

Université
de Liège



Faculté des Sciences

MATHEMATIQUE

« partim B »

Premier Bloc en

Biologie, Chimie, Géologie, Informatique, Physique

Introduction au cours

« Mathématique », partim B

Aperçu général du cours

L'enseignement auquel ces notes se réfèrent fait suite au cours intitulé *Mathématique*, figurant au programme du premier quadrimestre du premier bloc des bacheliers de la Faculté des Sciences (tous les bacheliers sauf celui en maths), du premier bloc de bachelier en sciences informatiques (Faculté des Sciences Appliquées) et du premier bloc de bachelier en Philosophie (option sciences, Faculté de Philosophie et Lettres).

Au second quadrimestre, le cours intitulé *Mathématique, partim B*, est enseigné aux mêmes sections (sauf philosophie et sauf géographie), mais la charge horaire et les crédits sont différents selon les filières. Les matières suivantes sont traitées, avec approfondissement en fonction du nombre d'heures au programme, des ECTS et de la perspective de la préparation aux cours de 2ème bloc.

- Calcul matriciel (et algèbre linéaire)
- Fonctions de plusieurs variables (représentation, dérivées partielles, intégration, ...)
- Approximations polynomiales et séries

Peu d'exercices sont proposés ici. Des listes sont disponibles via les pages web relatives au cours (listes de l'année en cours et des années précédentes, via les liens « archives »).

Cependant, on peut noter que deux fascicules leur sont consacrés : ils reprennent l'ensemble de la matière de l'année et leur structure n'a pas été tout à fait remise à jour lors de l'introduction du nouveau programme « tronc commun » du premier quadrimestre et la suite du second quadrimestre. Cependant, la matière qui y est présentée reste celle du nouveau programme ; ce sont les agencements de diverses parties de matière qui doivent simplement être réorganisées et présentées dans un autre cadre ; le fond reste le même.

Avertissement et liens utiles

- Il est conseillé de consulter les pages relatives à ce cours via l'adresse [http ://www.afo.ulg.ac.be/fb/](http://www.afo.ulg.ac.be/fb/) ; de nombreux documents y sont présents et il y a une mise à jour régulière.
 - Pour des compléments d'information, des exemples, des exercices supplémentaires, je conseille l'ouvrage - *CALCULUS, with analytic geometry*, R. Ellis, D. Gulick, Harcourt Brace Jovanovich Inc. 1993 et, pour les curieux, le fantastique - *WHAT IS MATHEMATICS ? An elementary approach to ideas and methods*, Richard Courant and Herbert Robbins, Oxford University Press (plusieurs éditions).
- Ces livres se trouvent notamment à la bibliothèque des sciences.□

Je remercie sincèrement tous ceux qui aident à l'encadrement des étudiants. Sans l'équipe, rien n'est possible ! Merci à vous de toujours répondre « présent » !

Je tiens à exprimer un merci particulier à Jacqueline Crasborn pour sa présence depuis quinze ans, pour ses relectures, ses avis et suggestions pertinents, et pour ses enthousiasme et dynamisme permanents, facteurs si importants au sein d'une équipe.

Je remercie infiniment ceux qui me sont les plus proches, le terme « proche » prenant une signification particulière pour chacun. Par vos perspicacité, finesse, honnêteté, intelligence et expérience des choses, offertes de façon si totale, spontanée, franche et subtile à la fois, par votre amitié si généreuse, vous me poussez constamment à aller de l'avant et à toujours tenter de construire et rassembler.

Françoise Bastin, 26 décembre 2015

Table des matières

1	Calcul matriciel	1
1.1	Matrices	1
1.1.1	Matrices : définitions générales et notations	1
1.1.2	Vecteurs rangées, vecteurs	3
1.1.3	Matrices associées	3
1.2	Opérations entre matrices	5
1.2.1	Addition de deux matrices du même type	5
1.2.2	Multiplication d'une matrice par un nombre complexe	5
1.2.3	Propriétés des deux opérations précédentes	5
1.2.4	Produit de matrices	6
1.2.5	Propriétés du produit matriciel	7
1.3	Déterminants	8
1.3.1	Définition	8
1.3.2	Propriétés	10
1.4	Inversion de matrices	13
1.5	Trace d'une matrice carrée	17
1.6	Compléments sur les vecteurs	17
1.6.1	Cas particulier du produit matriciel	18
1.6.2	(In)-dépendance linéaire de vecteurs	18
1.7	Valeurs propres et vecteurs propres d'une matrice carrée	20
1.7.1	Introduction	20
1.7.2	Définitions et premières propriétés	20
1.7.3	Exemples	23
1.8	Diagonalisation	28
1.8.1	Définition et propriétés	28
1.8.2	Exemples	30
1.9	Matrices particulières	33
1.10	Applications	35
1.10.1	Utilisation en statistique	35
1.10.2	Processus de Markov	35
1.10.3	Modes normaux de vibration	39
1.10.4	Changement de repère orthonormé	42
1.10.5	Ecriture matricielle des transformations linéaires	42
2	Fonctions de plusieurs variables réelles	43
2.1	Définitions, représentations	43
2.1.1	Définitions et premiers exemples	43
2.1.2	Courbes de niveau	44
2.1.3	Exemples	45
2.1.4	Surfaces de niveau	47
2.1.5	Surfaces quadriques	48
2.1.6	Opérations entre fonctions	51
2.2	Limites et continuité	51

2.2.1	Limites	51
2.2.2	Continuité	55
2.3	Dérivation	56
2.3.1	Définitions des dérivées partielles et premières propriétés	56
2.3.2	Lien entre dérivabilité et continuité	60
2.3.3	Dérivées multiples	61
2.3.4	Des opérateurs de dérivation fort utiles	61
2.4	Intégration	61
2.4.1	Intégration sur des ensembles bornés fermés	61
2.4.2	Intégration sur des ensembles non bornés fermés	66
2.4.3	Intégration par changement de variables	71
2.4.4	Applications	73
2.4.5	Intégrales triples	75
2.5	Annexe	78
2.5.1	Exemple où on ne peut pas permuter les dérivées	78
3	Approximations polynomiales et séries	80
3.1	Approximations polynomiales	80
3.1.1	Définitions	80
3.1.2	Propriétés	81
3.1.3	Recherche de la forme de l'approximation.	81
3.1.4	Retour aux polynômes	83
3.1.5	Exemples des fonctions sin et cos	84
3.1.6	Estimation du reste	86
3.2	Développements illimités-Séries	88
3.2.1	Rappels sur les suites numériques	88
3.2.2	Introduction aux développements illimités	88
3.2.3	Séries (définition)	89
3.2.4	Exemples	90
3.2.5	Cas de référence	92
3.2.6	Conditions de convergence	93
3.2.7	Séries de puissances	94
3.3	Fonction exponentielle (définie par une série)	95
3.3.1	Définition	95
3.3.2	Propriétés fondamentales	95
3.3.3	Exponentielle complexe	98
3.4	Quelques exemples fondamentaux de développements en série de puissances	101
3.5	Annexe	102
3.5.1	Approximations polynomiales	102
3.5.2	Critères de convergence pour les séries	104
3.5.3	Critères pratiques de convergence des séries	105
3.5.4	Une approximation de e	105

Chapitre 1

Calcul matriciel

1.1 Matrices

1.1.1 Matrices : définitions générales et notations

Définition 1.1.1 Une matrice est un tableau rectangulaire de nombres (réels ou complexes).

Les lignes et les colonnes de ce tableau sont appelées les rangées de la matrice.

Les nombres formant le tableau sont appelés les éléments de la matrice.

La longueur des lignes de la matrice (c'est-à-dire le nombre d'éléments des lignes) est égale au nombre de colonnes de la matrice. La longueur des colonnes de la matrice (c'est-à-dire le nombre d'éléments des colonnes) est égale au nombre de lignes de la matrice.

Deux matrices sont dites égales lorsqu'elles ont le même nombre de lignes, le même nombre de colonnes et que leurs éléments correspondants sont égaux.

Si une matrice possède p lignes et q colonnes, on dit que c'est une matrice de type $p \times q$ ou de format $p \times q$. Par convention, le premier naturel indique toujours le nombre de lignes et le second le nombre de colonnes. Si une matrice possède strictement plus de lignes que de colonnes, on utilise parfois la dénomination *matrice verticale*; si elle possède strictement plus de colonnes que de lignes, on utilise parfois la dénomination *matrice horizontale*.

On désigne souvent une matrice par une lettre majuscule. Si A désigne une matrice, l'élément qui se trouve sur la ligne numéro i et la colonne numéro j est désigné par

$$(A)_{i,j}.$$

Par exemple si

$$A = \begin{pmatrix} i & -1 & 2 & \sqrt{3} \\ \frac{1}{2} & 0 & -2 & \pi \\ i+1 & 1 & 2 & -4 \\ 5 & -1 & 2i & \sqrt{2} \\ 1 & -i & 0 & -i+3 \end{pmatrix}$$

cette matrice possède 4 colonnes, 5 lignes (elle est du type 5×4); chaque ligne a une longueur égale à 4; chaque colonne a une longueur égale à 5. L'élément qui se situe sur la 3ième ligne et la 2ième colonne est

$$(A)_{3,2} = 1$$

et celui qui se situe sur la 5ième ligne et 4ième colonne est

$$(A)_{5,4} = -i + 3.$$

Pour alléger les notations (surtout dans le cas des matrices dans lesquelles la longueur des lignes ou des colonnes est grande), on utilise la même lettre minuscule pour désigner tous les éléments de la matrice, mais celle-ci est indexée par deux indices indiquant le numéro de la ligne et de la colonne sur lesquelles

se trouve l'élément. Par convention, le premier indice est le numéro de la ligne et le second celui de la colonne. Par exemple, une matrice de type 4×6 est notée, en toute généralité

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \end{pmatrix}.$$

Une matrice nulle est une matrice dont tous les éléments sont nuls. Quel que soit son format, une telle matrice est notée 0 :

$$0 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Définition 1.1.2 Une matrice carrée est une matrice dont le nombre de lignes est égal à celui des colonnes. Par définition, ce nombre est la dimension de la matrice.

