text{rang}(A) = 2$.

1.3 Inversion

1.3.1 Définition

Définition 7 *Soit* $A \in M_n(\mathbb{K})$. *On dit que A est inversible s'il existe une matrice de* $\mathcal{M}_n(\mathbb{K})$, *notée* A^{-1}, *telle que* $AA^{-1} = A^{-1}A = I_n$

2. Cette proposition est vraie pour les lignes aussi mais son intérêt est bien moindre.

Par exemple :

$$\begin{pmatrix} 1 & 0 & -1 \\ 1 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix} * \begin{pmatrix} 1 & -1 & 1 \\ 1 & -2 & 1 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & -2 & 1 \\ 0 & -1 & 1 \end{pmatrix} * \begin{pmatrix} 1 & 0 & -1 \\ 1 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

L'inverse, s'il existe, est nécessairement unique [3]. Il suffit de trouver une matrice B telle que $AB = I_n$ ou $BA = I_n$ pour être sûr que A est inversible et que son inverse est B.

1.3.2 Matrice inverse et application réciproque

Si f est une application bijective de E dans E, son application réciproque, notée f^{-1}, est l'unique fonction vérifiant $f \circ f^{-1} = f^{-1} \circ f = 1$. Soit A la matrice associée à f dans une base donnée de E. A^{-1}est la matrice associée à f^{-1}, dans la même base.

Par exemple, si la matrice associée à une rotation d'angle θ de centre O est A, A^{-1} est la matrice associée à la rotation de centre O d'angle $-\theta$.

1.3.3 Critère d'inversibilité

La propostion suivante est d'une grande utilité pratique.

Proposition 15 *Soit A une matrice d'ordre n, A est inversible* $\Longleftrightarrow$ $\det A \neq 0$

1.3.4 Propriétés

Proposition 16 *Soient A et B deux matrices inversibles de $\mathcal{M}_n(\mathbb{K})$. Le produit AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.*

En effet, $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n$.

Si A est une matrice inversible, alors ${}^TA$ est inversible et $({}^TA)^{-1} = {}^T(A^{-1})$ [4]. On en déduit que :
- si A est inversible et symétrique alors A^{-1} est symétrique.
- si A est inversible et antisymétrique alors A^{-1} est antisymétrique.

Si A est inversible, la proposition 9 implique que $\det(A^{-1}) = 1/\det A$

Définition 8 *Si A est inversible, on définit $A^{-n} = (A^{-1})^n$*

1.3.5 Calcul de l'inverse

Proposition 17 *Soit $A^{-1} = (b_{ij})$,*

$$b_{ij} = \frac{{}^T\mathcal{A}_{ij}}{\det A}$$

Cette relation fait appel au déterminant de A et aux cofacteurs de la transposé A^T (voir définition 5).

En pratique, le choix de la méthode à utiliser dépend de la taille de la matrice.

3. En effet, soient B_1 et B_2 deux matrices telles que $AB_1 = B_1A = I_n$ et $AB_2 = B_2A = I_n$. En utilisant l'associativité, $B_1AB_2 = B_1(AB_2) = B_1I_n = B_1$, mais aussi $(B_1A)B_2 = I_nB_2 = B_2$. Donc $B_1 = B_2$.

4. $({}^TA)$ est inversible si et seulement si il existe B telle que $B\ ({}^TA) = 1$. Soit ${}^T(B\ ({}^TA)) = A\ {}^TB = 1 \Leftrightarrow {}^TB = A^{-1}$ soit $B = {}^T(A^{-1})$.

Pour des matrices de faible dimension

Utilisation directe de la définition (peu conseillé) :
Soit $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Le 4-uplet (a, b, c, d) doit vérifier $AA^{-1} = I_2$ avec $A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. a, b, c et d doivent donc vérifier le système linéaire de 4 équations à 4 inconnues

$$\left\{ \begin{array}{rrcl} a & +2c & = & 1 \\ b & +2d & = & 0 \\ 3a & +4c & = & 0 \\ 3b & +4d & = & 1 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rcl} a & = & -2 \\ b & = & 1 \\ c & = & \frac{3}{2} \\ d & = & -\frac{1}{2} \end{array} \right. \quad \text{donc} \quad \mathrm{A}^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

Méthode plus maligne :
On considère l'équation $Ax = b$ d'inconnnue $x = (x_1, x_2)$, $b = (b_1, b_2)$ étant un vecteur quelconque de $\mathbb{R}^2$. Résoudre cette équation revient à résoudre un système linéaire de 2 équations à 2 inconnues x_1 et x_2. Si A est inversible, $Ax = b \Leftrightarrow x = A^{-1}b$. Les coefficients de A^{-1} se lisent sur le système résolu.

$$\left\{ \begin{array}{rrcl} x_1 & +2x_2 & = & b_1 \\ 3x_1 & +4x_2 & = & b_2 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rrcll} x_1 & +2x_2 & = & b_1 & \\ & -2x_2 & = & -3b_1 & +b_2 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rcll} x_1 & = & -2b_1 & +b_2 \\ x_2 & = & \frac{3b_1}{2} & -\frac{b_2}{2} \end{array} \right.$$

Méthode assez rapide : celle donnée par la proposition 17.
Application à la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, ${}^TA = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$, $\det A = -2$
$b_{11} = \frac{4}{-2}$, $b_{12} = (-1)\frac{2}{-2}$, $b_{21} = (-1)\frac{3}{-2}$, $b_{22} = \frac{1}{-2}$

Pour des matrices de grande dimension

Analytiquement ou numériquement, la méthode du pivot de Gauss (voir paragraphe 2.1) est incontournable. Il s'agit comme dans l'exemple précédent de résoudre l'équation $AX = B$, d'inconnnue X en écrivant $B = (b_1, ..., b_n)$. Les coordonnées de X s'expriment alors comme une combinaison linéaire des b_i ; les coefficients de ces combinaisons linéaires sont les coefficients de la matrice A^{-1}.

1.3.6 Exemples d'utilisation

- Si A est inversible, x et y étant deux vecteurs, $y = Ax \Leftrightarrow x = A^{-1}y$. Résoudre un système d'équations linéaires revient donc à inverser une matrice (quand celle-ci est inversible) même si, en pratique, on calcule rarement l'inverse directement pour connaître x.
- L'inverse de la matrice de passage permet de calculer les coordonnées d'un vecteur dans une nouvelle base connaissant celles dans l'ancienne base.
- Connaissant la matrice de transfert T d'un quadripôle électrique, calculer son inverse permet de déduire le vecteur de sortie (U_s, I_s) en fonction du vecteur d'entrée (U_e, I_e).

Chapitre 2

Systèmes d'équations

2.1 Résolution de systèmes d'équations linéaires

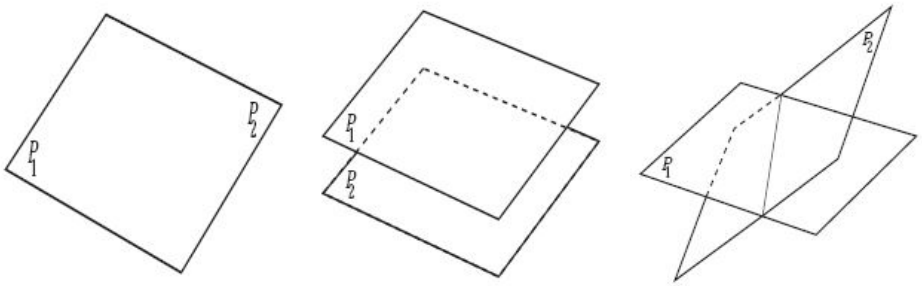
Une des applications importantes du calcul matriciel est la recherche de solutions de systèmes d'équations linéaires algébriques.

2.1.1 Introduction

Voici trois systèmes de deux équations à trois inconnues x, y et z.

$$(S1)\begin{cases} x & +y & +z & = & 1 \\ -x & -y & -z & = & -1 \end{cases} \quad (S2)\begin{cases} x & +y & +z & = & 1 \\ -x & -y & -z & = & 1 \end{cases} \quad (S3)\begin{cases} x & +y & +z & = & 1 \\ x & -y & +z & = & -1 \end{cases}$$

Ces trois cas types sont illustrés par la figure ci-dessous.



Soit le repère $(0, \vec{i}, \vec{j}, \vec{k})$.

Les deux équations du premier système représentent le même plan $\mathcal{P}$. En langage de l'algèbre linéaire, les solutions sont des n-uplets ou des vecteurs. En géométrie, les solutions sont des points. Dans le premier cas, on dira que l'ensemble des solutions est l'ensemble des triplets $(x, y, 1-x-y)$, $(x, y) \in \mathbb{R}^2$ ou que tout vecteur $X = (0, 0, 1) + x(1, 0, -1) + y(0, 1, -1), (x, y) \in \mathbb{R}^2$ est solution de (S_1). Si l'on fait de la géométrie, on dira que l'ensemble des solutions est le plan affine passant par $M_0(0, 0, 1)$ et de vecteurs de base $(1, 0, -1)$ et $(0, 1, -1)$ [1].

Le second système n'admet pas de solution. Il s'agit de deux plans parallèles, un passant par $M_0(0, 0, 1)$ et l'autre par $M_1(0, -1, 0)$: leur intersection est vide.

1. Interprétation géométrique : Soit un plan $\mathcal{P}$ d'équation $ax + by + cz = d$ passant par un point M_0. Rappelons que le vecteur $\vec{n} = a\vec{i} + b\vec{j} + c\vec{k}$ est normal au plan $\mathcal{P}$ car $M(x, y, z)$ appartient à ce plan si et seulement si $\overrightarrow{M_0M} \cdot \vec{n} = 0$. Dans l'exemple, les deux équations représentent le même plan $\mathcal{P}$(de vecteur normal $\vec{n} = \vec{i} + \vec{j} + \vec{k}$ et passant par le point $M_0(0, 0, 1)$).

Pour le troisième système, l'ensemble des solutions est l'ensemble des triplets $(x, 1, -x)$, $x \in \mathbb{R}$. Tout vecteur solution s'écrit $X = (0, 1, 0) + x(1, 0, -1)$. L'ensemble des solutions est une droite affine passant par $M_2(0, 1, 0)$ de vecteur directeur $(1, 0, -1)$ [2].

Si l'on ajoute une troisième équation au système (S_3) et que le nouveau plan croise la droite $\mathcal{D}$, la solution est unique, c'est un point.

$$(S4)\begin{cases} x & +y & +z & = & 1 \\ x & -y & +z & = & -1 \\ x & & & = & 0 \end{cases}$$

Le point $M_0(0, 1, 0)$ dans ce cas.

Si l'on avait rajouté l'équation $x = 1$ à la place de $x = 0$:

$$(S5)\begin{cases} x & +y & +z & = & 1 \\ x & -y & +z & = & -1 \\ x & & & = & 1 \end{cases}$$

le système aurait admis une autre solution unique, le point $M_2(1, 1, -1)$. Remarquez que dans les 2 derniers cas, la matrice d'ordre 3 $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ associée au système possède un déterminant non nul. Les 3 vecteurs colonnes sont linéairement indépendants.

Par contre le système

$$(S6)\begin{cases} x & +y & +z & = & 1 \\ x & -y & z & = & -1 \\ & y & & = & a \end{cases}$$

$(a \in \mathbb{R})$ est associé à la matrice $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ qui a un déterminant nul (les vecteurs colonnes 1 et 3 étant les mêmes). Si a vaut 1, la solution du système est la droite $\mathcal{D}$; si $a \neq 1$, le système n'admet aucune solution.

2.1.2 Définitions

On appelle système linéaire de n équations à p inconnues tout système d'équations de la forme :

$$(S)\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & ... & + & a_{1j}x_j & + & ... & + & a_{1p}x_p & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & ... & + & a_{2j}x_j & + & ... & + & a_{2p}x_p & = & b_2 \\ ... & & & & & & & & & & & & \\ ... & & & & & & & & & & & & \\ a_{i1}x_1 & + & a_{i2}x_2 & + & ... & + & a_{ij}x_j & + & ... & + & a_{ip}x_p & = & b_i \\ ... & & & & & & & & & & & & \\ a_{n1}x_1 & + & a_{n2}x_2 & + & ... & + & a_{nj}x_j & + & ... & + & a_{np}x_p & = & b_n \end{cases}$$

Les coefficients a_{ij}, les seconds membres b_i et les inconnues x_j sont des scalaires, réels ou complexes $(i = 1, ..., n,\ j = 1, ..., p)$. Une solution de (S) est un p-uplet $(x_1, x_2, ..., x_p)$ qui vérifie chacune des n égalités figurant dans (S). Résoudre le système c'est trouver l'ensemble des solutions.

À un système (S) on associe le système homogène (S_H) en annulant les seconds membres.

---

2. Interprétation géométrique : le second plan de (S_3) est celui passant par $M_2(0, 1, 0)$ et de vecteur normal $\vec{n} = \vec{i} - \vec{j} + \vec{k}$. L'intersection des deux plans est la droite $\mathcal{D}$ d'équation $\begin{cases} y & = & 1 \\ z & = & -x \end{cases}$. Rappelons que $M(x, y, z)$ appartient à la droite $\mathcal{D}$ si et seulement si $\overrightarrow{M_2M} = \lambda\vec{u}$ où u est un vecteur directeur de la droite, $\vec{u} = \vec{i} - \vec{k}$ ici.

2.1.3 Interprétation matricielle

Soit $A = (a_{ij})$, élément de $M_{np}(\mathbb{K})$, la matrice du système (S).

Soient $B = \begin{pmatrix} b_1 \\ b_2 \\ ... \\ ... \\ b_i \\ ... \\ b_n \end{pmatrix}$ et $\mathrm{X} = \begin{pmatrix} x_1 \\ ... \\ x_j \\ ... \\ ... \\ x_p \end{pmatrix}$

(S) s'écrit $\begin{pmatrix} a_{11} & ... & a_{1j} & ... & a_{1p} \\ a_{21} & ... & a_{2j} & ... & a_{2p} \\ ... & & & & \\ ... & & & & \\ a_{i1} & ... & a_{ij} & ... & a_{ip} \\ ... & & & & \\ a_{n1} & ... & a_{nj} & ... & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ ... \\ x_j \\ ... \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ ... \\ ... \\ b_i \\ ... \\ b_n \end{pmatrix}$, c'est-à-dire $AX = B$.