Un élément diagonal d'une matrice carrée est un élément de cette matrice qui se trouve sur une ligne et une colonne de même numéro. La diagonale principale d'une matrice carrée est formée de l'ensemble des éléments diagonaux de cette matrice.

Une matrice carrée est appelée matrice diagonale si tous ses éléments non diagonaux sont nuls.

La matrice carrée diagonale de dimension n dont les éléments diagonaux sont tous égaux à 1 est appelée matrice identité¹ de dimension n . Elle est notée I_n ou tout simplement I si la dimension est claire.

Voici des exemples de matrices carrées de dimension 2, 3, 4.

$$A = \begin{pmatrix} 4i & 9 \\ -1 & i+2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -2 & i \\ i^3 & \sqrt{5} & 6 \\ \frac{3}{4} & \frac{i}{2} + 1 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 3 & i \\ 1 & 2^2 & 3^2 & i^2 \\ 1 & 2^3 & 3^3 & i^3 \\ 1 & 2^4 & 3^4 & i^4 \end{pmatrix}.$$

La diagonale principale de A est formée des éléments $4i, i+2$, celle de B est formée des éléments $0, \sqrt{5}, 0$, celle de C est formée des éléments $1, 2^2, 3^3, i^4$ c'est-à-dire $1, 4, 27, 1$.

Voici deux exemples de matrices diagonales

$$\begin{pmatrix} i & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & \sqrt{3} \end{pmatrix}, \quad \begin{pmatrix} a_1 & 0 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 & 0 \\ 0 & 0 & a_3 & 0 & 0 \\ 0 & 0 & 0 & a_4 & 0 \\ 0 & 0 & 0 & 0 & a_5 \end{pmatrix}$$

où les $a_i, i = 1, \dots, 5$, sont des nombres complexes. Pour alléger les notations, on utilise la notation suivante pour une matrice diagonale de dimension n :

$$\text{diag}(a_1, a_2, \dots, a_{n-1}, a_n) = \begin{pmatrix} a_1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & a_2 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & a_3 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & a_4 & \dots & 0 & 0 \\ \vdots & \vdots & & & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & a_{n-1} & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & a_n \end{pmatrix}.$$

1. parfois on utilise aussi le terme "matrice unité"

Les matrices identité à deux et trois dimensions sont les matrices

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

On appelle *symbole de Kronecker* l'ensemble des symboles $\delta_{i,j}$ pour des indices naturels $i, j = 1, \dots, n$, où

$$\delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

En toute généralité, la matrice identité de dimension n est donc la matrice I_n définie par

$$(I_n)_{i,j} = \delta_{i,j}, \quad i, j = 1, \dots, n.$$

1.1.2 Vecteurs rangées, vecteurs

Les rangées d'une matrice A de format $p \times q$ constituent aussi des matrices : les colonnes sont des matrices de format $p \times 1$ et les lignes sont des matrices de format $1 \times q$. On les appelle "vecteurs colonnes" de A et "vecteurs lignes" de A .

Par exemple, pour

$$A = \begin{pmatrix} i & -1 & 2 & \sqrt{3} \\ \frac{1}{2} & 0 & -2 & \pi \\ i+1 & 1 & 2 & -4 \\ 5 & -1 & 2i & \sqrt{2} \\ 1 & -i & 0 & -i+3 \end{pmatrix}$$

les vecteurs lignes sont les matrices

$$\begin{pmatrix} i & -1 & 2 & \sqrt{3} \end{pmatrix} \\ \begin{pmatrix} \frac{1}{2} & 0 & -2 & \pi \end{pmatrix} \\ \begin{pmatrix} i+1 & 1 & 2 & -4 \end{pmatrix} \\ \begin{pmatrix} 5 & -1 & 2i & \sqrt{2} \end{pmatrix} \\ \begin{pmatrix} 1 & -i & 0 & -i+3 \end{pmatrix}$$

et les vecteurs colonnes sont les matrices

$$\begin{pmatrix} i \\ \frac{1}{2} \\ i+1 \\ 5 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 0 \\ 1 \\ -1 \\ -i \end{pmatrix}, \quad \begin{pmatrix} 2 \\ -2 \\ 2 \\ 2i \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \sqrt{3} \\ \pi \\ -4 \\ \sqrt{2} \\ -i+3 \end{pmatrix}.$$

Plus généralement, on utilise la définition suivante.

Définition 1.1.3 Une matrice qui possède n lignes et une colonne ou une ligne et n colonnes est appelée un vecteur de dimension n .

Si on veut préciser s'il s'agit d'une matrice horizontale ou verticale, on utilise respectivement le terme "vecteur ligne" et "vecteur colonne". En général le contexte dans lequel on utilisera ces vecteurs sera suffisamment clair pour que le type en soit bien déterminé. Par exemple, plus loin, dans la diagonalisation, il s'agira de vecteurs colonnes.

1.1.3 Matrices associées

Définition 1.1.4 Etant donné une matrice A de type $p \times q$, on lui associe trois autres matrices :

– la

matrice conjuguée de A ,

notée

$$\bar{A};$$

il s'agit de la matrice de même type que celui de A dont les éléments sont les conjugués de ceux de A :

$$(\bar{A})_{i,j} = \overline{(A)_{i,j}}, \quad i = 1, \dots, p, \quad j = 1, \dots, q$$

– la

matrice transposée de A ,

notée

$$\tilde{A};$$

il s'agit de la matrice de type $q \times p$ dont les lignes sont formées des colonnes de A ; les colonnes de cette matrice sont alors les lignes de A .

On peut aussi définir la matrice transposée en disant que ses colonnes sont les lignes de A ; dans ce cas, ses lignes sont les colonnes de A .

Par définition :

$$(\tilde{A})_{i,j} = (A)_{j,i}, \quad i = 1, \dots, q, \quad j = 1, \dots, p$$

– la

matrice adjointe de A ,

notée

$$A^*;$$

il s'agit de la matrice transposée et conjuguée de A (ou, ce qui revient au même, de la matrice conjuguée transposée de A), c'est-à-dire de la matrice de type $q \times p$ dont les lignes (resp. les colonnes) sont formées des colonnes (resp. des lignes) de la matrice A dans lesquelles on conjugue les éléments.

Par définition

$$A^* = \tilde{\bar{A}} = \overline{\tilde{A}}, \quad (A^*)_{i,j} = \overline{(\tilde{A})_{j,i}}, \quad i = 1, \dots, q, \quad j = 1, \dots, p$$

Par exemple, pour la matrice de type 4×3 suivante

$$A = \begin{pmatrix} 0 & -2 & i \\ -i & \sqrt{5} & 6 \\ \frac{3}{4} & \frac{i}{2} + 1 & 0 \\ -1 & 3 & 4 \end{pmatrix}$$

on a

$$\bar{A} = \begin{pmatrix} 0 & -2 & -i \\ i & \sqrt{5} & 6 \\ \frac{3}{4} & \frac{-i}{2} + 1 & 0 \\ -1 & 3 & 4 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 0 & -i & \frac{3}{4} & -1 \\ -2 & \sqrt{5} & \frac{i}{2} + 1 & 3 \\ i & 6 & 0 & 4 \end{pmatrix},$$

et

$$A^* = \begin{pmatrix} 0 & i & \frac{3}{4} & -1 \\ -2 & \sqrt{5} & \frac{-i}{2} + 1 & 3 \\ -i & 6 & 0 & 4 \end{pmatrix}$$

Pour la matrice identité, on a

$$\bar{I} = \tilde{I} = I^* = I.$$

En effet, les éléments de cette matrice sont tous réels (donc $\bar{I} = I$) et sa ligne numéro k est formée des mêmes éléments, dans le même ordre, que sa colonne numéro k (donc $\tilde{I} = I$).

1.2 Opérations entre matrices

1.2.1 Addition de deux matrices du même type

Définition 1.2.1 *Etant donné deux matrices A, B de format $p \times q$, on définit la somme de ces deux matrices, notée $A + B$, comme étant la matrice de format $p \times q$ dont les éléments sont les sommes des éléments correspondants de chacune des deux matrices. Ainsi, par définition :*

$$(A + B)_{i,j} = (A)_{i,j} + (B)_{i,j}, \quad i = 1, \dots, p, \quad j = 1, \dots, q.$$

Par exemple, pour

$$A = \begin{pmatrix} 1 & 2 & 0 & i \\ 0 & -1 & 1+i & \frac{1}{2} \\ 2 & 6 & -3 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -1 & -i & -i \\ 1 & -1 & 2 & i \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

on a

$$A + B = \begin{pmatrix} 1 & 1 & -i & 0 \\ 1 & -2 & 3+i & \frac{1}{2}+i \\ 3 & 7 & -2 & 1 \end{pmatrix}.$$

1.2.2 Multiplication d'une matrice par un nombre complexe

Définition 1.2.2 *Etant donné une matrice A de format $p \times q$ et un complexe c , on définit le produit de A par c , noté cA , comme étant la matrice de format $p \times q$ dont les éléments sont égaux à c fois les éléments correspondants de A . Ainsi, par définition :*

$$(cA)_{i,j} = c(A)_{i,j}, \quad i = 1, \dots, p, \quad j = 1, \dots, q.$$

Par exemple, pour

$$A = \begin{pmatrix} 1 & 2 & 0 & i \\ 0 & -1 & 1+i & \frac{1}{2} \\ 2 & 6 & -3 & 0 \end{pmatrix}, \quad c = i$$

on a

$$cA = \begin{pmatrix} i & 2i & 0 & -1 \\ 0 & -i & -1+i & \frac{i}{2} \\ 2i & 6i & -3i & 0 \end{pmatrix}.$$

1.2.3 Propriétés des deux opérations précédentes

On vérifie directement (c'est un simple calcul, basé sur les définitions précédentes et sur les propriétés de la somme et de la multiplication entre complexes) que les deux opérations précédentes vérifient les propriétés suivantes. L'importance de ces propriétés réside dans le fait que ce sont celles que doivent vérifier deux lois définies sur un ensemble de "choses" (addition de deux "choses" du même type et multiplication d'une "chose" par un nombre) pour que cet ensemble constitue un *espace vectoriel*. Nous avons déjà mis ces propriétés en évidence au cours de l'étude du calcul vectoriel.

Propriété 1.2.3 Propriétés relatives à l'addition.

Pour toutes matrices A, B, C de même format, on a

- $(A + B) + C = A + (B + C)$ (associativité de l'addition)
- $0 + A = A + 0 = A$ (la matrice nulle est un neutre pour l'addition)
- $A + A' = 0$ où A' est la matrice de même format que A et dont les éléments sont les opposés des éléments de A (pour toute matrice A , existence d'un symétrique)
- $A + B = B + A$ (commutativité de l'addition)

Propriétés liant les deux opérations.

Pour toutes matrices A, B de même format et pour tous complexes c, c' on a

- $1A = A$
- $c(c'A) = (cc')A$
- $c(A + B) = cA + cB$
- $(c + c')A = cA + c'A$

1.2.4 Produit de matrices

La définition du produit de deux matrices peut paraître artificielle. Il n'en est rien. Cette définition provient de l'étude de la représentation matricielle des opérateurs linéaires et plus précisément de la représentation matricielle de la composée de deux opérateurs linéaires.

Définition 1.2.4 *Soit*

A une matrice de format $p \times r$

et soit

B une matrice de format $r \times q$.

Le produit des matrices A et B, dans l'ordre,

est la matrice notée AB,

de format

$$p \times q$$

dont les éléments sont obtenus en faisant les produits "ligne par colonne" des matrices A et B (dans l'ordre) c'est-à-dire

$$(AB)_{i,j} = \sum_{k=1}^r (A)_{i,k} (B)_{k,j} \quad i = 1, \dots, p, \quad j = 1, \dots, q.$$

Vu cette définition, on peut donc toujours multiplier deux matrices carrées de même dimension.

Présentons quelques exemples de produits de deux matrices.

Soient

$$A = \begin{pmatrix} 1 & i & -1 \\ -2 & \frac{i}{2} & -i \end{pmatrix}, \quad B = \begin{pmatrix} -1 & i & -1 & 2 \\ \frac{1}{4} & -i & 0 & 1 \\ 0 & 2 & -1 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} i & 2 & 3 \\ 0 & 1 & 1+i \\ -1 & 1 & 2 \end{pmatrix}.$$

On a

$$\begin{aligned} AB &= \begin{pmatrix} -1 + \frac{i}{4} & i + 1 - 2 & -1 + 1 & 2 + i - 3 \\ 2 + \frac{i}{8} & -2i + \frac{1}{2} - 2i & 2 + i & -4 + \frac{i}{2} - 3i \end{pmatrix} \\ &= \begin{pmatrix} -1 + \frac{i}{4} & -1 + i & 0 & -1 + i \\ 2 + \frac{i}{8} & \frac{1}{2} - 4i & 2 + i & -4 - \frac{5}{2}i \end{pmatrix}, \end{aligned}$$

$$C^2 = \begin{pmatrix} -1 - 3 & 2i + 2 + 3 & 3i + 2 + 2i + 6 \\ -1 - i & 1 + 1 + i & 1 + i + 2 + 2i \\ -i - 2 & -2 + 1 + 2 & -3 + 1 + i + 4 \end{pmatrix} = \begin{pmatrix} -4 & 5 + 2i & 8 + 5i \\ -1 - i & 2 + i & 3 + 3i \\ -2 - i & 1 & 2 + i \end{pmatrix},$$

$$AC = \begin{pmatrix} i + 1 & 2 + i - 1 & 3 + i - 1 - 2 \\ -2i + i & -4 + \frac{i}{2} - i & -6 + \frac{i}{2} - \frac{1}{2} - 2i \end{pmatrix} = \begin{pmatrix} i + 1 & 1 + i & i \\ -i & -4 - \frac{1}{2}i & -\frac{13}{2} - \frac{3}{2}i \end{pmatrix},$$

et

$$\begin{aligned} CB &= \begin{pmatrix} -i + \frac{1}{2} & -1 - 2i + 6 & -i - 3 & 2i + 2 + 9 \\ \frac{1}{4} & -i + 2 + 2i & -1 - i & 1 + 3 + 3i \\ 1 + \frac{1}{4} & -i - i + 4 & 1 - 2 & -2 + 1 + 6 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} - i & 5 - 2i & -3 - i & 11 + 2i \\ \frac{1}{4} & 2 + i & -1 - i & 4 + 3i \\ \frac{5}{4} & 4 - 2i & -1 & 5 \end{pmatrix}. \end{aligned}$$

1.2.5 Propriétés du produit matriciel

Deux propriétés importantes du produit matriciel sont les suivantes.

Propriété 1.2.5 1) *Le produit matriciel est associatif.*

Soient A de format $n \times p$, B de format $p \times q$ et C de format $q \times r$. On a

$$(AB)C = A(BC).$$

2) *Le produit matriciel n'est PAS commutatif.*

Cela signifie que, même si les produits sont définis, on peut avoir $AB \neq BA$.

Preuve. 1) L'élément d'indices i, j du premier membre vaut

$$\sum_{k=1}^q (AB)_{i,k} (C)_{k,j} = \sum_{k=1}^q \left(\sum_{l=1}^p (A)_{i,l} (B)_{l,k} \right) (C)_{k,j} = \sum_{l=1}^p \sum_{k=1}^q (A)_{i,l} (B)_{l,k} (C)_{k,j}$$

et celui du second membre est

$$\sum_{l=1}^p (A)_{i,l} (BC)_{l,j} = \sum_{l=1}^p (A)_{i,l} \sum_{k=1}^q (B)_{l,k} (C)_{k,j} = \sum_{l=1}^p \sum_{k=1}^q (A)_{i,l} (B)_{l,k} (C)_{k,j}.$$

2) Pour montrer que le produit matriciel n'est pas commutatif, il suffit de donner un exemple :

$$\begin{pmatrix} \cdot & \cdot \\ 1 & \cdot \end{pmatrix} \begin{pmatrix} \cdot & 1 \\ \cdot & \cdot \end{pmatrix} \neq \begin{pmatrix} \cdot & 1 \\ \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ 1 & \cdot \end{pmatrix}.$$

□

L'associativité du produit matriciel permet de définir le produit d'un nombre fini de matrices dans un ordre donné.

En particulier, on peut définir les puissances entières d'une matrice carrée

$$A^m = \underbrace{A \dots A}_m, \quad m \in \mathbb{N}_0.$$

Par convention, $A^0 = I$. Plus généralement, si

$$p(z) = a_0 + a_1 z + \dots + a_m z^m$$

est un polynôme à coefficients complexes, on pose

$$p(A) = a_0 I + a_1 A + \dots + a_m A^m.$$

Voici quelques autres propriétés du produit matriciel.

1) *Si les produits sont définis, on a*

$$A0 = 0A = 0, \quad AI = IA = A.$$

Remarquons que, dans ces formules, 0 et I ne désignent pas les mêmes matrices à gauche et à droite de A si A n'est pas carré.

2) Si $c \in \mathbb{C}$, A est du type $n \times p$ et B du type $p \times q$, on a

$$c(AB) = (cA)B = A(cB).$$

3) Le produit matriciel est *distributif par rapport aux combinaisons linéaires de matrices*. Cela signifie que l'on a

$$(cA + c'B)C = cAC + c'BC \quad \text{et} \quad C(cA + c'B) = cCA + c'CB$$

si les produits matriciels ont un sens et pour tous $c, c' \in \mathbb{C}$.

4) Si A est du type $n \times p$ et si B est du type $p \times q$ on a

$$\widetilde{AB} = \widetilde{B}\widetilde{A}, \quad \overline{AB} = \overline{A}\overline{B}, \quad (AB)^* = B^*A^*.$$

De fait, on a par exemple, pour $i = 1, \dots, n, j = 1, \dots, q$:

$$(\widetilde{AB})_{i,j} = (AB)_{j,i} = \sum_{k=1}^p (A)_{j,k} (B)_{k,i} = \sum_{k=1}^p (\widetilde{B})_{i,k} (\widetilde{A})_{k,j} = (\widetilde{B}\widetilde{A})_{i,j}.$$

5) On a vu que le produit matriciel n'est pas une opération commutative : même lorsque AB et BA sont définis, il se peut que $AB \neq BA$.

Si AB et BA sont définis et A est du type $n \times p$, alors B est nécessairement du type $p \times n$. Pour que AB et BA soient de même type, il faut donc $n = p$. Ainsi seules des matrices carrées de même dimension peuvent vérifier la relation $AB = BA$. Lorsque c'est le cas, on dit que A et B *commutent*. Voici quelques exemples de matrices qui commutent.