Le système homogène (S_H) associé à (S) s'écrit : $AX = 0$.

2.2 Ensemble des solutions de l'équation $AX = B$

On cherche à définir la solution de l'équation $AX = B$ d'inconnue X. Soit (S) le système associé et (S_H) le système homogène.

Solutions du système homogène (S_H)

Proposition 18 *L'ensemble des solutions de (S_H) est un espace vectoriel E_H de dimension $k = p - \mathrm{rang(A)}$.*

- *Si* rang(A) = p, *l'unique solution de (S_H) est $(0, 0, ..., 0)$.*
- *Si* rang(A) < p *: (S_H) possède une infinité de solutions. Soient $X_{H1}, X_{H2}, ..., X_{Hk}$ k solutions particulières de (S_H) formant une base de E_H, toute solution de (S_H) s'écrit*

$$X_H = \lambda_1 X_{H1} + ... + \lambda_k X_{Hk}, \ \ \lambda_1 ... \lambda_k \in \mathbb{K}$$

On appelle X_H la solution générale du système homogène (S_H).

Pour des matrices carrées, le premier cas est caractérisé par det$A \neq 0$.

Solutions du système (S)

Proposition 19 *La solution générale de (S), si elle existe, s'écrit $X = X_0 + X_H$, X_0 étant une solution particulière de (S) et X_H la solution générale de (S_H).*

Démonstration :

- Soit X_0 une solution de (S), on a $AX_0 = B$. Soit X une solution quelconque de (S), on a $AX = B$. Donc $AX - AX_0 = 0$, ce qui est équivalent (grâce à la linéarité à droite du produit de matrices) à $A(X - X_0) = 0$. $X - X_0$ est donc une solution du système homogène (S_H) donc toute solution de (S) s'écrit comme la somme de X_0 et d'une solution de (S_H).

Réciproquement, si $X = X_0 + X_H$, alors $AX = A(X_0 + X_H) = A(X_0) + A(X_H) = B + 0 = B$.

Le théorème fondamental est le suivant :

Théorème 1 *Soit le système linéaire de n équations et p inconnues $AX = B$. On note $A|B$ la matrice obtenue en ajoutant le vecteur B comme colonne supplémentaire à A. Alors,*

- *si* rangA = rang(A|B) = p, *le système admet une seule solution.*
- *si* rangA = rang(A|B) < p, *le système admet une infinité de solutions.*
- *si* rangA < rang(A|B), *le système n'admet aucune solution.*

Pour les systèmes de n équations et n inconnues, on commencera toujours par calculer la valeur du déterminant :

- Si det$A \neq 0$, on est dans le premier cas. La matrice est inversible et la solution, unique, vaut $X = A^{-1}B$. L'unique solution du système homogène est le vecteur nul.

- Si det$A = 0$. Soit le système n'admet pas de solution, soit il en admet une infinité. Les solutions s'écrivent alors $X = X_0 + X_H$, X_0 étant une solution particulière de (S) et X_H étant la solution générale de (S_H) s'écrivant comme une combinaison linéaire de $n -$ rangA vecteurs de base.

Reprenons les exemples précédents :

(S_4) et (S_5) ont un déterminant non nul donc rangA = 3. On est dans le 1er cas où la solution est unique.

Pour (S_6), det$A = 0$, les vecteurs colonnes 1 et 3 sont identiques donc rangA < 3. D'après la proposition 14 rangA = 2.

- Si $a = 1$, le vecteur B est proportionnel au vecteur de la 2e colonne donc rang(A|B) = 2. La solution est un espace de dimension 3-2 =1. C'est une droite.

- Si $a \neq 1$, rang(A|B) = 3 > rangA : le système n'admet aucune solution.

Pour (S_1), rangA = rang(A|B) = 1 < 3, la solution est un espace de dimension 3-1 =2, c'est un plan.

Pour (S_2), rangA = 1 et rang(A|B) = 2, le système n'admet aucune solution.

Pour (S_3), rangA = rang(A|B) = 2 < 3, la solution est un espace de dimension 3-2 =1, c'est une droite.

2.2.1 Systèmes équivalents

Deux systèmes linéaires sont dits équivalents s'ils ont le même ensemble de solutions.

Les transformations suivantes changent tout système en un système équivalent :

1. échanger deux lignes,
2. multiplier une ligne par un scalaire non nul,
3. ajouter une ligne à une autre ligne.

Les opérations 2 et 3 s'effectuent souvent en même temps : $L_i \leftarrow L_i + \lambda L_k$.

Derrière cette notion d'équivalence, il y a la notion de matrices équivalentes. Faire ces opérations sur une matrice revient à changer les bases de départ et d'arrivée dans lesquelles on exprime l'application linéaire associée à la matrice.

2.3 Méthode de Gauss

Principe

La méthode de Gauss consiste à appliquer successivement ces transformations jusqu'à obtenir un système dit échelonné de la forme

$$(S_E)\begin{cases} p_1y_1 + c_{1,2}y_2 + ... + c_{1,j}y_j + ... + c_{1,p}y_p = d_1 \\ \quad p_2y_2 + ... + c_{2,j}y_j + ... + c_{2,p}y_p = d_2 \\ \quad\quad ... \\ \quad\quad p_iy_i + ... + c_{i,p}y_p = d_i \\ \quad\quad\quad 0 = d_{r+1} \\ \quad\quad\quad ... \\ \quad\quad\quad ... \\ \quad\quad\quad 0 = d_n \end{cases}$$

Soit A_E la matrice échelonnée associée à (S_E), on a $A_EY = D$. Les inconnues $y_1, .., y_n$ sont celles de (S) mais dans un ordre qui peut être différent.

Proposition 20 *Les coefficients p_i sont appelés les pivots. Le nombre de pivots non nuls est le rang de la matrice A.*

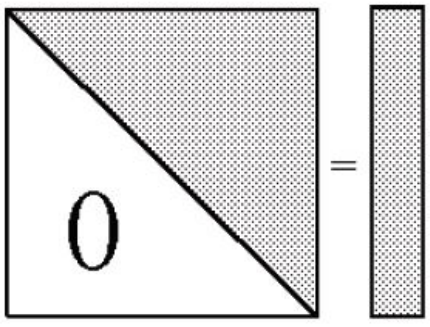
En effet, mis sous forme échelonnée, il est clair que les vecteurs colonnes avec des pivots non nuls sont linéairement indépendants. Le rang de A_E est donc le nombre de ces vecteurs colonnes avec pivots non nuls. Comme le rang est une caractéristique de l'application linéaire associée à la matrice, rang$A$=rangA_E= nombre de pivots non nuls.

La forme échelonnée associée à la méthode de factorisation d'un déterminant indique que le déterminant est égal au produit des pivots non nuls, au facteur $(-1)^r$ près, r étant le nombre d'échanges de lignes effectué.

Les différents cas possibles

Premier cas : Le système (S_E) se présente de manière symbolique comme ci-dessous.

Le système est triangulaire supérieur (système dit de Cramer) ; le nombre de pivots non nuls est égal au nombre d'inconnues. Le déterminant est égal au produit des pivots (multiplié par $(-1)^r$ où r est le nombre de permutations deux à deux de lignes ou de colonnes) donc det$A\neq 0$: la solution de (S) est unique. Elle s'obtient en résolvant (S_E) "en cascades".



L'exemple ci-dessous est une illustration de ce cas.

$$
\begin{array}{rcl}
 & (S_A) & \left\{\begin{array}{rrrrcr} x & +y & -3z & -4t & = & -1 \\ 2x & +2y & +2z & -3t & = & 2 \\ 3x & +6y & -2z & +t & = & 8 \\ 2x & +y & +5z & +t & = & 5 \end{array}\right. \\
 & \Longleftrightarrow & \\
\begin{array}{rcl} L2 & \leftarrow & L2-2L1 \\ L3 & \leftarrow & L3-3L1 \\ L4 & \leftarrow & L4-2L1 \end{array} & & \left\{\begin{array}{rrrrcr} x & +y & -3z & -4t & = & -1 \\ & & 8z & +5t & = & 4 \\ & 3y & +7z & +13t & = & 11 \\ & -y & +11z & +9t & = & 7 \end{array}\right. \\
 & \Longleftrightarrow & \\
\begin{array}{rcl} & & \\ L2 & \leftarrow & L4 \\ & & \\ L4 & \leftarrow & L2 \end{array} & & \left\{\begin{array}{rrrrcr} x & +y & -3z & -4t & = & -1 \\ & -y & +11z & +9t & = & 7 \\ & 3y & +7z & +13t & = & 11 \\ & & 8z & +5t & = & 4 \end{array}\right. \\
 & \Longleftrightarrow & \\
\begin{array}{rcl} & & \\ & & \\ L3 & \leftarrow & L3+3L2 \\ & & \end{array} & & \left\{\begin{array}{rrrrcr} x & +y & -3z & -4t & = & -1 \\ & -y & +11z & +9t & = & 7 \\ & & 40z & +40t & = & 32 \\ & & 8z & +5t & = & 4 \end{array}\right. \\
 & \Longleftrightarrow & \\
 & (S_{A,E}) & \left\{\begin{array}{rrrrcr} x & +y & -3z & -4t & = & -1 \\ & -y & +11z & +9t & = & 7 \\ & & 40z & +40t & = & 32 \\ & & & -3t & = & -12/5 \end{array}\right. \\
L4 \leftarrow L4-L3/5 & &
\end{array}
$$

On peut vérifier que $\det A = 1\times(-1)\times 40\times(-3)\neq 0$.

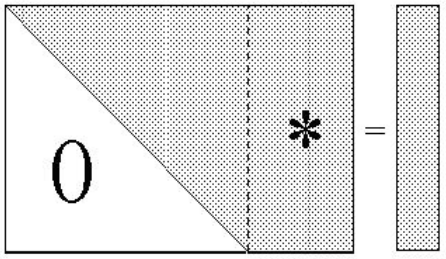
Le système échelonné montre 4 pivots non nuls donc $\text{rang}A=4$, égal au nombre d'inconnues. (S_A) n'admet donc qu'une seule solution d'après le théorème 1. Cette solution s'obtient par la cascade suivante :

$$
\begin{array}{rcl}
 & (S_A) & \Longleftrightarrow \\
\begin{array}{rcl} L2 & \leftarrow & -L2 \\ L3 & \leftarrow & (1/40)L3 \\ L4 & \leftarrow & -(1/3)L4 \end{array} & & \left\{\begin{array}{rrrrcr} x & +y & -3z & -4t & = & -1 \\ & y & -11z & -9t & = & -7 \\ & & z & +t & = & 4/5 \\ & & & t & = & 4/5 \end{array}\right. \\
 & \Longleftrightarrow & \\
\begin{array}{rcl} L1 & \leftarrow & L1+4L4 \\ L2 & \leftarrow & L2+9L4 \\ L3 & \leftarrow & L3-L4 \end{array} & & \left\{\begin{array}{rrrrcr} x & +y & -3z & & = & 11/5 \\ & y & -11z & & = & 1/5 \\ & & z & & = & 0 \\ & & & t & = & 4/5 \end{array}\right. \\
 & \Longleftrightarrow & \\
\begin{array}{rcl} L1 & \leftarrow & L1+3L3 \\ L2 & \leftarrow & L2+11L3 \\ & & \end{array} & & \left\{\begin{array}{rrrrcr} x & +y & & & = & 11/5 \\ & y & & & = & 1/5 \\ & & z & & = & 0 \\ & & & t & = & 4/5 \end{array}\right. \\
 & \Longleftrightarrow & \\
L1 \leftarrow L1-L2 & & \left\{\begin{array}{rrrrcr} x & & & & = & 2 \\ & y & & & = & 1/5 \\ & & z & & = & 0 \\ & & & t & = & 4/5 \end{array}\right.
\end{array}
$$

Le 4-uplet $(2, 1/5, 0, 4/5)$ est l'unique solution.

On aurait aussi pu calculer la matrice inverse après s'être assuré que le déterminant était non nul.