- Toute matrice commute avec 0 et I .
- Toute matrice commute avec elle-même.
- Des polynômes d'une même matrice commutent.
- Le produit de deux matrices diagonales est une matrice diagonale dont les éléments diagonaux sont les produits des éléments diagonaux correspondants des facteurs ; ce produit est commutatif,

$$\text{diag}(a_1, \dots, a_n) \text{diag}(b_1, \dots, b_n) = \text{diag}(a_1 b_1, \dots, a_n b_n).$$

6) La propriété qui suit est tout à fait différente de son analogue entre nombres complexes : *le produit de deux matrices peut être nul sans qu'aucun facteur ne soit nul*.

Par exemple

$$\begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}^2 = 0, \quad \begin{pmatrix} \cdot & 1 & -1 \\ -1 & \cdot & 1 \\ 1 & -1 & \cdot \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = 0.$$

1.3 Déterminants

Etant donné une matrice carrée, on lui associe un nombre complexe, appelé *déterminant de la matrice*.

La définition du déterminant d'une matrice carrée est donnée par récurrence sur la dimension de la matrice. Cela signifie que l'on donne la définition du déterminant d'une matrice quelconque de dimension 2 ; ensuite on donne la définition du déterminant d'une matrice de dimension 3, laquelle est basée sur les déterminants de matrices de dimension 2, etc.

Par souci de généralité, on donne aussi la définition du déterminant d'une matrice de dimension 1.

Si A est une matrice carrée, son déterminant est noté $\det A$.

1.3.1 Définition

Si la matrice est de dimension 1, c'est-à-dire si $A = (a)$ où $a \in \mathbb{C}$, on définit le déterminant de cette matrice par

$$\boxed{\det(a) = a.}$$

Si la matrice est de dimension 2, c'est-à-dire si

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

alors

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

Si la matrice est de dimension 3, c'est-à-dire si

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

alors

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

Appelons

cofacteur

de l'élément i, j d'une matrice carrée (c'est-à-dire de l'élément qui se trouve sur la ligne numéro i et la colonne numéro j), le déterminant du tableau carré obtenu en supprimant la ligne et la colonne qui contiennent cet élément, multiplié par $(-1)^{i+j}$. Par définition, le déterminant de A est donc la somme des produits des éléments de la première ligne par les cofacteurs correspondants :

$$\begin{aligned} & \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ &= (a_{11} \times \text{cofacteur de } a_{11}) + (a_{12} \times \text{cofacteur de } a_{12}) + (a_{13} \times \text{cofacteur de } a_{13}). \end{aligned}$$

De même pour le déterminant d'une matrice de dimension deux.

Ainsi, de proche en proche sur la dimension de la matrice, on définit le déterminant pour une matrice de dimension n :

le déterminant de A est la somme des produits des éléments de la première ligne par les cofacteurs correspondants

On appelle

ordre d'un déterminant

la dimension de la matrice carrée qui sert à le définir.

Pour les déterminants, on utilise aussi souvent la notation suivante : si

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

alors

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

De même en dimension 3 : si

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

alors

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

1.3.2 Propriétés

Une première propriété fort utile dans le calcul des déterminants est la suivante. Elle se démontre directement (mais les calculs sont un peu longs).

Propriété 1.3.1 (Première loi des mineurs) *Le déterminant d'une matrice carrée est égal à la somme des produits des éléments d'une rangée (quelconque) par les cofacteurs correspondants.*

Cette propriété signifie donc que, si l'on effectue le même "développement" que celui qui a été fait dans la définition, mais en suivant une ligne quelconque ou une colonne quelconque, on trouve la même valeur. Cette propriété est très utile par exemple lorsque la matrice a des éléments nuls situés sur une même rangée.

Ainsi par exemple le déterminant de la matrice

$$A = \begin{pmatrix} i & \sqrt{3} & -\frac{1}{2} \\ 2 & i+2 & 0 \\ \frac{1}{3} & -i & 0 \end{pmatrix}$$

se calcule rapidement lorsqu'on le développe selon la troisième colonne : en effet, comme deux éléments de cette colonne sont nuls, le calcul nécessite seulement le calcul d'un déterminant d'ordre 2 :

$$\begin{aligned} \det A &= \frac{-1}{2} (-1)^{1+3} \det \begin{pmatrix} 2 & i+2 \\ \frac{1}{3} & -i \end{pmatrix} \\ &= \frac{-1}{2} \left(-2i - \frac{i}{3} - \frac{2}{3} \right) \\ &= \frac{-1}{2} \left(-\frac{2}{3} - \frac{7i}{3} \right) \\ &= \frac{1}{3} + \frac{7i}{6}. \end{aligned}$$

Voici quelques autres propriétés. Nous les admettons.

Propriété 1.3.2 1. *Propriété de linéarité : si une colonne C (resp. une ligne L) d'une matrice carrée A est une combinaison linéaire (c'est-à-dire une somme de multiples) de vecteurs colonnes (resp. vecteurs lignes) alors le déterminant de A est égal à la somme des multiples des déterminants des matrices obtenues en remplaçant C (resp. L) par les vecteurs colonnes (resp. lignes) intervenant dans la combinaison linéaire.*

2. *Le déterminant d'une matrice change de signe lorsqu'on permute deux rangées parallèles dans la matrice.*

En conséquence, le déterminant d'une matrice qui a deux rangées parallèles égales est nul.

3. *Une conséquence des deux propriétés précédentes est la suivante : si, à une rangée d'une matrice, on ajoute une combinaison linéaire des autres rangées parallèles, on définit une nouvelle matrice qui a le même déterminant que celui de la matrice de départ.*

Illustrons ces propriétés par des exemples.

Si

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

la première colonne de A est égale à

$$s \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + r \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}$$

donc, vu la première propriété, on a

$$\det A = s \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + r \det \begin{pmatrix} b_{11} & a_{12} \\ b_{21} & a_{22} \end{pmatrix}.$$

Un cas très utile est celui où on ne fait que multiplier une rangée par un nombre :

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

On a aussi

$$\det(rA) = \det \begin{pmatrix} ra_{11} & ra_{12} & ra_{13} \\ ra_{21} & ra_{22} & ra_{23} \\ ra_{31} & ra_{32} & ra_{33} \end{pmatrix} = r^3 \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

et, si la matrice est de dimension deux

$$\det(rA) = r^2 \det A.$$

Par la deuxième propriété, on a

$$\begin{aligned}
 \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= -\det \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\
 &= \det \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \end{pmatrix} \\
 &= -\det \begin{pmatrix} a_{13} & a_{12} & a_{11} \\ a_{23} & a_{22} & a_{21} \\ a_{33} & a_{32} & a_{31} \end{pmatrix} \\
 &\text{etc}
 \end{aligned}$$

La troisième propriété permet de simplifier grandement le calcul de déterminants ; elle est spécialement utile lorsqu'on demande une factorisation. Ainsi par exemple

$$\begin{aligned}
 \det \begin{pmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{pmatrix} &= \det \begin{pmatrix} 1 & a & a^2 \\ 0 & b-a & b^2-a^2 \\ 0 & c-a & c^2-a^2 \end{pmatrix} \\
 &= \det \begin{pmatrix} b-a & b^2-a^2 \\ c-a & c^2-a^2 \end{pmatrix} \\
 &= \det \begin{pmatrix} b-a & (b-a)(b+a) \\ c-a & (c-a)(c+a) \end{pmatrix} \\
 &= (b-a)(c-a) \det \begin{pmatrix} 1 & b+a \\ 1 & c+a \end{pmatrix} \\
 &= (b-a)(c-a) [(c+a) - (b+a)] \\
 &= (b-a)(c-a)(c-b) \\
 &= (a-b)(b-c)(c-a).
 \end{aligned}$$

La propriété qui suit, jointe à la première loi des mineurs, permettra d'inverser les matrices.

Propriété 1.3.3 (Seconde loi des mineurs) *La somme des produits des éléments d'une rangée d'une matrice carrée par les cofacteurs des éléments correspondants d'une rangée parallèle est nulle.*

□

Illustrons cette propriété sur deux exemples. Soit la matrice

$$A = \begin{pmatrix} 1 & i & -2 \\ i+2 & \frac{1}{2} & 1 \\ 3 & 1 & \frac{i}{2} \end{pmatrix}.$$

Calculons la somme des produits des éléments de la troisième ligne par les cofacteurs de la deuxième ligne. De même, calculons la somme des produits des éléments de la deuxième colonne par les cofacteurs des éléments de la première colonne.

Les cofacteurs des éléments de la deuxième ligne sont

$$\begin{aligned} \text{cofacteur de } a_{21} &= (-1)^{2+1} \begin{vmatrix} i & -2 \\ 1 & \frac{i}{2} \end{vmatrix} = -\left(-\frac{1}{2} + 2\right) = -\frac{3}{2} \\ \text{cofacteur de } a_{22} &= (-1)^{2+2} \begin{vmatrix} 1 & -2 \\ 3 & \frac{i}{2} \end{vmatrix} = 6 + \frac{i}{2} \\ \text{cofacteur de } a_{23} &= (-1)^{2+3} \begin{vmatrix} 1 & i \\ 3 & 1 \end{vmatrix} = -(1 - 3i) = -1 + 3i. \end{aligned}$$

La somme des produits des éléments de la troisième ligne par les cofacteurs correspondants de la deuxième ligne est donc

$$\begin{aligned} &a_{31} \times \text{cofacteur de } a_{21} + a_{32} \times \text{cofacteur de } a_{22} + a_{33} \times \text{cofacteur de } a_{23} \\ &= 3 \times \left(-\frac{3}{2}\right) + 1 \times \left(6 + \frac{i}{2}\right) + \frac{i}{2} \times (-1 + 3i); \end{aligned}$$

un simple calcul donne

$$3 \times \left(-\frac{3}{2}\right) + 1 \times \left(6 + \frac{i}{2}\right) + \frac{i}{2} \times (-1 + 3i) = -\frac{9}{2} + 6 + \frac{i}{2} - \frac{i}{2} - \frac{3}{2} = 0.$$

De même, les cofacteurs des éléments de la première colonne sont

$$\begin{aligned} \text{cofacteur de } a_{11} &= (-1)^{1+1} \begin{vmatrix} \frac{1}{2} & 1 \\ 1 & \frac{i}{2} \end{vmatrix} = \left(\frac{i}{4} - 1\right) \\ \text{cofacteur de } a_{21} &= (-1)^{2+1} \begin{vmatrix} i & -2 \\ 1 & \frac{i}{2} \end{vmatrix} = -\left(\frac{-1}{2} + 2\right) = -\frac{3}{2} \\ \text{cofacteur de } a_{31} &= (-1)^{3+1} \begin{vmatrix} i & -2 \\ \frac{1}{2} & 1 \end{vmatrix} = i + 1. \end{aligned}$$

La somme des produits des éléments de la deuxième colonne par les cofacteurs des éléments de la première colonne est donc

$$\begin{aligned} &a_{12} \times \text{cofacteur de } a_{11} + a_{22} \times \text{cofacteur de } a_{21} + a_{32} \times \text{cofacteur de } a_{31} \\ &= i \times \left(\frac{i}{4} - 1\right) + \frac{1}{2} \times \left(-\frac{3}{2}\right) + 1 \times (i + 1); \end{aligned}$$

un simple calcul donne

$$i \times \left(\frac{i}{4} - 1\right) + \frac{1}{2} \times \left(-\frac{3}{2}\right) + 1 \times (i + 1) = -\frac{1}{4} - i - \frac{3}{4} + i + 1 = 0.$$

On montre directement (par calcul) la propriété suivante concernant le déterminant d'une matrice diagonale.

Propriété 1.3.4 Si

$$A = \text{diag}(a_1, a_2, \dots, a_n)$$

alors

$$\det A = a_1 a_2 \dots a_n.$$

En particulier,

$$\det I = 1.$$

□ Ce résultat s'écrit, en dimension deux :

$$\begin{vmatrix} a_1 & 0 \\ 0 & a_2 \end{vmatrix} = a_1 a_2;$$

en dimension trois :

$$\begin{vmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{vmatrix} = a_1 a_2 a_3;$$

en dimension quatre

$$\begin{vmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & a_3 & 0 \\ 0 & 0 & 0 & a_4 \end{vmatrix} = a_1 a_2 a_3 a_4.$$

La propriété précédente peut être généralisée au cas de matrices triangulaires inférieures (resp. supérieures), c'est-à-dire au cas des matrices dont les éléments situés au-dessus (resp. en dessous) de la diagonale principale sont tous nuls. Le déterminant d'une matrice de ce type est encore égal au produit des éléments diagonaux. Pour la dimension 3 et une matrice triangulaire supérieure, ce résultat s'écrit

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix} = a_{11} a_{22} a_{33}.$$

Enfin vis-à-vis du produit de deux matrices carrées de même dimension, on a la propriété suivante.

Propriété 1.3.5 Si A, B sont deux matrices carrées de même dimension, on a

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

□

Attention, il se peut que AB soit une matrice carrée sans que A, B le soient (si A est de type $p \times q$, la matrice AB est définie et est carrée si et seulement si B est du type $q \times p$). La propriété précédente n'est plus correcte dans ce cas. On peut aussi donner une méthode de calcul du déterminant de AB par l'intermédiaire du calcul de déterminants de sous-matrices de A et de B (*formule de Binet Cauchy*). Ainsi, par exemple, on établit le résultat suivant : si la matrice A est verticale, de type $p \times q$ (donc $p > q$) et si la matrice B est du type $q \times p$, alors $\det(AB) = 0$.

1.4 Inversion de matrices

Le nombre 1 joue le rôle de neutre dans la multiplication entre complexes : on a $1z = z1 = z$ pour tout complexe z . Dans le cas des matrices carrées de même dimension, la matrice identité joue aussi le rôle de neutre pour la multiplication entre matrices : on a $AI = IA = A$ pour toute matrice carrée A .

Dans le cadre de la théorie des nombres complexes, on s'est posé la question de savoir si, étant donné un complexe z , il existe un autre complexe z' tel que

$$zz' = z'z = 1.$$

La réponse est oui si z n'est pas nul. Le complexe z' qui réalise ces égalités est de plus unique et est appelé l'inverse du complexe z .

Il est naturel de se poser la même question au sein des matrices carrées de même dimension : étant donné une matrice carrée² A , existe-t-il une matrice carrée A' de même dimension que A , telle que

$$AA' = I \text{ et } A'A = I.$$

Remarquons ici que le produit matriciel n'étant pas commutatif, le fait d'avoir $AA' = I$ n'implique pas, à priori, que $A'A = I$ comme c'est le cas pour les nombres complexes.

On adopte alors de façon naturelle la définition suivante.

Définition 1.4.1 *Soit A une matrice carrée de dimension n . On appelle matrice inverse de A une matrice carrée A' de dimension n qui vérifie les deux égalités suivantes*

$$AA' = I = A'A.$$

Si, étant donné A , il existe une telle matrice A' , on dit simplement que A admet un inverse.

Recherchons alors sous quelle(s) condition(s) une matrice admet une matrice inverse, si cette matrice inverse est unique et, si possible, quelle est sa forme explicite.

Nous obtenons immédiatement une condition nécessaire à l'existence d'une matrice inverse, comme le montre la propriété ci-dessous.

Propriété 1.4.2 *Si A admet un inverse alors son déterminant est non nul.*

Preuve. Soit une matrice A' telle que $AA' = I$. En prenant le déterminant, on obtient

$$\det(AA') = 1 = \det A \det A'$$

donc $\det A \neq 0$. $\square$

En fait, les lois des mineurs vont nous permettre de montrer que cette condition nécessaire à l'existence d'un inverse (à savoir $\det A \neq 0$) est aussi suffisante.

Définition 1.4.3 *Soit une matrice carrée A . Notons*

$$\mathcal{A}$$

la matrice dont les éléments sont les cofacteurs de la matrice A , c'est-à-dire

$$(\mathcal{A})_{i,j} = \text{cofacteur de l'élément } (A)_{i,j}, \quad i, j = 1, \dots, n.$$

En toute généralité, pour les dimensions deux et trois, on a donc

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

2. Le problème de l'existence d'un inverse (à gauche ou à droite) d'une matrice non carrée est un peu plus lourd à développer. Dans un premier temps, ce cours ne l'abordera pas.

et

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

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

Par exemple, si

$$A = \begin{pmatrix} 0 & i \\ 1 & -3 \end{pmatrix} \quad \text{alors} \quad \mathcal{A} = \begin{pmatrix} -3 & -1 \\ -i & 0 \end{pmatrix};$$

si

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix} \quad \text{alors} \quad \mathcal{A} = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \\ 2 & -1 & -1 \end{pmatrix}.$$

Cela étant, énonçons la propriété qui n'est en fait rien d'autre que l'expression conjointe des deux lois des mineurs.

Propriété 1.4.4 *Soit une matrice carrée A . On a*

$$A \tilde{\mathcal{A}} = (\det A) I \quad \text{et} \quad \tilde{\mathcal{A}} A = (\det A) I.$$

Cela s'écrit, à deux dimensions

$$A \tilde{\mathcal{A}} = (\det A) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \det A & 0 \\ 0 & \det A \end{pmatrix} = \tilde{\mathcal{A}} A,$$

et à trois dimensions

$$A \tilde{\mathcal{A}} = (\det A) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \det A & 0 & 0 \\ 0 & \det A & 0 \\ 0 & 0 & \det A \end{pmatrix} = \tilde{\mathcal{A}} A.$$

Preuve. Examinons le cas du produit $A \tilde{\mathcal{A}}$. L'élément de la ligne numéro i et de la colonne numéro j de ce produit est, par définition du produit matriciel, la somme des produits des éléments de la ligne numéro i de A avec ceux de la colonne numéro j de $\tilde{\mathcal{A}}$, c'est-à-dire la somme des produits des éléments de la ligne numéro i de A avec ceux de la ligne numéro j de $\mathcal{A}$. Lorsque $i = j$, il s'agit donc de la somme des produits des éléments d'une ligne de A par les cofacteurs des éléments de cette ligne. Vu la première loi des mineurs, cette somme vaut $\det A$. Lorsque $i \neq j$, il s'agit donc de la somme des produits des éléments d'une ligne de A par les cofacteurs des éléments correspondants d'une autre ligne. Par la seconde loi des mineurs, cette somme est nulle. On a donc obtenu

$$(A \tilde{\mathcal{A}})_{i,j} = \begin{cases} 0 & \text{si } i \neq j \\ \det A & \text{si } i = j \end{cases}$$

c'est-à-dire

$$A \tilde{\mathcal{A}} = (\det A) I.$$

Le même raisonnement peut être tenu pour établir la seconde égalité. L'élément de la ligne numéro i et de la colonne numéro j du produit matriciel $\tilde{\mathcal{A}} A$ est, par définition du produit matriciel, la somme des produits des éléments de la ligne numéro i de $\tilde{\mathcal{A}}$ par ceux de la colonne numéro j de A , c'est-à-dire

la somme des produits des éléments de la colonne numéro i de $\mathcal{A}$ avec ceux de la colonne numéro j de A . Lorsque $i = j$, cette somme vaut $\det A$ par la première loi des mineurs. Lorsque $i \neq j$, cette somme est nulle en vertu de la seconde loi des mineurs. On a donc

$$\tilde{\mathcal{A}} A = (\det A) I.$$

□

Il s'ensuit le résultat ci-dessous.