Second cas : Le système (S_E) se présente de manière symbolique comme ceci :



On est dans le cas où $rang(A) = rang(A|b) < p$. Il y a ici des inconnues en surnombre (zone marquée du symbole *). Ces inconnues excédentaires (ou non principales) seront reportées au second membre, où elles jouent le rôle de paramètres arbitraires. On résoud alors le système par rapport aux inconnues principales. La présence de ces valeurs arbitraires montre que (S) possède une infinité de solutions. C'est le cas de l'exemple ci-dessous.

$$
\begin{array}{rcl}
 & (S_B) & \left\{\begin{array}{rrrrrcr} x & +3y & +5z & -2t & -7u & = & 3 \\ 3x & +y & +z & -2t & -u & = & 1 \\ 2x & -y & -3z & +7t & +5u & = & 2 \\ 3x & -2y & -5z & +7t & +8u & = & 1 \end{array}\right. \\
 & \Longleftrightarrow & \\
L2 \leftarrow L2-3L1 & & \left\{\begin{array}{rrrrrcr} x & +3y & +5z & -2t & -7u & = & 3 \\ & -8y & -14z & +4t & +20u & = & -8 \\ & -7y & -13z & +11t & +19u & = & -4 \\ & -11y & -20z & +13t & +29u & = & -8 \end{array}\right. \\
L3 \leftarrow L3-2L1 & & \\
L4 \leftarrow L4-3L1 & & \\
 & \Longleftrightarrow & \\
 & & \left\{\begin{array}{rrrrrcr} x & +3y & +5z & -2t & -7u & = & 3 \\ & -8y & -14z & +4t & +20u & = & -8 \\ & & -6z & +60t & +12u & = & 24 \\ & & -6z & +60t & +12u & = & 24 \end{array}\right. \\
L3 \leftarrow 8L3-7L2 & & \\
L4 \leftarrow 8L4-11L2 & & \\
 & \Longleftrightarrow & \\
 & & \left\{\begin{array}{rrrrrcr} x & +3y & +5z & -2t & -7u & = & 3 \\ & -8y & -14z & +4t & +20u & = & -8 \\ & & -6z & +60t & +12u & = & 24 \\ & & & & 0 & = & 0 \end{array}\right. \\
 & (S_{B,E}) & \\
L4 \leftarrow L4-L3 & &
\end{array}
$$

Le système (S_B) se réduit donc à 3 équations. Il possède 3 pivots non nuls : son rang est 3. rang(A|b) vaut aussi 3 puisque le vecteur $(3, -8, 24)$, vecteur de $\mathbb{R}^3$, est forcément combinaison linéaire des 3 vecteurs colonnes $(1, 0, 0)$, $(3, -8, 0)$, $(5, -14, 6)$. En traitant t et u comme des paramètres arbitraires, on poursuit la résolution avec les inconnues x, y et z.

$$
\begin{array}{rcl}
 & & (S_B) \\
L2 & \leftarrow & -L2/2 \\
L3 & \leftarrow & -L3/6
\end{array}
\left\{\begin{array}{rrrcrrr}
x & +3y & +5z & = & 3 & +2t & +7u \\
 & 4y & +7z & = & 4 & +2t & +10u \\
 & & z & = & -4 & +10t & +2u
\end{array}\right.
$$

$$\Longleftrightarrow$$

$$
\begin{array}{rcl}
L1 & \leftarrow & L1-5L3 \\
L2 & \leftarrow & L2-7L3
\end{array}
\left\{\begin{array}{rrrcrrr}
x & +3y & & = & 23 & -48t & -3u \\
 & 4y & & = & 32 & -68t & -4u \\
 & & z & = & -4 & +10t & +2u
\end{array}\right.
$$

$$\Longleftrightarrow$$

$$
\begin{array}{rcl}
 & & \\
L2 & \leftarrow & L2/4
\end{array}
\left\{\begin{array}{rrrcrrr}
x & +3y & & = & 23 & -48t & -3u \\
 & y & & = & 8 & -17t & -u \\
 & & z & = & -4 & +10t & +2u
\end{array}\right.
$$

$$\Longleftrightarrow$$

$$
\begin{array}{rcl}
L1 & \leftarrow & L1-3L2
\end{array}
\left\{\begin{array}{rrrcrrr}
x & & & = & -1 & +3t & \\
 & y & & = & 8 & -17t & -u \\
 & & z & = & -4 & +10t & +2u
\end{array}\right.
$$

Les solutions de (S_B) sont les 5-uplets $X = (x, y, z, t, u)$ qui s'écrivent :

$$X = (-1+3t,\ 8-17t-u,\ -4+10t+2u,\ t,\ u)\,(t,u) \in \mathbb{R}^2$$

Dit autrement, la solution générale de (S_B) est l'ensemble des vecteurs

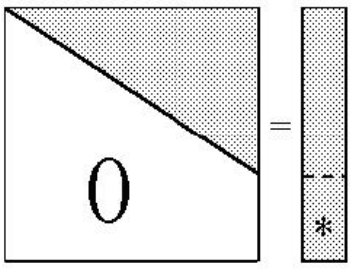
$$X = (-1,8,-4,0,0) + t(3,-17,10,1,0) + u(0,-1,2,0,1),\ \ (t,u) \in \mathbb{R}^2$$

.

Il s'agit du plan défini par les 2 vecteurs de base $\overrightarrow{u} = (3,-17,10,1,0)$ et $\overrightarrow{v} = (0,-1,2,0,1)$ et passant par le point $A_0 = (-1,8,-4,0,0)$.

On voit bien ici la structure de l'ensemble des solutions : à la solution particulière $A_0 = (-1,8,-4,0,0)$ (qui est obtenue pour $t = u = 0$), on ajoute la solution générale du système homogène (S_H) associé à (S_B).

Troisième cas : Le système (S_E) seprésente de manière symbolique comme suit :



Ici il y a des équations en surnombre, dites non principales. Leurs premiers membres sont nuls. Tout dépend alors de leurs seconds membres (zone marquée d'une *) :

1er cas : si l'un d'eux est non nul, alors (S_E) et donc (S) n'ont pas de solution. On est dans le cas où $rang(A) < rang(A|b)$. Voici un exemple :

$$
\begin{array}{rcl}
 & & (S_C) \\
 & & \Longleftrightarrow \\
L2 & \leftarrow & L2 - L1 \\
L3 & \leftarrow & L3 + L1 \\
L4 & \leftarrow & L4 - 2L1 \\
 & & \Longleftrightarrow \\
L3 & \leftarrow & L3 - 3L2 \\
L4 & \leftarrow & L4 + 3L2 \\
 & & \Longleftrightarrow \\
 & z \leftrightarrow t & \\
 & & \Longleftrightarrow \\
 & & (S_{C,E}) \\
L4 & \leftarrow & L4 + \frac{7}{9}L3
\end{array}
\left\{
\begin{array}{rrrrcr}
x & +2y & -z & +t & = & 1 \\
x & +3y & +z & -t & = & 2 \\
-x & +y & +7z & +2t & = & 3 \\
2x & +y & -8z & +t & = & 4
\end{array}
\right.
\left\{
\begin{array}{rrrrcr}
x & +2y & -z & +t & = & 1 \\
 & y & +2z & -2t & = & 1 \\
 & 3y & +6z & +3t & = & 4 \\
 & -3y & -6z & -t & = & 2
\end{array}
\right.
\left\{
\begin{array}{rrrrcr}
x & +2y & -z & +t & = & 1 \\
 & y & +2z & -2t & = & 1 \\
 & & & +9t & = & 1 \\
 & & & -7t & = & 5
\end{array}
\right.
\left\{
\begin{array}{rrrrcr}
x & +2y & +t & -z & = & 1 \\
 & y & -2t & +2z & = & 1 \\
 & & 9t & & = & 1 \\
 & & -7t & & = & 5
\end{array}
\right.
\left\{
\begin{array}{rrrrcr}
x & +2y & +t & -z & = & 1 \\
 & y & -2t & +2z & = & 1 \\
 & & 9t & & = & 1 \\
 & & & 0 & = & 52/9
\end{array}
\right.
$$

2nd cas : si tous les seconds membres des équations non principales sont nuls, (S_E) se réduit à ses équations principales, qui forment un système de Cramer. On est dans le cas où $rang(A) = rang(A|b) < p$: le système admet une infinité de solutions. Il se résoud "en cascades" comme le montre l'exemple suivant :

$$
\begin{array}{rcl}
 & & (S_D) \\
 & & \Longleftrightarrow \\
L2 & \leftarrow & L2 - 5L1 \\
L3 & \leftarrow & L3 + 3L1 \\
L4 & \leftarrow & L4 - 6L1 \\
 & & \Longleftrightarrow \\
L3 & \leftarrow & L3 + L2 \\
L4 & \leftarrow & L4 - L2 \\
 & & \Longleftrightarrow \\
 & & (S_{D,E})
\end{array}
\left\{
\begin{array}{rrrrcr}
x & -y & +z & +t & = & 3 \\
5x & +2y & -z & -3t & = & 5 \\
-3x & -4y & +3z & +2t & = & 1 \\
6x & +y & & -2t & = & 8
\end{array}
\right.
\left\{
\begin{array}{rrrrcr}
x & -y & +z & +t & = & 3 \\
 & 7y & -6z & -8t & = & -10 \\
 & -7y & +6z & +5t & = & 10 \\
 & 7y & -6z & -8t & = & -10
\end{array}
\right.
\left\{
\begin{array}{rrrrcr}
x & -y & +z & +t & = & 3 \\
 & 7y & -6z & -8t & = & -10 \\
 & & & -3t & = & 0 \\
 & & & 0 & = & 0
\end{array}
\right.
\left\{
\begin{array}{rrrrcr}
x & -y & +t & +z & = & 3 \\
 & 7y & -8t & -6z & = & -10 \\
 & & -3t & & = & 0 \\
 & & & 0 & = & 0
\end{array}
\right.
$$

La forme échelonnée montre que le nombre d'équations principales est 3. z devient un paramètre secondaire que l'on passe dans le membre de droite. Résolvons le système réduit à ses équations

principales :

$$
\begin{array}{rcl}
 & (S_D) & \left\{\begin{array}{rrrcrr} x & -y & +t & = & 3 & -z \\ & 7y & -8t & = & -10 & -6z \\ & & -3t & = & 0 & \end{array}\right. \\
 & \Longleftrightarrow & \\
\begin{array}{rcr} & & \\ L2 & \leftarrow & (1/7)L2 \\ L3 & \leftarrow & -(1/3)L3 \end{array} & & \left\{\begin{array}{rrrcrr} x & -y & +t & = & 3 & -z \\ & y & -(8/7)t & = & -10/7 & +(6/7)z \\ & & t & = & 0 & \end{array}\right. \\
 & \Longleftrightarrow & \\
\begin{array}{rcr} L1 & \leftarrow & L1 - L3 \\ L2 & \leftarrow & L2 + (8/7)L3 \end{array} & & \left\{\begin{array}{rrcrr} x & -y & = & 3 & -z \\ & y & = & -10/7 & +(6/7)z \\ & t & = & 0 & \end{array}\right. \\
 & \Longleftrightarrow & \\
\begin{array}{rcr} L1 & \leftarrow & L1 + L2 \end{array} & & \left\{\begin{array}{rrcrr} x & & = & 11/7 & -(1/7)z \\ & y & = & -10/7 & +(6/7)z \\ & t & = & 0 & \end{array}\right.
\end{array}
$$

Les solutions de (S_D) sont les 4-uplets $X = (x, y, z, t)$ qui s'écrivent :

$$X = (11/7 - (1/7)z,\ -10/7 + (6/7)z,\ z,\ 0),\ z \in \mathbb{R}$$

(S_D) admet une infinité de solutions, dépendant du paramètre z. L'ensemble des solutions s'écrit

$$S = \{(11/7, -10/7, 0, 0) + z(-1/7, 6/7, 1, 0), z \in \mathbb{R}\}$$

Comme prévu par le théorème 1, comme $rang(A) = rang(A|b) = 3 < 4$, le nombre de solutions est infini de dimension 4-3=1. L'ensemble des solutions de (S_D) est la droite affine, passant par la solution particulière $(11/7, -10/7, 0, 0)$, de vecteur directeur $(-1/7, 6/7, 1, 0)$.

Chapitre 3

Diagonalisation

3.1 Bases et changements de base

Dans cette section, nous rappelons de manière très (trop ?) rapide quelques notions essentielles. Sans définir un espace vectoriel, citons ceux que vous manipulez régulièrement :

- $\mathbb{R}^n$: l'espace des n-uplets de réels $(x_1, ..., x_n)$
- $\mathbb{C}$: l'ensemble des nombres complexes z
- L'ensemble des polynômes dans $\mathbb{R}$: $\{a_0 + a_1x + a_2x^2 + ... + a_nx^n, (a_0, ..., a_n) \in \mathbb{R}^{n+1}\}$
- L'ensemble des applications continues de $\mathbb{R}$ dans $\mathbb{R}$.

Un élément d'un espace vectoriel s'appelle un vecteur.

La structure d'espace vectoriel sur $\mathbb{R}$ assure que la somme de deux éléments de l'ensemble appartient à l'ensemble, et la multiplication par un réel aussi. Cette propriété des espaces vectoriels est essentielle.

Soient $v_1, ..., v_n$ n vecteurs d'un espace vectoriel E. On appelle combinaison linéaire de $v_1, ..., v_n$, tout vecteur s'écrivant :

$$v = \lambda_1 v_1 + ... + \lambda_n v_n = \sum_{i=1}^{n} \lambda_i v_i \text{ avec } \lambda_i \in \mathbb{K}$$

Une famille $v_1, ..., v_n$ de vecteurs d'un espace vectoriel E est dite génératrice si tout vecteur de E s'écrit comme une combinaison linéaire de ces n vecteurs.

Des vecteurs sont linéairement indépendants si et seulement si on ne peut pas écrire l'un comme une combinaison linéaire des autres. On parle de famille de vecteurs libres (sinon la famille est liée). Cela se traduit par : $v_1, ..., v_n$ sont n vecteurs linéairement indépendants si et seulement si

$$\lambda_1, ..., \lambda_n \in \mathbb{K}, \ \sum_{i=1}^{n} \lambda_i v_i = 0 \Longrightarrow \lambda_i = 0$$

On appelle base de E toute famille de vecteurs qui est libre et génératrice.

Toutes les bases de E (si E est non nul) contiennent le même nombre de vecteurs (si ce nombre est fini). Ce nombre définit la dimension de E.

Tout vecteur v de E s'écrit comme une combinaison linéaire de ces n vecteurs de base : $v = \sum_{i=1}^{n} \lambda_i v_i$, $\lambda_i \in \mathbb{K}$. Les coordonnées de v sont les coefficients λ_i de cette combinaison linéaire. Ces coefficients sont uniques pour une base donnée.

Exemples :

- $(1, i)$ est une base de $\mathbb{C}$. Tout nombre complexe s'écrit $z = x + iy$, $(x, y) \in \mathbb{R}^2$.
- Les couples $(0, 1)$ et $(1, 0)$ forment une base de $\mathbb{R}^2$. Tout élément de $\mathbb{R}^2$ s'écrit $(x, y) = x(1, 0)+y(0, 1)$, $(x, y) \in \mathbb{R}^2$. Si l'on identifie le plan $\mathcal{P}$ muni du repère $(O, \overrightarrow{i}, \overrightarrow{j})$ à l'ensemble $\mathbb{R}^2$, à tout point M de ce plan est associé un couple unique de coordonnées (x, y). Remarque : dans ce plan, tout couple de vecteurs non colinéaires forme une base dans laquelle le point M a d'autres coordonnées.
- Les vecteurs $(1, 0, 0)$, $(0, 1, 0)$ et $(0, 0, 1)$ forment une base de $\mathbb{R}^3$ (appelée base canonique car tous ses vecteurs ont des composantes nulles, sauf une).
- Les fonctions $x \to 1$, $x \to x$, $x \to x^2$, $x \to x^3$ forment une base de l'ensemble des polynômes de degrés inférieurs ou égaux à 3.
- Le couple de fonctions cos et sin est une base de l'espace des fonctions numériques qui s'écrivent $f(x) = a\cos x + b\sin x$, $(a, b) \in \mathbb{R}^2$. Ceci explique pourquoi on peut affirmer que, si pour tout x, $a\cos x + b\sin x = 0$ alors $a = b = 0$.
- Les fonctions 1, $\exp(x)$ et $\exp(2x)$ forment une base de l'espace des fonctions qui s'écrivent $f(x) = a + b\exp(x) + c\exp(2x)$, $(a, b, c) \in \mathbb{R}^3$.
- L'ensemble des applications de $\mathbb{R}$ dans $\mathbb{R}$ continues sur $I \subset \mathbb{R}$ est un espace vectoriel mais de dimension infinie !

3.1.1 Matrices de passage et changement de base

Par définition, **la matrice de passage** P de la base $\{\vec{e}_i\}$ vers la base $\{\vec{e}\,'_i\}$ est telle que sa ième colonne est formée des composantes de $\vec{e}\,'_i$ exprimées dans la base $\{\vec{e}_i\}$, soit :

$$\boxed{\vec{e}\,'_j \quad = \quad \sum_{i=1}^{N} P_{ij}\,\vec{e}_i,}$$

et inversement $\vec{e}_j = \sum_{i=1}^{N} P^{-1}_{ij}\,\vec{e}\,'_i$.

Un vecteur quelconque $\vec{V} \in E$ peut s'écrire indifféremment dans l'une ou l'autre base de E :

$$\vec{V} = \sum_{j=1}^{N} x_j\,\vec{e}_j = \sum_{j=1}^{N} x'_j\,\vec{e}\,'_j$$

En utilisant les formules précédentes, on est conduit aux relations suivantes entre les composantes de $\vec{V}$ dans les deux bases :

$$x'_i = \sum_{j=1}^{N} P^{-1}_{ij}\,x_j, \quad \text{et} \quad x_i = \sum_{j=1}^{N} P_{ij}\,x'_j \tag{3.1}$$

Considérons maintenant une application linéaire $f : E \to E$ [1] représenté dans la base $\{\vec{e}_i\}$ par la matrice $A : \vec{x} \in E \mapsto A\vec{x} \in E$. La jème colonne de A est le vecteur $A\,\vec{e}_j$:

$$A\,\vec{e}_j = \sum_{k=1}^{N} A_{kj}\,\vec{e}_k$$

1. Les physiciens parlent plutôt d'opérateur, les mathématiciens plutôt d'endomorphisme.

Rappelons quelle est l'expression de A dans la base $\{\vec{e}'_i\}$. Soient A'_{kj} les coefficients de A dans cette base. Par définition, on a $A\,\vec{e}'_j = \sum_k A'_{kj}\,\vec{e}'_k$, mais on a aussi :

$$A\,\vec{e}'_j \;=\; A\left(\sum_{l=1}^{N} P_{lj}\,\vec{e}_l\right) = \sum_{l,m=1}^{N} P_{lj}\,A_{ml}\,\vec{e}_m = \sum_{k=1}^{N}\left(\sum_{l,m=1}^{N} P_{lj}\,A_{ml}\,P^{-1}_{km}\right)\vec{e}'_k,$$

soit en égalisant les deux expressions de $A\,\vec{e}'_j$:

$$A'_{kj} = \sum_{m,l=1}^{N} P^{-1}_{km}\,A_{ml}\,P_{lj}, \tag{3.2}$$

Réécrivons (3.1) et (3.2) sous forme matricielle :

Proposition 21 Soit E un espace vectoriel de dimension finie N. Soient P la matrice de passage de la base $\{\vec{e}_i\}$ vers la base $\{\vec{e}'_i\}$, $\vec{V}$ un vecteur de E et f une application linéaire de E dans E.
Les matrices (ou vecteurs colonnes) $X = \begin{pmatrix} x_1 \\ ... \\ x_N \end{pmatrix}$ et $X' = \begin{pmatrix} x'_1 \\ ... \\ x'_N \end{pmatrix}$ sont les représentations de $\vec{V}$ respectivement dans la base $\{\vec{e}_i\}$ et $\{\vec{e}'_i\}$.
Les matrices carré A et A', d'ordre N, sont les représentations de f respectivement dans la base $\{\vec{e}_i\}$ et $\{\vec{e}'_i\}$.
On a les relations :

$$X' = P^{-1}\,X \quad \textbf{et} \quad A' = P^{-1}\,A\,P$$

Un vecteur met donc en jeu une seule matrice de passage (P^{-1}) lors d'un changement de base, tandis qu'il en faut 2 (P et P^{-1}) pour transformer un opérateur linéaire. Un nombre réel, invariant par changement de base, n'en met aucune en jeu [2].

Deux matrices A et B sont dites similaires, ou **semblables**, s'il existe une matrice inversible P telle que $B = P^{-1}AP$. Les matrices A et A' définies ci-dessus sont donc un cas particulier de matrices semblables. Ces 2 matrices sont deux représentations d'un même opérateur décrit dans 2 bases différentes.

3.2 Diagonalisation

Définition 9 Une matrice carrée A est diagonalisable s'il existe une matrice inversible P telle que la matrice $D = P^{-1}AP$ soit diagonale.

Lorsqu'une matrice est diagonalisable, les calculs algébriques sont simplifiés. Le calcul de A^n en est un exemple. Comme $PP^{-1} = I$, on a $A = PDP^{-1}$, $A^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1}$ et finalement $A^n = PD^nP^{-1}$, D^n étant la matrice diagonale de coefficients égaux à ceux de D élevés à la puissance n. Ainsi, tout polynôme en A se calcule aisément lorsque A est diagonalisable.

2. Le calcul tensoriel systématise ce point de vue en considérant des êtres mathématiques nouveaux : *les tenseurs*, qui se transforment lors d'un changement de base en mettant en jeu un nombre quelconque de matrices de passage P et P^{-1}. L'algèbre (et l'analyse) tensorielle constitue une généralisation de l'algèbre (et de l'analyse) linéaire élémentaire que les mathématiciens appellent *algèbre multilinéaire*.

Les vecteurs qui se trouvent seulement contractés ou dilatés dans leur direction sous l'action de l'opérateur linéaire f associé à la matrice A jouent un rôle spécifique pour la diagonalisation. Ces vecteurs sont appelés vecteurs propres.

Définition 10 Le scalaire (réel ou complexe) λ est dit valeur propre de l'application linéaire f si et seulement si il existe un vecteur non nul $\vec{v}$ tel que $f(\vec{v}) = \lambda\,\vec{v}$. Un tel vecteur $\vec{v}$ est appelé vecteur propre de f associé à la valeur propre λ.

3.2.1 Recherche des valeurs propres

λ est le coefficient de dilatation ou de contraction de $\vec{v}$ dans sa propre direction.

Soit une application linéaire f exprimée par la matrice A dans une certaine base. Déterminer le problème aux valeurs propres (i.e. trouver les vecteurs et valeurs propres) de f revient à résoudre le système linéaire homogène :

$$(A - \lambda\, I)\,\vec{v} = 0$$

Les solutions non triviales (i.e. telle que $\vec{v} \neq \vec{0}$) sont telles que $\det(A - \lambda\, I) = 0$. Pour A d'ordre n connue, $\det(A - \lambda\, I)$ est un polynôme en λ d'ordre n, appelé **polynôme caractéristique** de f et noté $P(\lambda)$. Les valeurs propres λ sont donc les racines de ce polynôme :

$$P(\lambda) = \det(A - \lambda\, I) = 0$$

Cas particulier des matrices 2 × 2

Dans le cas des matrices d'ordre 2, il est facile de montrer que le polynôme caractéristique s'exprime simplement en fonction de la trace et du déterminant de A. En effet, soit A une matrice 2×2 telle que $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Comme $P(\lambda) = \lambda^2 - \lambda(a+d) + ad - bc = \lambda^2 - \mathrm{tr}\,A\,\lambda + \mathrm{detA}$, le polynôme caractéristique ne dépend que de la trace et du déterminant de A. On obtient aussitôt :

$$\lambda_{\pm} = \frac{\mathrm{tr}A \pm \sqrt{(\mathrm{trA})^2 - 4\,\mathrm{detA}}}{2}$$

Noter que $\det A = \Pi_i \lambda_i$ et $\mathrm{trA} = \sum_i \lambda_i$. Cherchons des valeurs propres réelles :

Si $(\mathrm{trA})^2 - 4\,\mathrm{detA}$ est :

- positif, il existe 2 valeurs propres simples ; $P(\lambda) = (\lambda - \lambda_+)(\lambda - \lambda_-) = 0$.
- négatif, l'application f n'admet pas de valeur propre réelle.
- nul, il existe une valeur propre double $\lambda_0 = \frac{\mathrm{tr}A}{2}$; $P(\lambda) = (\lambda - \lambda_0)^2 = 0$.

Cas général de la dimension n

Dans le cas d'un problème de dimension n, le nombre de valeurs propres possibles dépend du corps, typiquement $\mathbb{R}$ ou $\mathbb{C}$ dans lequel on cherche les racines de l'équation $P(\lambda) = 0$. Dans $\mathbb{R}$, tout est possible, depuis aucune jusqu'à n solutions. Dans $\mathbb{C}$, le théorème de d'Alembert nous garantit l'existence de n valeurs propres, éventuellement multiples. Dans ce dernier cas (racine multiple), on dit que **la valeur propre est dégénérée**.

3.2.2 L'ensemble des vecteurs propres

Une fois trouvées les valeurs propres, la résolution des systèmes linéaires associés à chaque valeur propre permet d'obtenir les vecteurs propres correspondants. Compte tenu de la forme linéaire $A\vec{v} = \lambda\,\vec{v}$, il est clair que les vecteurs propres sont déterminés à un facteur multiplicatif près (si $\vec{v}$ est vecteur propre, $\alpha\vec{v}$ avec $\alpha \in \mathbb{C}$ l'est aussi).

Considérons la matrice P dont les colonnes sont formées par les vecteurs propres de la matrice A. Puisque $A\vec{v}_i = \lambda_i\,\vec{v}_i$, le produit matriciel AP est une matrice dont les colonnes sont les vecteurs

$\lambda_i \vec{v}_i$. D'autre-part, AP est aussi égal à PD où D est la matrice diagonale de coefficients égaux aux valeurs propres λ_i. La matrice P satisfait donc la relation $AP = PD$.

Deux cas peuvent se présenter pour l'ensemble des vecteurs propres $(\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n)$ associé à l'application linéaire f :

- Ils sont linéairement indépendants. Ils constituent alors une base de l'espace vectoriel E. On a $\det P \neq 0$, la matrice P est inversible et l'on peut donc écrire $D = P^{-1}AP$ où, rappelons-le, P a pour colonnes les vecteurs propres de A.
- Ils ne sont pas indépendants. On a $\det P = 0$, la matrice P n'est pas inversible. On peut montrer qu'il n'existe pas d'autre matrice P qui satisfasse la relation $D = P^{-1}AP$: la matrice A n'est donc pas diagonalisable.

Il en résulte le théorème fondamental suivant :

Théorème 1 Soit une application linéaire représentée par la matrice carré A, d'ordre n, dans la base E. Cette matrice est diagonalisable si et seulement si elle possède n vecteurs propres linéairement indépendants.
La ième colonne de la matrice P telle que $P^{-1}AP = D$ est formée par les composantes du ième vecteur propre exprimé dans la base E, le ième coefficient de la matrice D étant la valeur propre correspondante.

3.2.3 Condition suffisante pour qu'une matrice soit diagonalisable

Suivant les exigences de ce théorème, on doit s'attendre à ce que beaucoup de matrices ne soient pas diagonalisables. C'est bien le cas. Il existe cependant certaines matrices, fréquemment rencontrées dans les applications physiques, pour lesquelles la diagonalisation est garantie. Il s'agit par exemple des matrices $n \times n$ ayant n valeurs propres distinctes.

Théorème 2 Une matrice ayant des valeurs propres toutes distinctes est diagonalisable.

La démonstration se fait par l'absurde. Supposons que les vecteurs propres $(\vec{v}_1, \cdots, \vec{v}_p)$ soient des vecteurs liés. Un vecteur propre particulier, disons $\vec{v}_1$ s'exprime en fonction des $(p-1)$ autres supposés indépendants : $\vec{v}_1 = \sum_{i=2,p} c_i \vec{v}_i$ avec c_i non tous nuls. On a par définition $(A - \lambda_1)\vec{v}_1 = 0 = \sum_{i=2,p} c_i(\lambda_i - \lambda_1)\vec{v}_i$. Comme $(\vec{v}_2, \cdots, \vec{v}_p)$ sont indépendants, il en résulte que $c_i(\lambda_i - \lambda_1) = 0, \forall i = 2...p$. Mais $\lambda_i \neq \lambda_1, \forall i = 2...p$ puisque par hypothèse les valeurs propres sont toutes distinctes, donc $c_i = 0, \forall i = 2...p$ en contradiction avec le fait que les c_i étaient supposés non tous nuls.

3.3 Application : deux particules couplées élastiquement

Considérons 2 particules identiques de masse m couplées élastiquement et contraintes de se déplacer selon une seule direction dont les équations du mouvement s'écrivent :

$$\begin{aligned} m\,\ddot{x}_1(t) &= k(x_2 - 2x_1), \\ m\,\ddot{x}_2(t) &= k(x_1 - 2x_2) \end{aligned}$$

où x_1 et x_2 désignent les déplacements des particules par rapport à leurs positions d'équilibre.