Proposition 1.4.5 *Si A est une matrice carrée dont le déterminant n'est pas nul, alors la matrice*

$$A' = \frac{1}{\det A} \tilde{\mathcal{A}}$$

vérifie

$$AA' = I \quad \text{et} \quad A'A = I.$$

Montrons que cette matrice $\frac{1}{\det A} \tilde{\mathcal{A}}$ est l'unique matrice A' à avoir la propriété $A'A = I = AA'$. On a en fait un résultat beaucoup plus fort. Il est énoncé et démontré ci-dessous.

Propriété 1.4.6 *Soit A une matrice carrée.*

Si A' est une matrice carrée de même dimension que A telle que

$$A'A = I \quad (\text{resp. } AA' = I)$$

alors

$$\det A \neq 0$$

et

$$A' = \frac{1}{\det A} \tilde{\mathcal{A}}.$$

Preuve. Par propriété du déterminant du produit de deux matrices carrées, on a

$$1 = \det I = \det (A'A) = \det A' \det A \quad (\text{resp. } 1 = \det I = \det (AA') = \det A \det A')$$

donc le déterminant de A n'est pas nul.

Notons

$$B = \frac{1}{\det A} \tilde{\mathcal{A}}.$$

On a

$$\begin{aligned} A' &= A'I = A'(AB) = (A'A)B = IB = B \\ (\text{resp. } A' &= IA' = (BA)A' = B(AA') = BI = B. \end{aligned}$$

□

Vu le résultat d'unicité, on utilise alors la notation suivante : si A est une matrice carrée de déterminant non nul, on pose

$$A^{-1} = \frac{1}{\det A} \tilde{\mathcal{A}}.$$

Cette matrice est appelée

l'inverse de la matrice A .

Une matrice dont le déterminant est non nul (resp. nul) est dite *non singulière* (resp. *singulière*) ou encore *inversible* (resp. *non inversible*).

Voici quelques propriétés de l'inverse d'une matrice. Elles se démontrent directement.

Propriété 1.4.7 a) Les matrices $A, \bar{A}, \tilde{A}$ et A^* sont simultanément inversibles et on a

$$(\bar{A})^{-1} = \overline{A^{-1}}, (\tilde{A})^{-1} = \widetilde{A^{-1}}, (A^*)^{-1} = (A^{-1})^*.$$

b) Si λ est un nombre complexe non nul et A est inversible alors λA est inversible et

$$(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}.$$

c) Si A et B sont inversibles et de même dimension, alors AB est inversible et

$$(AB)^{-1} = B^{-1}A^{-1}.$$

d) Une matrice diagonale est inversible si et seulement si ses éléments diagonaux sont non nuls. De plus

$$\text{diag}(a_1, \dots, a_n)^{-1} = \text{diag}\left(\frac{1}{a_1}, \dots, \frac{1}{a_n}\right).$$

En particulier, $I^{-1} = I$.

e) L'inverse d'une matrice est toujours inversible. On a

$$\det A^{-1} = \frac{1}{\det A} \quad \text{et} \quad (A^{-1})^{-1} = A.$$

f) Si $AB = AC$ (ou $BA = CA$) et si A est inversible, alors $B = C$.

Preuve. a) Les déterminants de ces matrices sont simultanément nuls. Les inverses proposés conviennent. Par exemple, dans le cas de la matrice transposée, on a

$$\tilde{A}^{-1} \tilde{A} = \widetilde{AA^{-1}} = I.$$

c) De fait, $(B^{-1}A^{-1})AB = B^{-1}IB = I$.

e) En effet, $\det(A^{-1}A) = \det A^{-1} \det A = \det I = 1$; de plus, $AA^{-1} = I$ donne $(A^{-1})^{-1} = A$.

f) Cela revient en effet à multiplier à gauche (ou à droite) la relation de départ par A^{-1} . $\square$

1.5 Trace d'une matrice carrée

Définition 1.5.1 Soit A une matrice carrée de dimension n . La trace de A , notée $\text{tr}(A)$ est la somme des éléments de sa diagonale principale :

$$\text{tr}(A) = \sum_{i=1}^n (A)_{i,i}.$$

En voici quelques propriétés. Elles se démontrent directement en repassant à la définition.

1. $\text{tr}(\tilde{A}) = \text{tr}(A)$, $\text{tr}(\bar{A}) = \overline{\text{tr}(A)}$, $\text{tr}(A^*) = \overline{\text{tr}(A)}$.
2. $\text{tr}(\sum_{j=1}^J c_j A_j) = \sum_{j=1}^J c_j \text{tr}(A_j)$, pour tous complexes c_j ($j = 1, \dots, J$) et toutes matrices carrées A_j ($j = 1, \dots, J$) de même dimension.
3. $\text{tr}(AB) = \text{tr}(BA)$ pour toutes matrices A, B carrées de même dimension.

1.6 Compléments sur les vecteurs

Dans cette partie et jusqu'à la partie réservée à la diagonalisation, nous ne nous intéresserons qu'aux matrices carrées et aux vecteurs de dimension deux et trois. Une étude analogue pourrait être effectuée dans le cas des matrices carrées et des vecteurs de dimension quelconque.

1.6.1 Cas particulier du produit matriciel

Soit A une matrice carrée de dimension 2 (resp. 3) et soit X un vecteur colonne de dimension 2 (resp. 3). Si

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

et si on note

$$C_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad C_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$

respectivement les première et seconde colonnes de A , on a

$$A = (C_1, C_2), \quad AX = xC_1 + yC_2 = \begin{pmatrix} a_{11}x + a_{12}y \\ a_{21}x + a_{22}y \end{pmatrix}.$$

De même si

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

et si on note

$$C_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \quad C_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \quad C_3 = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}$$

respectivement les première, deuxième et troisième colonnes de A , on a

$$A = (C_1, C_2, C_3), \quad AX = xC_1 + yC_2 + zC_3 = \begin{pmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{pmatrix}.$$

1.6.2 (In)-dépendance linéaire de vecteurs

Définition 1.6.1 On dit que 2 vecteurs de dimension deux (resp. trois) sont linéairement dépendants si l'un est un multiple de l'autre.

On dit que 3 vecteurs de dimension deux (resp. trois) sont linéairement dépendants si l'un est une somme de multiples des deux autres.

En toute généralité, on dit que J vecteurs (J étant un naturel au moins égal à 2) sont linéairement dépendants si l'un d'entre eux est une somme de multiples des autres.

Des vecteurs qui ne sont pas linéairement dépendants sont appelés vecteurs linéairement indépendants.

En guise d'interprétation, on peut considérer que les vecteurs dont il est question ici (c'est-à-dire des matrices de type particulier) sont les tableaux des composantes de vecteurs libres (ou liés) dans une base fixée du plan (pour les vecteurs de dimension deux) ou de l'espace (pour les vecteurs de dimension trois).

On a les propriétés suivantes. Elles sont aisées à démontrer mais nous les accepterons sans preuve en première lecture. Elles se généralisent à des vecteurs de plus grande dimension.

L'interprétation à l'aide des vecteurs libres ou liés est utile pour une aide à la compréhension des notions de dépendance et d'indépendance linéaire, de même qu'aux propriétés générales liées à ces notions.

Propriété 1.6.2 1) Trois vecteurs de dimension deux sont toujours linéairement dépendants.

Quatre vecteurs de dimension trois sont toujours linéairement dépendants.

2) Si deux vecteurs sont multiples d'un autre vecteur, ces deux vecteurs sont linéairement dépendants.

Si chacun des trois vecteurs X_1, X_2, X_3 est une somme de multiples de deux vecteurs Y_1, Y_2 (ces deux vecteurs étant les mêmes pour X_1, X_2, X_3), alors les vecteurs X_1, X_2, X_3 sont linéairement dépendants.

□

Insistons sur le fait que les propriétés ci-dessus signifient géométriquement que

- 1) étant donné trois vecteurs du plan (resp. quatre vecteurs de l'espace), il y en a toujours un qui est une somme de multiples des deux (resp. trois) autres,
- 2) deux vecteurs parallèles à une même droite sont parallèles entre eux (resp. si trois vecteurs de l'espace sont dans un même plan, alors il y a un vecteur qui est une somme de multiples des deux autres).

En ce qui concerne la relation entre la dépendance linéaire de vecteurs et l'annulation de déterminants, on a notamment le résultat suivant.

Théorème 1.6.3 *Le déterminant d'une matrice carrée est nul si et seulement si les colonnes (resp. les lignes) de cette matrice sont des vecteurs linéairement dépendants.*

Preuve. Le résultat affirmant que si les vecteurs colonnes sont linéairement dépendants alors le déterminant est nul est très aisé à démontrer en utilisant les propriétés relatives au déterminant et énoncées précédemment. En effet, supposons par exemple que A soit une matrice de dimension 3 telle que $C_3 = cC_1 + dC_2$ où C_1, C_2, C_3 sont les vecteurs colonnes de A et où c, d sont des complexes. On a donc, en utilisant la propriété de linéarité du déterminant

$$\begin{aligned}\det A = \det (C_1, C_2, C_3) &= \det (C_1, C_2, cC_1 + dC_2) \\ &= c \det (C_1, C_2, C_1) + d \det (C_1, C_2, C_2).\end{aligned}$$

Les deux termes du membre de droite sont nuls car il s'agit de déterminants de matrices ayant deux rangées parallèles égales.

La démonstration du résultat disant que l'annulation du déterminant entraîne la dépendance linéaire des rangées n'est pas difficile mais un peu longue pour être présentée ici. □

Illustrons ce théorème.

D'une part (illustration du sens $\Leftarrow$), si

$$A = \begin{pmatrix} a & ra \\ b & rb \end{pmatrix}$$

alors

$$\det A = a \times (rb) - b \times (ra) = rab - rab = 0.$$

D'autre part (illustration du sens $\Rightarrow$), avec

$$A = \begin{pmatrix} 1 & 2 & -\frac{5}{6} \\ 2 & 1 & \frac{1}{3} \\ -3 & 1 & -\frac{13}{6} \end{pmatrix}$$

on a

$$\begin{aligned}\det A &= -2 \begin{vmatrix} 2 & \frac{1}{3} \\ -3 & -\frac{13}{6} \end{vmatrix} + \begin{vmatrix} 1 & -\frac{5}{6} \\ -3 & -\frac{13}{6} \end{vmatrix} - \begin{vmatrix} 1 & -\frac{5}{6} \\ 2 & \frac{1}{3} \end{vmatrix} \\ &= -2 \left(-\frac{13}{3} + 1 \right) + \left(-\frac{13}{6} - \frac{5}{2} \right) - \left(\frac{1}{3} + \frac{5}{3} \right) \\ &= \frac{20}{3} - \frac{28}{6} - \frac{6}{3} \\ &= \frac{20 - 14 - 6}{3} \\ &= 0.\end{aligned}$$

Vu le théorème précédent, cela implique qu'une colonne est une somme de multiples des deux autres et qu'une ligne est une somme de multiples des deux autres. Après quelques calculs, on trouve en effet

explicitement

$$C_3 = \begin{pmatrix} -\frac{5}{6} \\ \frac{1}{3} \\ -\frac{13}{6} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} - \frac{2}{3} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{2}C_1 - \frac{2}{3}C_2$$

et

$$\begin{aligned} L_3 &= \begin{pmatrix} -3 & 1 & -\frac{13}{6} \end{pmatrix} \\ &= \frac{5}{3} \begin{pmatrix} 1 & 2 & -\frac{5}{6} \end{pmatrix} - \frac{7}{3} \begin{pmatrix} 2 & 1 & \frac{1}{3} \end{pmatrix} \\ &= \frac{5}{3}L_1 - \frac{7}{3}L_2. \end{aligned}$$

1.7 Valeurs propres et vecteurs propres d'une matrice carrée

1.7.1 Introduction

Dans de nombreuses situations (axes principaux d'inertie en mécanique du solide, matrices de dispersion en statistique, réduction de systèmes d'équations différentielles pour les molécules vibrantes, ...), on rencontre le problème suivant : étant donné une matrice carrée A , déterminer les complexes λ et les vecteurs non nuls X tels que

$$AX = \lambda X.$$

1.7.2 Définitions et premières propriétés

Définition 1.7.1 *Un complexe λ pour lequel il existe un vecteur non nul X tel que $AX = \lambda X$ s'appelle une valeur propre de A .*

L'ensemble des valeurs propres de A est appelé le spectre de A .

Un vecteur non nul X vérifiant $AX = \lambda X$ est appelé

vecteur propre de A de valeur propre λ .

Proposition 1.7.2 *1) Les valeurs propres de A sont les zéros du polynôme*

$$P(\lambda) = \det(A - \lambda I).$$

Ce polynôme est appelé

polynôme caractéristique de A

et l'équation $P(\lambda) = 0$ est appelée équation caractéristique de A .

2) Une matrice de dimension 2 (resp. 3) a exactement deux (resp. trois) valeurs propres, si on les compte avec leur multiplicité.

3) Le produit (resp. la somme) des valeurs propres d'une matrice (en les répétant selon leur multiplicité) est égal au déterminant (resp. à la trace) de la matrice.

Preuve. 1) Si λ_0 est une valeur propre de la matrice A , il existe un vecteur colonne non nul X_0 tel que

$$AX_0 = \lambda_0 X_0$$

ou encore

$$(A - \lambda_0 I)X_0 = 0.$$

Si le déterminant de $A - \lambda_0 I$ était non nul, on aurait, en appliquant l'inverse de la matrice $A - \lambda_0 I$ aux deux membres de l'égalité :

$$0 = (A - \lambda_0 I)^{-1}(A - \lambda_0 I)X_0 = X_0$$

ce qui est contradictoire. Dès lors, le complexe λ_0 est bien un zéro du polynôme $P(\lambda) = \det(A - \lambda I)$.

Démontrons à présent que si le complexe λ_0 vérifie $\det(A - \lambda_0 I) = 0$ alors c'est une valeur propre de A .

Supposons A de dimension 2. Comme $\det(A - \lambda_0 I) = 0$, les colonnes de $A - \lambda_0 I$, notées C_1, C_2 , sont des vecteurs linéairement dépendants. Il existe donc un nombre c tel que, par exemple, $C_1 = cC_2$ ou encore

$$C_1 - cC_2 = 0.$$

Comme

$$C_1 - cC_2 = (C_1, C_2)X_0 = (A - \lambda_0 I)X_0$$

avec

$$X_0 = \begin{pmatrix} 1 \\ -c \end{pmatrix}$$

on a donc bien trouvé un vecteur non nul X_0 tel que

$$(A - \lambda_0 I)X_0 = 0.$$

On procède de même lorsque A est de dimension 3 : comme $\det(A - \lambda_0 I) = 0$, les colonnes de $A - \lambda_0 I$, notées C_1, C_2, C_3 , sont des vecteurs linéairement dépendants. Il existe donc des nombres c, d tels que, par exemple, $C_2 = cC_1 + dC_3$ ou encore

$$-cC_1 + C_2 - dC_3 = 0.$$

Comme

$$-cC_1 + C_2 - dC_3 = (C_1, C_2, C_3)X_0 = (A - \lambda_0 I)X_0$$

avec

$$X_0 = \begin{pmatrix} -c \\ 1 \\ -d \end{pmatrix}$$

on a donc bien trouvé un vecteur non nul X_0 tel que

$$(A - \lambda_0 I)X_0 = 0.$$

2) Cela résulte de la théorie générale des polynômes à coefficients et variable complexe (admis) qui dit que tout polynôme P_n de degré n possède exactement n zéros $z_1, \dots, z_n$, comptés avec leur multiplicité et que

$$P_n(z) = c_n \prod_{j=1}^n (z - z_j)$$

où c_n est le coefficient de z^n dans P_n .

3) Par exemple, pour une matrice de dimension trois et si les valeurs propres de A sont notées $\lambda_1, \lambda_2, \lambda_3$, on a (vu la preuve de (2))

$$\det(A - \lambda I) = (\lambda_1 - \lambda) (\lambda_2 - \lambda) (\lambda_3 - \lambda)$$

donc, pour $\lambda = 0$, on obtient

$$\det A = \lambda_1 \lambda_2 \lambda_3.$$

□

Remarquons que si A est une matrice diagonale, tout se simplifie.

Propriété 1.7.3 *Si*

$$A = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$$

alors ses valeurs propres sont

$$a_1, a_2$$

et

$$X_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

sont des vecteurs propres de valeur propre a_1, a_2 respectivement.

Si

$$A = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}$$

alors ses valeurs propres sont

$$a_1, a_2, a_3$$

et les vecteurs

$$X_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sont des vecteurs propres de valeur propre a_1, a_2, a_3 respectivement.

Preuve. En effet, à deux dimensions, on a

$$A - \lambda I = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_1 - \lambda & 0 \\ 0 & a_2 - \lambda \end{pmatrix}$$

donc

$$\det(A - \lambda I) = (a_1 - \lambda)(a_2 - \lambda).$$

De plus

$$AX_1 = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ 0 \end{pmatrix} = a_1 X_1, \quad AX_2 = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ a_2 \end{pmatrix} = a_2 X_2.$$

De même à trois dimensions, on a

$$A - \lambda I = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_1 - \lambda & 0 & 0 \\ 0 & a_2 - \lambda & 0 \\ 0 & 0 & a_3 - \lambda \end{pmatrix}$$

donc

$$\det(A - \lambda I) = (a_1 - \lambda)(a_2 - \lambda)(a_3 - \lambda).$$

De plus

$$AX_1 = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix} = a_1 X_1$$

$$AX_2 = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ a_2 \\ 0 \end{pmatrix} = a_2 X_2$$

$$AX_3 = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ a_3 \end{pmatrix} = a_3 X_3.$$

□

Terminons cette partie par des propriétés des vecteurs propres relatifs à une même valeur propre, puis à des valeurs propres différentes.

Propriété 1.7.4 Si les vecteurs $X_1, X_2, \dots, X_J$ sont des vecteurs propres relatifs à la même valeur propre λ_0 , alors, pour tous complexes $c_1, \dots, c_J$, le vecteur

$$X = c_1 X_1 + \dots + c_J X_J = \sum_{j=1}^J c_j X_j,$$

s'il n'est pas nul, est aussi un vecteur propre relatif à la valeur propre λ_0 .

Preuve. Vu les propriétés du produit matriciel, on a

$$AX = A(c_1 X_1 + \dots + c_J X_J) = c_1 AX_1 + \dots + c_J AX_J;$$

comme les vecteurs X_j ($j = 1, \dots, J$) sont tous des vecteurs propres relatifs à la valeur propre λ_0 , on a

$$AX_j = \lambda_0 X_j \quad \text{pour tout } j = 1, \dots, J.$$

Il s'ensuit que

$$\begin{aligned} AX &= c_1 AX_1 + \dots + c_J AX_J \\ &= c_1 \lambda_0 X_1 + \dots + c_J \lambda_0 X_J \\ &= \lambda_0 (c_1 X_1 + \dots + c_J X_J) \\ &= \lambda_0 X. \end{aligned}$$

□

Propriété 1.7.5 Des vecteurs propres relatifs à des valeurs propres distinctes sont toujours linéairement indépendants.