Posons $\omega^2 \equiv k/m$. Ces équations s'écrivent encore sous la forme du système différentiel suivant :

$$\begin{aligned} \ddot{x}_1 + 2\omega^2 x_1 &= \omega^2 x_2, \\ \ddot{x}_2 + 2\omega^2 x_2 &= \omega^2 x_1 \end{aligned}$$

3.3.1 Résolution par découplage des variables

Ce système de 2 équations différentielles inhomogènes couplées peut être facilement découplé en introduisant les 2 nouvelles variables $x \equiv x_1 + x_2$ et $y \equiv x_1 - x_2$ qui vérifient les équations différentielles homogènes :

$$\begin{aligned} \ddot{x} + \omega^2 x &= 0, \\ \ddot{y} + 3\omega^2 y &= 0 \end{aligned}$$

Les solutions générales pouvant s'écrire sous la forme $x(t) = A\,\cos(\omega t + \varphi)$ et $y(t) = B\,\cos(\sqrt{3}\omega t + \psi)$, on en déduit la solution du système différentiel initial :

$$x_1(t) = A_1\,\cos(\omega t + \varphi) + B_1\,\cos(\sqrt{3}\omega t + \psi) \quad \text{et} \quad x_2(t) = A_1\,\cos(\omega t + \varphi) - B_1\,\cos(\sqrt{3}\omega t + \psi)$$

Il est clair que si les conditions initiales sont telles que $B_1 = 0$, les 2 masses vibrent en phase (à la fréquence ω), tandis qu'elles vibrent en opposition de phase (à la fréquence $\sqrt{3}\omega$) si les conditions initiales sont telles que $A_1 = 0$. En général, le mouvement des masses est la superposition de ces 2 mouvements de base que l'on appelle les modes normaux de vibrations.

3.3.2 Résolution matricielle

Soit $\vec{x}$ le vecteur de composantes x_1 et x_2 solution du système différentiel initial. Ce système peut se mettre sous la forme $\ddot{\vec{x}} = A\vec{x}$ où A est la matrice

$$A = \begin{pmatrix} -2\omega^2 & \omega^2 \\ \omega^2 & -2\omega^2 \end{pmatrix}$$

La solution générale de cette équation est une combinaison linéaire de deux solutions particulières $\vec{x}_a$ et $\vec{x}_b$: $\vec{x} = a\vec{x}_a + b\vec{x}_b$, où a et b sont des constantes complexes dépendant des conditions initiales. D'une façon générale, pour déterminer les modes normaux de vibrations, on écrit les solutions particulières ($p = a, b$) sous la forme $\vec{x}_p(t) = \vec{u}_p\,\mathrm{e}^{i\alpha_p t}$. Cette écriture permet de déterminer $\vec{x}_a$ et $\vec{x}_b$ en résolvant un problème aux valeurs propres. $\vec{u}_p$ sera un vecteur propre et α_p dépendra de la valeur propre correspondante.

En écrivant $\ddot{\vec{x}}_p = A\vec{x}_p$, on obtient $A\,\vec{u}_p = -\alpha_p^2\,\vec{u}_p$. Les 2 valeurs propres de A sont $-\omega^2$ et $-3\omega^2$, auxquelles correspondent les 2 fréquences caractéristiques $\alpha_a = \omega \quad \text{et} \quad \alpha_b = \sqrt{3}\,\omega$. Les 2 vecteurs propres u_a et u_b correspondant sont respectivement colinéaires à $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. La matrice étant symétrique, les 2 vecteurs propres sont, comme attendu, orthogonaux.

La solution générale s'écrit alors :

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = a \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mathrm{e}^{i\omega t} + b \begin{pmatrix} 1 \\ -1 \end{pmatrix} \mathrm{e}^{i\sqrt{3}\omega t}$$

Deuxième partie

Analyse vectorielle

Chapitre 4

Systèmes de coordonnées et champs

La plupart des grandeurs physiques rencontrées en Physique sont des fonctions de plusieurs variables. Ainsi en va-t-il des *champs scalaires* : température, pression ... et des *champs vectoriels* : champs électriques, champs de vitesses, ..., qui, en représentation eulérienne, sont des fonctions des variables d'espace $\vec{r} \in \mathbb{R}^3$ et du temps t.

On considèrera dans la suite, pour fixer les idées, et simplifier l'écriture, des fonctions de 3 variables indépendantes, par exemple x, y et z. La plupart des résultats donnés dans ce chapitre se généralise aisément aux cas d'un nombre quelconque (mais fini) de variables.

4.1 Coordonnées cartésiennes, cylindriques, sphériques

Un point dans $\mathbb{R}^3$ sera noté par un vecteur $\boldsymbol{r}$ avec les composantes (x, y, z) qui sont les coordonnées cartésiennes de ce point. On peut décrire ce point de plusieurs manières différentes. Les coordonnées sont alors modifiées mais aussi les vecteurs de la base ou encore les mesures d'intégrations. Cette section fait un rapide résumé des trois systèmes de coordonnées les plus utilisés à savoir les coordonnées cartésiennes, cylindriques et sphériques

4.1.1 Coordonnées cylindriques

Dans les *coordonnées cylindriques*, ce même point sera présenté par les paramètres (ρ, ϕ, z) où :

$$\begin{cases} \rho = \sqrt{x^2 + y^2} \\ \phi = \arctan(y/x) \\ z = z \end{cases} \quad \Leftrightarrow \quad \begin{cases} x = \rho \cos\phi \\ y = \rho \sin\phi \\ z = z \end{cases} \tag{4.1}$$

Ce type de coordonnées permet de décrire de manière efficace un cylindre, comme présenté en Figure 4.1.

4.1.2 Coordonnées sphériques

Dans les *coordonnées sphériques*, $\boldsymbol{r}$ sera présenté par les paramètres (r, θ, ϕ) où :

$$\begin{cases} r = \sqrt{x^2 + y^2 + z^2} \\ \theta = \arctan(\sqrt{x^2 + y^2}/z) \\ \phi = \arctan(y/x) \end{cases} \quad \Leftrightarrow \quad \begin{cases} x = r \sin\theta \cos\phi \\ y = r \sin\theta \sin\phi \\ z = r \cos\theta \end{cases} \tag{4.2}$$

Ce type de coordonnées permet de décrire de manière efficace une sphère, comme présenté en Figure 4.2.

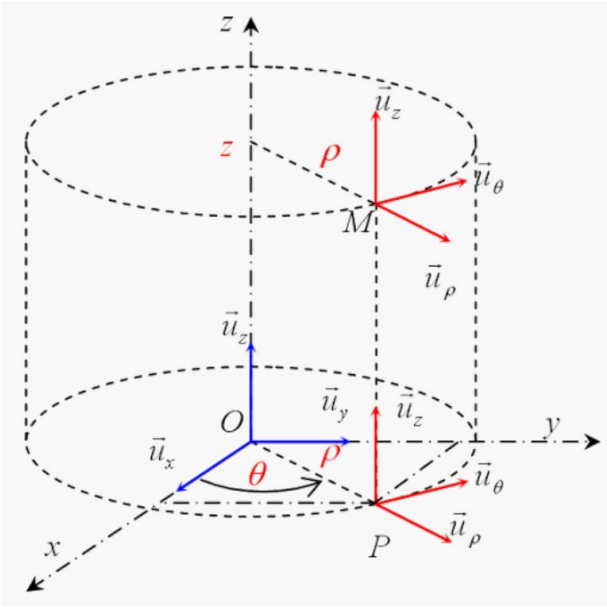


FIGURE 4.1 – Coordonnées cylindriques. Attention, leur θ correspond à notre ϕ. Crédit : Michel HENRY - Université du Maine

4.1.3 Bases mobiles

En coordonnées cartésiennes, un vecteur quelconque $\boldsymbol{p}$ avec des composantes (p_x, p_y, p_z) peut être présenté soit en colonne, soit par la décomposition dans les vecteurs de la base canonique

$$\boldsymbol{p} = p_x \boldsymbol{e}_x + p_y \boldsymbol{e}_y + p_z \boldsymbol{e}_z$$

La base $\{\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z\}$ est dite orthonormée car

$$||\boldsymbol{e}_x|| = ||\boldsymbol{e}_y|| = ||\boldsymbol{e}_z|| = 1 \tag{4.3}$$

$$\boldsymbol{e}_x.\boldsymbol{e}_y = \boldsymbol{e}_x.\boldsymbol{e}_z = \boldsymbol{e}_y.\boldsymbol{e}_z = 0 \tag{4.4}$$

On note généralement $\boldsymbol{e}_x = \boldsymbol{i}$, $\boldsymbol{e}_y = \boldsymbol{j}$ et $\boldsymbol{e}_z = \boldsymbol{k}$.

Il s'agit d'une *base fixe* : quelque soit le point de l'espace considéré, les vecteurs de base sont les même. Ce n'est pas le cas pour les coordonnées cylindriques et sphériques où les vecteurs de la base sont redéfinis à chaque point de l'espace, comme on peut le voir dans les Figures 4.1 et 4.2. C'est ce qu'on appelle une *base mobile*. Regardons maintenant comment obtenir ces bases.

Base cylindrique

On remarque d'abord que

$$\boldsymbol{r}(\rho, \phi, z) = \begin{pmatrix} x(\rho, \phi, z) \\ y(\rho, \phi, z) \\ z(\rho, \phi, z) \end{pmatrix} = \begin{pmatrix} \rho \cos\phi \\ \rho \sin\phi \\ z \end{pmatrix}. \tag{4.5}$$

Alors, le vecteur de base étant une petite variation dans la direction de la coordonnée d'intérêt, on l'obtient en dérivant le vecteur position dans cette même direction :

$$\boldsymbol{e}_\rho = \frac{\partial \boldsymbol{r}}{\partial \rho} = \frac{\partial}{\partial \rho} \begin{pmatrix} \rho \cos\phi \\ \rho \sin\phi \\ z \end{pmatrix} = \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix}. \tag{4.6}$$

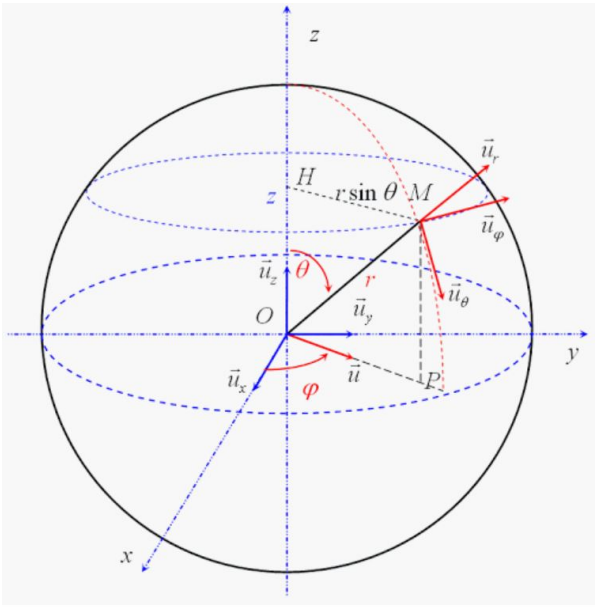


FIGURE 4.2 – Coordonnées sphériques. Crédit : Michel HENRY - Université du Maine

On conclut que

$$\boldsymbol{e}_\rho = \cos\phi \boldsymbol{e}_x + \sin\phi \boldsymbol{e}_y. \tag{4.7}$$

De la même manière

$$\boldsymbol{e}_\phi = \frac{\partial \boldsymbol{r}}{\partial \phi} = \frac{\partial}{\partial \phi}\begin{pmatrix}\rho\cos\phi\\ \rho\sin\phi\\ z\end{pmatrix} = \begin{pmatrix}-\rho\sin\phi\\ \rho\cos\phi\\ 0\end{pmatrix}, \tag{4.8}$$

à partir de quoi on conclut que

$$\boldsymbol{e}_\phi = -\rho\sin\phi \boldsymbol{e}_x + \rho\cos\phi \boldsymbol{e}_y. \tag{4.9}$$

Un calcul similaire permet de montrer que $\boldsymbol{e}_z$ est inchangé.

On peut alors aisément vérifier que

$$\boldsymbol{e}_\rho.\boldsymbol{e}_\phi = \boldsymbol{e}_\rho.\boldsymbol{e}_z = \boldsymbol{e}_\phi.\boldsymbol{e}_z = 0. \tag{4.10}$$

Par exemple, explicitement,

$$\boldsymbol{e}_\rho.\boldsymbol{e}_\phi = -\rho\cos\phi\sin\phi + \rho\cos\phi\sin\phi = 0. \tag{4.11}$$

La base est donc orthogonale. Il est souvent utile de travailler également avec une base normalisée (qu'on appelle orthonormale). Pour se faire, on définit

$$\hat{\boldsymbol{e}}_\rho = \frac{\boldsymbol{e}_\rho}{||\boldsymbol{e}_\rho||} \quad \hat{\boldsymbol{e}}_\phi = \frac{\boldsymbol{e}_\phi}{||\boldsymbol{e}_\phi||} \quad \hat{\boldsymbol{e}}_z = \frac{\boldsymbol{e}_z}{||\boldsymbol{e}_z||} \tag{4.12}$$

de sorte à ce que

$$||\hat{\boldsymbol{e}}_\rho|| = ||\hat{\boldsymbol{e}}_\phi|| = ||\hat{\boldsymbol{e}}_z|| = 1 \tag{4.13}$$

En conclusion, on peut retenir que la base cylindrique s'écrit en fonction des vecteurs de la base cartésienne comme

$$\boxed{\begin{aligned}\hat{\boldsymbol{e}}_\rho &= \cos\phi \boldsymbol{e}_x + \sin\phi \boldsymbol{e}_y \\ \hat{\boldsymbol{e}}_\phi &= -\sin\phi \boldsymbol{e}_x + \cos\phi \boldsymbol{e}_y \\ \hat{\boldsymbol{e}}_z &= \boldsymbol{e}_z\end{aligned}} \tag{4.14}$$

$$\hat{\boldsymbol{e}}_\phi = -\sin\phi \boldsymbol{e}_x + \cos\phi \boldsymbol{e}_y \tag{4.15}$$

$$\hat{\boldsymbol{e}}_z = \boldsymbol{e}_z \tag{4.16}$$

Base sphérique

De la même manière, on peut construire la base sphérique. On remarque d'abord que

$$\boldsymbol{r}(r,\theta,\phi) = \begin{pmatrix} x(r,\theta,\phi) \\ y(r,\theta,\phi) \\ z(r,\theta,\phi) \end{pmatrix} = \begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix}. \tag{4.17}$$

Alors,

$$\boldsymbol{e}_r = \frac{\partial \boldsymbol{r}}{\partial r} = \frac{\partial}{\partial r}\begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix} \tag{4.18}$$

$$\boldsymbol{e}_\theta = \frac{\partial \boldsymbol{r}}{\partial \theta} = \frac{\partial}{\partial \theta}\begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix} = \begin{pmatrix} r\cos\theta\cos\phi \\ r\cos\theta\sin\phi \\ -r\sin\theta \end{pmatrix} \tag{4.19}$$

$$\boldsymbol{e}_\phi = \frac{\partial \boldsymbol{r}}{\partial \phi} = \frac{\partial}{\partial \phi}\begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix} = \begin{pmatrix} -r\sin\theta\sin\phi \\ r\sin\theta\cos\phi \\ 0 \end{pmatrix} \tag{4.20}$$

On peut alors vérifier que

$$\boldsymbol{e}_r.\boldsymbol{e}_\theta = \boldsymbol{e}_r.\boldsymbol{e}_\phi = \boldsymbol{e}_\theta.\boldsymbol{e}_\phi = 0. \tag{4.21}$$

La base est donc orthogonale. Il reste à la normaliser :

$$\hat{\boldsymbol{e}}_r = \frac{\boldsymbol{e}_r}{||\boldsymbol{e}_r||} \quad \hat{\boldsymbol{e}}_\theta = \frac{\boldsymbol{e}_\theta}{||\boldsymbol{e}_\theta||} \quad \hat{\boldsymbol{e}}_\phi = \frac{\boldsymbol{e}_\phi}{||\boldsymbol{e}_\phi||} \tag{4.22}$$

de sorte à ce que

$$||\hat{\boldsymbol{e}}_r|| = ||\hat{\boldsymbol{e}}_\theta|| = ||\hat{\boldsymbol{e}}_\phi|| = 1 \tag{4.23}$$

En conclusion, on peut retenir que la base sphérique s'écrit en fonction des vecteurs de la base cartésienne comme

$$\boxed{\begin{aligned}\hat{\boldsymbol{e}}_r &= \sin\theta\cos\phi \boldsymbol{e}_x + \sin\theta\sin\phi \boldsymbol{e}_y + \cos\theta \boldsymbol{e}_z \\ \hat{\boldsymbol{e}}_\theta &= \cos\theta\cos\phi \boldsymbol{e}_x + \cos\theta\sin\phi \boldsymbol{e}_y - \sin\theta \boldsymbol{e}_z \\ \hat{\boldsymbol{e}}_\phi &= -\sin\phi \boldsymbol{e}_x + \cos\phi \boldsymbol{e}_y\end{aligned}} \tag{4.24}$$

$$\hat{\boldsymbol{e}}_\theta = \cos\theta\cos\phi \boldsymbol{e}_x + \cos\theta\sin\phi \boldsymbol{e}_y - \sin\theta \boldsymbol{e}_z \tag{4.25}$$

$$\hat{\boldsymbol{e}}_\phi = -\sin\phi \boldsymbol{e}_x + \cos\phi \boldsymbol{e}_y \tag{4.26}$$

4.1.4 Facteurs géométriques

Un dernier point important est le fait que les *facteurs géométriques* tels que la *mesure d'intégration* changent suivant le type de coordonnées utilisées. Nous nous en tiendrons ici à une description utilitaire. La *mesure d'intégration*, aussi appelé *volume élementaire*, est l'opérateur $\mathrm{d}V$ qui apparaît dans les intégrales multiples

$$\int_{\mathbb{R}^3} \mathrm{d}V f(\boldsymbol{r})$$

.

On a

$$\underline{\text{Coordonnées cartésiennes}} : \quad \mathrm{d}V = \mathrm{d}x\mathrm{d}y\mathrm{d}z \tag{4.27}$$

$$\underline{\text{Coordonnées cylindrique}} : \quad \mathrm{d}V = \rho\mathrm{d}\rho\mathrm{d}\phi\mathrm{d}z \tag{4.28}$$

$$\underline{\text{Coordonnées sphérique}} : \quad \mathrm{d}V = r^2 \sin\theta \mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi \tag{4.29}$$

4.2 Champ scalaire, Champ vectoriel

4.2.1 Champ scalaire

Un champ scalaire $f(\boldsymbol{r}) = f(x, y, z)$ est une fonction qui fait correspondre à chaque point de $\mathbb{R}^3$ un nombre réel ou complexe. On peut penser par exemple au champ de température qui a chaque point de l'espace associe une température. Il peut également s'agir du potentiel de Coulomb généré par une charge électrique q

$$U(\boldsymbol{r}) = \frac{q}{4\pi\varepsilon_0}\frac{1}{r} \tag{4.30}$$

où $r = ||\boldsymbol{r}||$ et ε_0 est la permittivité du vide. Enfin, on citera comme dernier exemple le cas du potentiel électrique créé par un petit dipôle placé à l'origine

$$U(\boldsymbol{r}) = \frac{q}{4\pi\varepsilon_0}\frac{\boldsymbol{p}\boldsymbol{r}}{r} \tag{4.31}$$

où $\boldsymbol{p}$ est le moment électrique du dipôle.

4.2.2 Champ vectoriel

Un champ vectoriel ou champ de vecteurs est une fonction qui fait correspondre à chaque point de $\mathbb{R}^3$ un vecteur. Cela permet par exemple de décrire la vitesse d'un fluide dans chaque direction en chacun de ses points. On peut également penser au champ électrique créé par une charge q placée à l'origine

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{\boldsymbol{r}}{r^3} \tag{4.32}$$

On remarque que

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{1}{r^3}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} E_x(x,y,z) \\ E_y(x,y,z) \\ E_z(x,y,z) \end{pmatrix}. \tag{4.33}$$

Autrement dit, un champ vectoriel tel que $\boldsymbol{E}$ est un ensemble de trois champs scalaires E_x, E_y et E_z.

4.3 Dérivées

4.3.1 Dérivées partielles

La dérivée partielle d'une fonction de plusieurs variables est définie d'une façon très comparable au cas des fonctions d'une seule variable ; la dérivée partielle $\frac{\partial f}{\partial x}$ de la fonction f par rapport à la variable x est définie par la relation :

$$\boxed{\frac{\partial f}{\partial x} \equiv \partial_x f = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y, z) - f(x, y, z)}{\Delta x}}$$

et de même pour les autres composantes. Cette définition montre que la dérivée partielle est évaluée lorsqu'une seule variable varie, les autres restant fixées pendant l'opération.

En général les dérivées partielles peuvent elles-même dépendre des variables x, y, z. Il est donc possible de définir par le même procédé des dérivées successives d'une variable : $\frac{\partial^2 f}{\partial x^2}$, et à la différence des fonctions d'une seule variable des dérivées croisées : $\frac{\partial^2 f}{\partial x \partial y}$. Une question très naturelle se pose de savoir si l'ordre dans lequel on prend les dérivées importe. Le résultat mathématique précis est le suivant :

Théorème 3 Si $f \in \mathcal{C}^2$, l'ensemble des fonctions deux fois dérivables aux dérivées secondes continues, alors $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$.

4.3.2 Fonctions composées

Considérons la fonction $f(x, y, z)$ dans le cas où x, y et z sont elles-mêmes des fonctions d'une autre variable que nous noterons t. Cette situation se rencontre fréquemment en mécanique où $\vec{r}(t) \equiv (x(t), y(t), z(t))$ décrit la trajectoire d'une particule au cours du temps. On peut donc définir une fonction $g = f \circ \vec{r}$ dépendant de la variable t par la relation :

$$g(t) \equiv f\left(x(t), y(t), z(t)\right)$$

Il est facile d'établir le résultat suivant qui généralise la relation bien connue pour la composition de 2 fonctions u et v d'une seule variable : $(u \circ v)' = (u' \circ v)\, v'$

Théorème 4 Supposons que $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ et $\frac{\partial f}{\partial z}$ soient continues et x, y, z des fonctions dérivables de la variable t, alors la fonction $g : t \mapsto f\left(x(t), y(t), z(t)\right)$ est dérivable et sa dérivée vaut :

$$\frac{dg}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}.$$

Un abus de notation, très fréquent en physique, consiste à confondre les fonctions f et g, ce qui conduit à écrire la relation précédente sous la forme :

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

Cette pratique est justifiée dans le cadre de la physique, où les symboles mathématiques sont associés à des grandeurs physiques bien déterminées, mais il convient toutefois d'être conscient que les *fonctions* f et g sont différentes. Par opposition aux dérivées partielles, df/dt est appelée *dérivée totale.*

Une situation un peu plus générale peut être rencontrée lorsque les fonctions x, y et z dépendent elles-mêmes de plusieurs variables. Par exemple, à partir de la fonction $f(x, y, z)$, lorsque $x = x(s, t), y = y(s, t), z = z(s, t)$, on peut définir une fonction des 2 variables s et t :

$$g(s, t) \equiv f\left(x(s, t), y(s, t), z(s, t)\right)$$

qui admet donc 2 dérivées partielles données par les relations

$$\begin{aligned}
\frac{\partial g}{\partial s} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial s}, \\
\frac{\partial g}{\partial t} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t}.
\end{aligned}$$

Ce genre de dérivée en chaîne, porte le nom de *chain rule* dans la littérature anglo-saxonne.

4.3.3 Application : dérivée totale de $T(\vec{r}(t), t)$

Supposons que l'on mesure la température T à intervalles de temps réguliers le long d'une trajectoire définie par la fonction $\vec{r}(t)$. On connaît donc la fonction $T(\vec{r}(t), t)$; les variations infinitésimales de température sont données par la dérivée totale (avec l'abus de notation signalé plus haut) :

$$\frac{dT}{dt} = \frac{\partial T}{\partial t}\frac{dt}{dt} + \frac{\partial T}{\partial x}\frac{dx}{dt} + \frac{\partial T}{\partial y}\frac{dy}{dt} + \frac{\partial T}{\partial z}\frac{dz}{dt},$$

qu'il est plus courant d'écrire sous la forme[1] :

$$\frac{dT}{dt} = \frac{\partial T}{\partial t} + \frac{d\vec{r}}{dt}.\overrightarrow{\nabla} T,$$

où l'on a introduit le gradient de T : $\overrightarrow{\nabla} T \equiv (\partial T/\partial x, \partial T/\partial y, \partial T/\partial z)$, objet que l'on va être amené à manipuler très prochainement.

Cette forme met clairement en évidence le fait que les variations de température pendant le temps dt ont 2 origines distinctes, une qui provient des variations de T au point considéré, et une autre qui est une conséquence du déplacement effectué à la vitesse $\overrightarrow{v} = \frac{d\vec{r}}{dt}$ pendant le temps dt.

En mécanique des fluides, cette dérivée totale est appelée dérivée particulaire ou dérivée convective. Si $\overrightarrow{A}(\overrightarrow{r}(t), t)$ est une fonction vectorielle, alors $\frac{d\overrightarrow{A}}{dt} = \frac{\partial \overrightarrow{A}}{\partial t} + (\overrightarrow{v}.\overrightarrow{\nabla})\overrightarrow{A}$, $\overrightarrow{v}.\overrightarrow{\nabla}$ étant un opérateur qui joue un rôle fondamental en mécanique des fluides.

1. pour souligner la différence avec les dérivées partielles, on note quelquefois les dérivées totales par le symbole D/Dt

Chapitre 5

Opérateurs différentiels

Les opérateurs différentiels agissent comme leur nom l'indique par dérivation sur des champs scalaires ou vectoriels (voire tensoriels). Ces opérateurs sont dits du 1er ordre si leur définition ne met en jeu que des dérivées partielles du premier ordre, du second ordre si apparaîssent des dérivées secondes...

Les opérateurs différentiels étudiés sont des opérateurs locaux; on entend par là que les opérateurs sont définis en tout point de l'espace.

5.1 Opérateurs usuels

Nous rappelons la définition usuelle des opérateurs *gradient*, *divergence* et *rotationnel* en coordonnées cartésiennes par rapport à la base canonique $(\overrightarrow{e}_x, \overrightarrow{e}_y, \overrightarrow{e}_z)$:

$$\begin{array}{rcl} \overrightarrow{\text{grad}}\, f(\overrightarrow{r}) & \equiv & \overrightarrow{\nabla}\, f(\overrightarrow{r}) \equiv \frac{\partial f}{\partial x}\,\overrightarrow{e}_x + \frac{\partial f}{\partial y}\,\overrightarrow{e}_y + \frac{\partial f}{\partial z}\,\overrightarrow{e}_z \\ \text{div}\,\overrightarrow{F}(\overrightarrow{r}) & \equiv & \overrightarrow{\nabla}\,.\,\overrightarrow{F}(\overrightarrow{r}) \equiv \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \\ \overrightarrow{\text{rot}}\,\overrightarrow{F}(\overrightarrow{r}) & \equiv & \overrightarrow{\nabla} \wedge \overrightarrow{F}(\overrightarrow{r}) \equiv \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\overrightarrow{e}_x + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)\overrightarrow{e}_y + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\overrightarrow{e}_z \end{array}$$

L'opérateur différentiel du second ordre le plus courant est le *laplacien*, qui s'applique aux fonctions scalaires ou vectorielles :

$$\begin{array}{rcl} \triangle f(\overrightarrow{r}) & \equiv & \text{div}\left[\overrightarrow{\text{grad}}\, f(\overrightarrow{r})\right] \equiv \overrightarrow{\nabla}\,.\,\overrightarrow{\nabla}\, f(\overrightarrow{r}) \equiv \nabla^2 f(\overrightarrow{r}) \equiv \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \\ \overrightarrow{\triangle}\,\overrightarrow{F}(\overrightarrow{r}) & \equiv & \triangle F_x(\overrightarrow{r})\,\overrightarrow{e}_x + \triangle F_y(\overrightarrow{r})\,\overrightarrow{e}_y + \triangle F_z(\overrightarrow{r})\,\overrightarrow{e}_z \end{array}$$

L'opérateur de l'équation des ondes, ou d'*alembertien*, noté $\Box$, qui agit sur les fonctions scalaires des variables $\overrightarrow{r}$ et t est défini par la relation

$$\Box \equiv \triangle - \frac{1}{c^2}\,\frac{\partial^2}{\partial t^2}$$

On définit également un d'*alembertien* vectoriel, qui lui agit sur des fonctions vectorielles, à l'aide du laplacien vectoriel.

Dans toutes ces définitions, il convient de noter le rôle permanent de l'opérateur $\overrightarrow{\nabla} \equiv (\frac{\partial}{\partial x}, \frac{\partial}{\partial x}, \frac{\partial}{\partial z})$ (on dit "nabla"). La notation la plus répandue des opérateurs différentiels utilise cet opérateur, et non pas les notations $\overrightarrow{\text{grad}}$, div, $\overrightarrow{\text{rot}}$ qui ne sont guère utilisées qu'en France.

La valeur prise par ces opérateurs, lorsqu'ils sont appliqués à une fonction donnée, est indépendante du système de coordonnées. Ci-dessous, on donne l'expression de ces opérateurs dans les systèmes de coordonnées cartésiennes, cylindriques et sphériques.

	cartésien (x, y, z)	cylindrique (ρ, ϕ, z)	sphérique (r, θ, ϕ)
$\overrightarrow{\text{grad}} f$	$\begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix}$	$\begin{pmatrix} \frac{\partial f}{\partial \rho} \\ \frac{1}{\rho}\frac{\partial f}{\partial \phi} \\ \frac{\partial f}{\partial z} \end{pmatrix}$	$\begin{pmatrix} \frac{\partial f}{\partial r} \\ \frac{1}{r}\frac{\partial f}{\partial \theta} \\ \frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi} \end{pmatrix}$
$\text{div}\overrightarrow{F}$	$\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$	$\frac{1}{\rho}\frac{\partial(\rho F_\rho)}{\partial \rho} + \frac{1}{\rho}\frac{\partial F_\phi}{\partial \phi} + \frac{\partial F_z}{\partial z}$	$\frac{1}{r^2}\frac{\partial(r^2 F_r)}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial(\sin\theta F_\theta)}{\partial \theta} + \frac{1}{r\sin\theta}\frac{\partial F_\phi}{\partial \phi}$
$\overrightarrow{\text{rot}}\overrightarrow{F}$	$\begin{pmatrix} \frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \\ \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \\ \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{\rho}\frac{\partial F_z}{\partial \phi} - \frac{\partial F_\phi}{\partial z} \\ \frac{\partial F_\rho}{\partial z} - \frac{\partial F_z}{\partial \rho} \\ \frac{1}{\rho}\frac{\partial(\rho F_\phi)}{\partial \rho} - \frac{1}{\rho}\frac{\partial F_\rho}{\partial \phi} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{r\sin\theta}\frac{\partial(\sin\theta F_\phi)}{\partial \theta} - \frac{1}{r\sin\theta}\frac{\partial F_\theta}{\partial \phi} \\ \frac{1}{r\sin\theta}\frac{\partial F_r}{\partial \phi} - \frac{\partial(rF_\phi)}{\partial r} \\ \frac{1}{r}\frac{\partial(rF_\theta)}{\partial r} - \frac{1}{r}\frac{\partial F_r}{\partial \theta} \end{pmatrix}$
Δf	$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$	$\frac{\partial^2 f}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial f}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2}$	$\frac{\partial^2 f}{\partial r^2} + \frac{2}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\cos\theta}{r^2\sin\theta}\frac{\partial f}{\partial \theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \phi^2}$

5.2 Identités vectorielles

Il existe un certain nombre d'identités vectorielles qui permettent de simplifier les calculs. Tout d'abord, on retiendra que

$$\begin{aligned} \text{div}\mathbf{grad}(f) &= \Delta f & (5.1)\\ \mathbf{rot}.\mathbf{grad}(f) &= \mathbf{0} & (5.2)\\ \text{div}.\mathbf{rot}(\boldsymbol{A}) &= 0 & (5.3) \end{aligned}$$

Ensuite, il peut-être utile de savoir que

$$\mathbf{grad}(fg) = g\mathbf{grad}(f) + f\mathbf{grad}(g) \tag{5.4}$$

où f et g sont des champs scalaires. On a également

$$\begin{aligned} \text{div}(f\boldsymbol{A}) &= \mathbf{grad}(f).\boldsymbol{A} + f.\text{div}(\boldsymbol{A}) & (5.5)\\ \mathbf{rot}(f\boldsymbol{A}) &= \mathbf{grad}(f) \wedge \boldsymbol{A} + f.\mathbf{rot}(\boldsymbol{A}) & (5.6) \end{aligned}$$

avec $\boldsymbol{A}$ un champ vectoriel.

Enfin, on a également

$$\mathbf{rot}(\mathbf{rot}\boldsymbol{A}) = \mathbf{grad}(\text{div}\boldsymbol{A}) - \Delta\boldsymbol{A}. \tag{5.7}$$

5.3 Un exemple en électrostatique : cas d'une boule chargée uniformément

Les opérateurs différentiels présentés dans ce chapitre prennent toute leur utilité en hydrodynamique et en électromagnétisme où ils sont des éléments centraux du formalisme. Pour vous en convaincre, voici un court exemple en électrostatique.

Les équations de Maxwell qui permettent d'obtenir les champs électrique $\boldsymbol{E}(\boldsymbol{r})$ et magnétique $\boldsymbol{B}(\boldsymbol{r})$ en tout point de l'espace s'expriment simplement à l'aide des opérateurs introduits dans ce

chapitre.

$$\begin{aligned} \mathrm{div}\boldsymbol{E} &= \frac{\rho}{\varepsilon_0}; & \mathrm{div}\boldsymbol{B} &= 0; \\ \mathbf{rot}\boldsymbol{E} &= -\frac{\partial \boldsymbol{B}}{\partial t}; & \mathbf{rot}\boldsymbol{B} &= \mu_0 \boldsymbol{j} + \frac{1}{c^2}\frac{\partial \boldsymbol{E}}{\partial t} \end{aligned}$$

où $\rho(\boldsymbol{r})$ est la densité de charge, $\boldsymbol{j}(\boldsymbol{r})$ la densité de charge, c la vitesse de la lumière, ε_0 le permittivité du vide et μ_0 la perméabilité du vide.

L'électrostatique est un cas particulier pour lequel :

- le champ magnétique est nul : $\boldsymbol{B}(\boldsymbol{r}) = 0$
- le champ électrique est statique : $\partial_t \boldsymbol{E} = 0$
- il n'y a pas de courant : $\boldsymbol{j}(\boldsymbol{r}) = 0$

Il reste alors à résoudre

$$\begin{cases} \mathrm{div}\boldsymbol{E} &= \rho/\varepsilon_0 \\ \mathbf{rot}\boldsymbol{E} &= \mathbf{0} \end{cases} \tag{5.8}$$

Généralement, la distribution de charge est connue et il faut déterminer le champ électrique généré par cette distribution. On remarque qu'étant donné que $\mathbf{rot}\boldsymbol{E} = \mathbf{0}$, on peut utiliser l'Equation 5.2 pour déduire que $\boldsymbol{E}$ peut s'écrire comme [1]

$$\boldsymbol{E}(\boldsymbol{r}) = -\mathbf{grad}V(\boldsymbol{r}) \tag{5.10}$$

où $V(\boldsymbol{r})$ est le potentiel électrique. Alors $\mathrm{div}\boldsymbol{E} = \rho/\varepsilon_0$ devient

$$-\mathrm{div}(\mathbf{grad}V(\boldsymbol{r})) = \rho/\varepsilon_0 \tag{5.11}$$

Finalement, avec l'aide de l'Equation 5.1, on que l'on doit résoudre le système d'équations

$$\begin{cases} \Delta V(\boldsymbol{r}) &= -\rho/\varepsilon_0 \\ \boldsymbol{E}(\boldsymbol{r}) &= -\mathbf{grad}V(\boldsymbol{r}) \end{cases} \tag{5.12}$$

Prenons alors l'exemple d'une boule chargée uniformément de rayon a :

$$\rho(\boldsymbol{r}) = \rho(r) = \begin{cases} \rho_0 & \text{pour } \ r \leq a \\ 0 & \text{pour } \ r > a \end{cases} \tag{5.13}$$

Nous allons chercher le potentiel électrique $V(\boldsymbol{r})$ généré par la boule chargée. Tout d'abord, $\rho(r)$ est une fonction radiale : elle ne dépend que de r et pas de θ ou ϕ. Par conséquent, $V(r)$ doit également être une fonction radiale. On conclut que le laplacien réduit à

$$\Delta V(r) = \frac{1}{r^2}\partial_r(r^2\partial_r V(r)) \tag{5.14}$$

1. De la même manière, en dehors du cadre de l'électrostatique, quand $\boldsymbol{B}(\boldsymbol{r}) \neq 0$, comme $\mathrm{div}\boldsymbol{B} = 0$, on peut conclure en utilisant l'Equation 5.3 que

$$\boldsymbol{B}(\boldsymbol{r}) = \mathbf{rot}\boldsymbol{A}(\boldsymbol{r}) \tag{5.9}$$

On doit donc résoudre

$$\begin{cases} \frac{1}{r^2}\partial_r(r^2\partial_r V(r)) = -\rho_0/\varepsilon_0 & \text{pour } \; r \leq a \\ \frac{1}{r^2}\partial_r(r^2\partial_r V(r)) = 0 & \text{pour } \; r > a \end{cases} \tag{5.15}$$

Pour $r \leq a$, la première équation devient

$$\partial_r(r^2\partial_r V(r)) = -\frac{\rho_0}{\varepsilon_0} r^2 \tag{5.16}$$

$$r^2\partial_r V(r) = -\frac{\rho_0}{\varepsilon_0}\frac{r^3}{3} + C_1 \tag{5.17}$$

$$\partial_r V(r) = -\frac{\rho_0}{\varepsilon_0}\frac{r}{3} + \frac{C_1}{r^2} \tag{5.18}$$

$C_1 = 0$ pour des raisons physiques car sinon, $V(r)$ divergerait à l'origine en $r = 0$. Finalement, on conclut que

$$V(r) = -\frac{\rho_0}{\varepsilon_0}\frac{r^2}{6} + C_2. \tag{5.19}$$

On garde C_2 pour le moment.

Pour la région $r > a$, on a

$$\partial_r(r^2\partial_r V(r)) = 0 \tag{5.20}$$

$$r^2\partial_r V(r) = C_3 \tag{5.21}$$

$$\partial_r V(r) = \frac{C_3}{r^2} \tag{5.22}$$

On garde C_3 car $r > a > 0$ donc il n'y pas de risque de divergence. On conclut que

$$V(r) = -\frac{C_3}{r} + C_4. \tag{5.23}$$

On a la liberté de choisir C_4 car $V(r)$ n'apparaissant physiquement que comme $\boldsymbol{E} = -\mathbf{grad}V$, le potentiel est défini à une constante prêt [2]. On choisit donc $C_4 = 0$ de sorte à ce que $V(r) \to 0$ quand $r \to \infty$.

Il ne reste plus qu'à assurer la continuité de $\partial_r V(r)$ et de $V(r)$. Pour $\partial_r V(r)$,

$$\partial_r V(r)|_a = -\frac{\rho_0}{\varepsilon_0}\frac{a}{3} = \frac{C_3}{a^2} \tag{5.24}$$

duquel on déduit que

$$C_3 = -\frac{\rho_0}{\varepsilon_0}\frac{a^3}{3}. \tag{5.25}$$

Pour $V(r)$,

$$V(r)|_a = -\frac{\rho_0}{\varepsilon_0}\frac{a^2}{6} + C_2 = -\frac{C_3}{a} = \frac{\rho_0}{\varepsilon_0}\frac{a^2}{3} \tag{5.26}$$

Donc

$$C_2 = \frac{\rho_0}{\varepsilon_0}\frac{a^3}{2} + \frac{\rho_0}{\varepsilon_0}\frac{a^2}{6} = \frac{\rho_0}{\varepsilon_0}\frac{a^2}{2}. \tag{5.27}$$

En injectant ces résultats dans les expressions du potentiel, on trouve finalement que

2. Dans le sens où $\boldsymbol{E} = -\mathbf{grad}V = -\mathbf{grad}V'$ si $V' = V + C$ où C est une constante

- Pour $r \leq a$:

$$V(r) = \frac{\rho_0}{\varepsilon_0}\left(\frac{a^2}{2} - \frac{r^2}{6}\right). \tag{5.28}$$

- Pour $r > a$:

$$V(r) = \frac{\rho_0}{\varepsilon_0}\frac{a^3}{3r}. \tag{5.29}$$

La Figure 5.1 permet de visualiser ce potentiel.

Enfin, comme $V(r)$ est radial, on note que $\boldsymbol{E}(\boldsymbol{r}) = E(r)\boldsymbol{e}_r$ avec $E(r) = -\partial_r V(r)$. Par conséquent

- Pour $r \leq a$:

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{\rho_0}{\varepsilon_0}\frac{\boldsymbol{r}}{3}. \tag{5.30}$$

- Pour $r > a$:

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{\rho_0}{\varepsilon_0}\frac{a^3}{3}\frac{\boldsymbol{r}}{r^3}. \tag{5.31}$$

La Figure 5.2 permet de visualiser ce le champ électrique. On note que que le champ électrique présente une discontinuité en $r = a$.

Enfin, nous allons réécrire $V(r)$ et $\boldsymbol{E}(\boldsymbol{r})$ comme exprimés en fonction de q la charge totale de la boule dans la région $r > a$:

$$q = \frac{4}{3}\pi a^3 \rho_0 \qquad \Leftrightarrow \qquad \rho_0 = \frac{3}{4\pi}\frac{q}{a^3} \tag{5.32}$$

En utilisant l'expression pour ρ_0, on trouve alors

$$V(r) = \frac{q}{4\pi\varepsilon_0}\frac{1}{r} \tag{5.33}$$

l'habituel potentiel de Coulomb créé par une charge q. Pour le champ électrique, on trouve comme attendu :

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{q}{4\pi\varepsilon_0}\frac{\boldsymbol{r}}{r^3}. \tag{5.34}$$

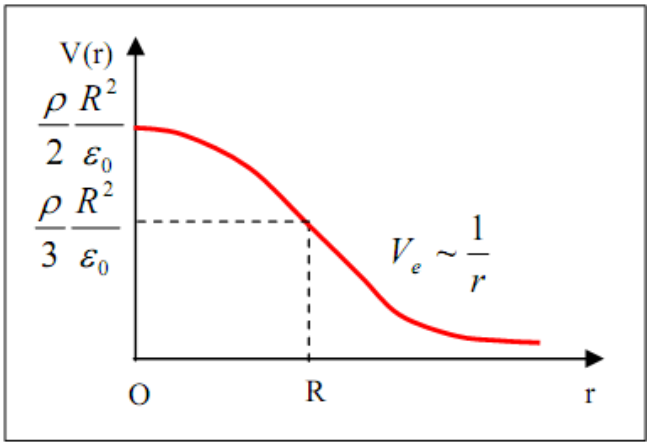


FIGURE 5.1 – Potentiel électrique $V(r)$ engendré par une boule chargée de rayon R. Attention, leur R correspond à notre a. Crédit : www.cours-et-exercices.com

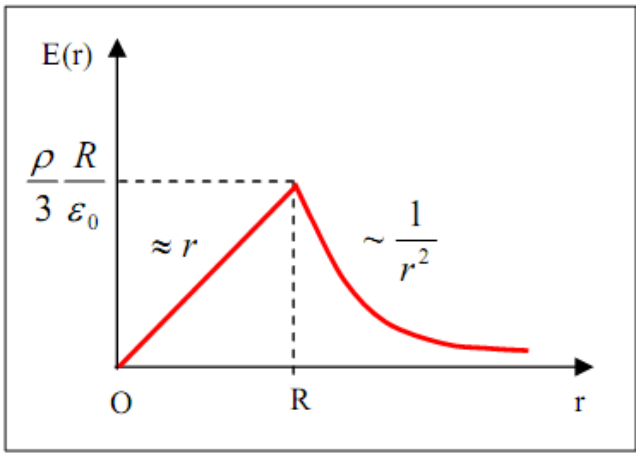


FIGURE 5.2 – Champ électrique $E(r)$ engendré par une boule chargée de rayon R. Attention, leur R correspond à notre a. Crédit : www.cours-et-exercices.com

Chapitre 6

Théorèmes de Green-Ostrogradski et de Stokes

L'un des résultats fondamentaux relatif à l'intégration des fonctions d'une variable est la formule

$$\int_a^b f'(x)\,dx = f(b) - f(a).$$

Cette relation admet une généralisation en dimension supérieure connue sous le nom de *formule de Stokes*. On verra que les rôles de l'intervalle $[a,b]$ et du couple de point (a,b) seront tenus par une variété (volume, surface) et son bord (surface fermée, contour fermé), tandis que le rôle de f' sera tenu par les opérateurs différentiels divergence et rotationnel.

Bien qu'une formulation unifiée des deux théorèmes utilisés en Physique :

- le théorème de Green-Ostrogradski (appelé aussi théorème de Gauss ou théorème de la divergence)

- le théorème de Stokes

soit possible, nous privilégierons dans la suite deux énoncés distincts pour ces deux théorèmes afin de se rapprocher de leur usage en Physique. Avant cela, il nous faut définir l'intégration dans $\mathbb{R}^2$ et $\mathbb{R}^3$

6.1 Intégrales doubles et intégrales triples

Commençons par un bref rappel sur les intégrales dans $\mathbb{R}$, $\mathbb{R}^2$ et $\mathbb{R}^3$, c'est-à-dire sur le calcul des nombres notés

$$\int_{U_1} f(x)\,dx, \quad \int_{U_2} f(x,y)\,dS, \quad \int_{U_3} f(x,y,z)\,dV$$

D'un point de vue pratique, on se débrouille en général pour ramener le calcul des intégrales à 2 ou 3 dimensions au calcul d'intégrales à une seule dimension. Dans cette réduction, un point capital tient dans le fait que les éléments différentiels dS et dV peuvent s'interpréter comme des éléments de surface $dS = dxdy$ et de volume $dV = dxdydz$. Soit un domaine $D = \{(x,y) \in \mathbb{R}^2\ ,\ a \leq x \leq b\ ,\ g(x) \leq y \leq h(x)\}$ où g et h sont des fonctions continues sur l'intervalle $]a,b[$,

$$\int\int_D f(x,y)dxdy = \int_a^b \left(dx \int_{g(x)}^{h(x)} f(x,y)dy \right)$$

On obtient une formule équivalente en échangeant le rôle de x et y. La formule se généralise aisément à 3 dimensions.

6.1.1 Circulation et intégrale curviligne

Supposons que l'on s'intéresse à l'intégrale $\int_{\mathcal{C}} \vec{F}(\vec{r}).\, d\vec{r}$ qui correspond à la circulation du vecteur $\vec{F}$ le long de la trajectoire représentée par le chemin orienté $\mathcal{C}$ (un travail si $\vec{F}$ est une force) [1].

Il existe deux façons équivalentes de calculer cette intégrale selon l'information dont on dispose :
1. Si la trajectoire (dans le plan $0xy$ pour simplifier) est connue sous la forme d'une fonction $y = f(x)$, il suffit d'exprimer le produit scalaire dans la base cartésienne $(\vec{e}_x, \vec{e}_y)$ pour ramener l'intégrale à une somme d'intégrales d'une seule variable (x ou y) :

$$\int_{\mathcal{C}} \vec{F}(\vec{r}).\, d\vec{r} = \int_{\mathcal{C}} F_x(x,y)\, dx + \int_{\mathcal{C}} F_y(x,y)\, dy$$

Comme les composantes du vecteur $\vec{F}$ dépendent en général des variables x et y, il convient dans ce calcul d'utiliser l'équation de la courbe $y = f(x)$ afin d'éliminer x ou y dans chacune des intégrales. Selon la géométrie du problème étudié, on peut également être amené à utiliser un autre système de coordonnées.

2. Si la trajectoire associée au chemin $\mathcal{C}$ est connue sous forme paramétrique (c'est souvent le cas en mécanique), la position le long de la trajectoire est donnée par le vecteur $\vec{r}(t) = x(t)\,\vec{e}_x + y(t)\,\vec{e}_y + z(t)\,\vec{e}_z$, où x, y et z sont des fonctions connues du paramètre t qui varie entre les deux valeurs t_a et t_b correspondant respectivement aux extrémités du chemin $\mathcal{C}$. Dans ce cas l'intégrale est calculée par la formule suivante :

$$\int_{\mathcal{C}} \vec{F}(\vec{r}).\, d\vec{r} \equiv \int_{t_a}^{t_b} \vec{F}(\vec{r}).\, \frac{d\vec{r}}{dt}\, dt$$

Si t est interprété comme un temps, on peut assimiler le vecteur $\vec{v} \equiv d\vec{r}/dt$ à une vitesse. Le produit scalaire $\vec{F}.\,\vec{v}$ peut alors être calculé dans un système de coordonnées adaptées de sorte que l'intégrale curviligne s'écrit encore une fois comme l'intégrale d'une fonction d'une seule variable.

En général la valeur de $\int_{\mathcal{C}} \vec{F}(\vec{r}).\, d\vec{r}$ dépend du chemin $\mathcal{C}$; nous considérerons plus loin le cas important des intégrales curvilignes dont le résultat ne dépend pas du chemin d'intégration.

6.1.2 Flux et intégrale surfacique

Le flux du champ $\overrightarrow{F}$ à travers une surface S régulière [2] (ou au moins régulière par morceaux) vaut

$$\int_S \overrightarrow{F}(\overrightarrow{r}).\, \overrightarrow{dS}$$

$\overrightarrow{dS}$ est un élément différentiel de surface appartenant à la surface S, de direction normale à cette surface, et orienté vers l'extérieur de celle-ci.

Il existe un lien profond entre le flux d'un champ à travers une surface et sa divergence. On dit que la divergence est la limite infinitésimale du flux car

$$\mathrm{d}F_{\mathcal{S}}\,[\boldsymbol{A}] = \mathrm{div}\boldsymbol{A}(\boldsymbol{r}).\mathrm{d}V \tag{6.1}$$

où $\mathrm{d}F_{\mathcal{S}}$ est une petite variation du flux à travers la surface $\mathcal{S}$.

1. Ce type d'intégrale est appelé "intégrale curviligne" ou "intégrale de chemin" par les physiciens. Les mathématiciens nomment "intégrale curviligne" $\int_{\mathcal{C}} f(\vec{r})\, dr$ où f est un champ scalaire et dr un élément de longueur infinitésimal.

2. Une surface régulière est une surface qui admet un plan tangent en tout point.

6.2 Théorèmes de Green-Ostrogradski et de Stokes

6.2.1 Théorèmes de Green-Ostrogradski

Théorème 5 Soit S une surface fermée dont les surfaces élémentaires $\overrightarrow{dS}$ sont orientées vers l'extérieur, régulière par morceaux, qui contient le volume V. Si $\overrightarrow{F}(\overrightarrow{r})$ possède des dérivées partielles continues dans V, alors

$$\oint_S \overrightarrow{F} . \overrightarrow{dS} = \iiint_V \left(\overrightarrow{\nabla} . \overrightarrow{F} \right) dV$$

Ce théorème relie la divergence d'un champ de vecteur, intégrée sur un volume V, au flux de ce vecteur à travers la surface S qui délimite ce volume V. Alors que l'intégrale volumique fait intervenir les valeurs de div $\overrightarrow{F}$ en chaque point du volume, l'intégrale surfacique ne fait appel qu'aux valeurs de $\overrightarrow{F}$ sur la périphérie du volume.

Propriété locale de l'opérateur *divergence* :

Considérons un point quelconque de l'espace et, autour de ce point, un volume $\Delta\Omega$ (dont la forme pourra dépendre des symétries du problème étudié) limité par une surface fermée S.

Grâce à ce théorème, on voit que la divergence d'un champ vectoriel peut être définie par la relation suivante :

$$\text{div } \overrightarrow{F}(\overrightarrow{r}) \equiv \lim_{\Delta\Omega \to 0} \frac{\oint_S \overrightarrow{F}(\overrightarrow{r}) \, . \, \overrightarrow{dS}}{\Delta\Omega}$$

Il apparaît que les notions de divergence et de flux sont intimement liées puisque *la divergence d'un champ de vecteurs en un point P de l'espace est égale au flux sortant par les surfaces d'un volume infinitésimal situé autour de P, par unité de volume.*

6.2.2 Théorème de Stokes

Théorème 6 Soit S une surface orientée, régulière par morceaux, et limitée par une courbe fermée $\mathcal{C}$. L'orientation du vecteur élémentaire $\overrightarrow{dS}$ est donnée par la règle du tire-bouchon appliquée au sens de parcours de $\mathcal{C}$. Si $\overrightarrow{F}(\overrightarrow{r})$ possède des dérivées partielles continues dans S, alors

$$\oint_{\mathcal{C}} \overrightarrow{F} . d\overrightarrow{r} = \iint_S \left(\overrightarrow{\nabla} \wedge \overrightarrow{F} \right) . \overrightarrow{dS}$$

Ce théorème relie le rotationnel d'un champ de vecteur intégré sur une surface S à la circulation de ce vecteur sur le contour $\mathcal{C}$ délimitant cette surface. Alors que l'intégrale surfacique fait intervenir les valeurs de $\overrightarrow{\text{rot}}\ \overrightarrow{F}$ en chaque point de la surface, l'intégrale de chemin ne fait appel qu'aux valeurs de $\overrightarrow{F}$ sur la périphérie de la surface.

Propriété locale de l'opérateur *rotationnel* :

Soit une surface ΔS autour du point considéré limitée par un contour fermé $\mathcal{C}$. La composante du rotationnel dans la direction de la normale $\overrightarrow{n}$ à l'élément de surface $\Delta \overrightarrow{S} = \Delta S \, \overrightarrow{n}$ peut être obtenue par la relation :

$$\left(\overrightarrow{\nabla} \wedge \overrightarrow{F} \right) . \overrightarrow{n} = \lim_{\Delta S \to 0} \frac{\oint_{\mathcal{C}} \overrightarrow{F} . d\overrightarrow{r}}{\Delta S}$$

On peut donc dire que *la composante du rotationnel d'un champ de vecteurs, en un point P de l'espace, dans la direction arbitraire $\vec{n}$, est égale à la circulation le long d'un contour infinitésimal situé autour de P dans un plan orthogonal à $\vec{n}$, par unité de surface.*

6.3 Un exemple en électrostatique : le théorème de Gauss

De manière générale, le théorème de Green-Ostrogradski reflète une loi de conservation et apparaît dans de très nombreux domaines de la physique comme l'hydrodynamique ou l'électromagnétisme. Le théorème indique qu'une dispersion au sein d'un volume (exprimée par la divergence) s'accompagne nécessairement d'un flux total équivalent sortant de sa frontière.

On peut donc l'utiliser pour obtenir une version intégrale de l'*équation de Maxwell-Gauss*

$$\mathrm{div}\boldsymbol{E} = \frac{\rho}{\varepsilon_0} \tag{6.2}$$

En intégrant sur tout l'espace, on trouve

$$\iiint_V \mathrm{div}\boldsymbol{E}\mathrm{d}V = \frac{1}{\varepsilon_0}\iiint_V \rho \mathrm{d}V \tag{6.3}$$

$$= \frac{Q_{\mathrm{int}}}{\varepsilon_0}. \tag{6.4}$$

De plus, en utilisant le théorème de Green-Ostrogradski, on observe que

$$\iiint_V \mathrm{div}\boldsymbol{E}\mathrm{d}V = \oint_S \boldsymbol{E}.\boldsymbol{dS} \tag{6.5}$$

On en conclut que

$$\oint_S \boldsymbol{E}.\boldsymbol{dS} = \frac{Q_{\mathrm{int}}}{\varepsilon_0}. \tag{6.6}$$

Ainsi, le flux du champ électrique à travers une surface fermée S contenant un volume V est proportionnelle à la charge contenu dans ce volume. C'est le *théorème de Gauss*, particulièrement utile pour calculer le champ électrique en présence de symétries.