Preuve. C'est un simple calcul. □

Cette propriété signifie notamment que si X_1 est un vecteur propre relatif à la valeur propre λ_1 et X_2 un vecteur propre relatif à la valeur propre λ_2 , avec $\lambda_1 \neq \lambda_2$, alors X_1, X_2 sont linéairement indépendants.

De même, si les valeurs propres λ_i ($i = 1, 2, 3$) sont distinctes et si X_i ($i = 1, 2, 3$) sont des vecteurs propres respectifs, alors X_1, X_2, X_3 sont linéairement indépendants.

1.7.3 Exemples

1) Les valeurs propres de la matrice

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

sont 1 et -1 .

En effet, le polynôme caractéristique de A est

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1).$$

Cherchons les vecteurs propres associés à la valeur propre 1. On doit trouver les vecteurs

$$X = \begin{pmatrix} x \\ y \end{pmatrix}$$

non nuls tels que

$$(A - I)X = (A - I)X = 0;$$

comme

$$(A - I)X = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} X = \begin{pmatrix} -x + y \\ x - y \end{pmatrix}$$

on a

$$(A - I)X = 0 \Leftrightarrow x - y = 0 \Leftrightarrow X = x \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Il s'ensuit que l'ensemble des vecteurs propres associés à la valeur propre 1 sont les vecteurs

$$X = c \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad c \in \mathbb{C}_0.$$

Cherchons les vecteurs propres associés à la valeur propre -1 . On doit trouver les vecteurs

$$X = \begin{pmatrix} x \\ y \end{pmatrix}$$

non nuls tels que

$$(A - (-1)I)X = (A + I)X = 0;$$

comme

$$(A + I)X = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} X = \begin{pmatrix} x + y \\ x + y \end{pmatrix}$$

on a

$$(A + I)X = 0 \Leftrightarrow x + y = 0 \Leftrightarrow X = y \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Il s'ensuit que l'ensemble des vecteurs propres associés à la valeur propre -1 sont les vecteurs

$$X = c \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad c \in \mathbb{C}_0.$$

2) Les valeurs propres de la matrice

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

sont i et $-i$.

En effet, le polynôme caractéristique de A est

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = \lambda^2 - i^2 = (\lambda - i)(\lambda + i).$$

Cherchons les vecteurs propres associés à la valeur propre i . On doit trouver les vecteurs

$$X = \begin{pmatrix} x \\ y \end{pmatrix}$$

non nuls tels que

$$(A - iI)X = 0;$$

comme

$$(A - iI)X = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} X = \begin{pmatrix} -ix - y \\ x - iy \end{pmatrix}$$

on a

$$(A - iI)X = 0 \Leftrightarrow x - iy = 0 \Leftrightarrow X = y \begin{pmatrix} i \\ 1 \end{pmatrix}.$$

Il s'ensuit que l'ensemble des vecteurs propres associés à la valeur propre i sont les vecteurs

$$X = c \begin{pmatrix} i \\ 1 \end{pmatrix}, \quad c \in \mathbb{C}_0.$$

Cherchons les vecteurs propres associés à la valeur propre $-i$. On doit trouver les vecteurs

$$X = \begin{pmatrix} x \\ y \end{pmatrix}$$

non nuls tels que

$$(A - (-i)I)X = (A + iI)X = 0;$$

comme

$$(A + iI)X = \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} X = \begin{pmatrix} ix - y \\ x + iy \end{pmatrix}$$

on a

$$(A + iI)X = 0 \Leftrightarrow x + iy = 0 \Leftrightarrow X = y \begin{pmatrix} -i \\ 1 \end{pmatrix}.$$

Il s'ensuit que l'ensemble des vecteurs propres associés à la valeur propre $-i$ sont les vecteurs

$$X = c \begin{pmatrix} -i \\ 1 \end{pmatrix}, \quad c \in \mathbb{C}_0.$$

3) Les valeurs propres de la matrice

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

sont $0, 0$, c'est-à-dire que cette matrice possède la valeur propre double 0 .

En effet, le polynôme caractéristique de A est

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2.$$

Cherchons les vecteurs propres associés à cette valeur propre. On doit trouver les vecteurs

$$X = \begin{pmatrix} x \\ y \end{pmatrix}$$

non nuls tels que

$$(A - 0I)X = AX = 0;$$

comme

$$AX = \begin{pmatrix} y \\ 0 \end{pmatrix}$$

on a

$$AX = 0 \Leftrightarrow y = 0.$$

L'ensemble des vecteurs propres recherchés est l'ensemble des vecteurs qui s'écrivent

$$X = \begin{pmatrix} x \\ 0 \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad x \in \mathbb{C}_0.$$

4) Les valeurs propres de la matrice

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

sont $0, 0, 1$, c'est-à-dire que cette matrice possède la valeur propre double 0 et la valeur propre simple 1 .

En effet, le polynôme caractéristique de A est

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & 1 - \lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix} = \lambda^2(1 - \lambda).$$

Cherchons les vecteurs propres associés à la valeur propre 0 . On doit trouver les vecteurs

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

non nuls tels que

$$(A - 0I)X = AX = 0;$$

comme

$$AX = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} X = \begin{pmatrix} y \\ y \\ 0 \end{pmatrix}$$

on a

$$AX = 0 \Leftrightarrow y = 0 \Leftrightarrow X = \begin{pmatrix} x \\ 0 \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Il s'ensuit que l'ensemble des vecteurs propres associés à la valeur propre 0 sont les vecteurs

$$X = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad x, z \in \mathbb{C}, \text{ non simultanément nuls.}$$

Cherchons les vecteurs propres associés à la valeur propre 1. On doit trouver les vecteurs

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

non nuls tels que

$$(A - I)X = (A - I)X = 0;$$

comme

$$(A - I)X = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} X = \begin{pmatrix} -x + y \\ 0 \\ -z \end{pmatrix}$$

on a

$$(A - I)X = 0 \Leftrightarrow \begin{cases} x = y \\ z = 0 \end{cases} \Leftrightarrow X = \begin{pmatrix} x \\ x \\ 0 \end{pmatrix} = x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Il s'ensuit que l'ensemble des vecteurs propres associés à la valeur propre 1 sont les vecteurs

$$X = x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad x \in \mathbb{C}_0.$$

5) Les valeurs propres de la matrice

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

sont 1, 1, 1, c'est-à-dire que cette matrice possède la valeur propre triple 1.

En effet, le polynôme caractéristique de A est

$$\det(A - \lambda I) = \begin{vmatrix} 1 - \lambda & 1 & 0 \\ 0 & 1 - \lambda & 0 \\ 0 & 0 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^3.$$

Cherchons les vecteurs propres associés à cette valeur propre 1. On doit trouver les vecteurs

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

non nuls tels que

$$(A - I)X = (A - I)X = 0;$$

comme

$$(A - I)X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} X = \begin{pmatrix} y \\ 0 \\ 0 \end{pmatrix}$$

on a

$$(A - I)X = 0 \Leftrightarrow y = 0 \Leftrightarrow X = \begin{pmatrix} x \\ 0 \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Il s'ensuit que l'ensemble des vecteurs propres associés à la valeur propre 1 sont les vecteurs

$$X = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad x, z \in \mathbb{C} \text{ non simultanément nuls.}$$

6) Les valeurs propres de la matrice

$$\begin{pmatrix} 7 & -4 & 8 \\ -2 & 3 & -2 \\ -5 & 4 & -6 \end{pmatrix}$$

sont $-1, 2, 3$.

En effet, le polynôme caractéristique de A est

$$\det(A - \lambda I) = \begin{vmatrix} 7 - \lambda & -4 & 8 \\ -2 & 3 - \lambda & -2 \\ -5 & 4 & -6 - \lambda \end{vmatrix}.$$

La valeur de ce déterminant ne change pas si on remplace la première colonne par la somme de la première colonne et de l'opposé de la troisième; on a donc

$$\det(A - \lambda I) = \begin{vmatrix} -1 - \lambda & -4 & 8 \\ 0 & 3 - \lambda & -2 \\ 1 + \lambda & 4 & -6 - \lambda \end{vmatrix};$$

on effectue alors une mise en évidence (du facteur $1 + \lambda$) puis on remplace la troisième ligne par la somme de la première et de la troisième : on obtient

$$\begin{aligned} \det(A - \lambda I) &= (1 + \lambda) \begin{vmatrix} -1 & -4 & 8 \\ 0 & 3 - \lambda & -2 \\ 1 & 4 & -6 - \lambda \end{vmatrix} \\ &= (1 + \lambda) \begin{vmatrix} -1 & -4 & 8 \\ 0 & 3 - \lambda & -2 \\ 0 & 0 & 2 - \lambda \end{vmatrix} \\ &= -(1 + \lambda)(2 - \lambda)(3 - \lambda). \end{aligned}$$

Cherchons les vecteurs propres associés à la valeur propre -1 . On doit trouver les vecteurs

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

non nuls tels que

$$(A - (-1)I)X = (A + I)X = 0;$$

comme

$$(A + I)X = \begin{pmatrix} 8 & -4 & 8 \\ -2 & 4 & -2 \\ -5 & 4 & -5 \end{pmatrix} X = \begin{pmatrix} 8x - 4y + 8z \\ -2x + 4y - 2z \\ -5x + 4y - 5z \end{pmatrix}$$

on a (on fait disparaître facilement le terme en y)

$$(A + I)X = 0 \Leftrightarrow \begin{cases} 6x + 6z = 0 \\ 3x + 3z = 0 \\ 2x - y + 2z = 0 \end{cases} \Leftrightarrow \begin{cases} x + z = 0 \\ y = 0 \end{cases} \Leftrightarrow X = \begin{pmatrix} x \\ 0 \\ -x \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Il s'ensuit que l'ensemble des vecteurs propres associés à la valeur propre -1 sont les vecteurs

$$X = x \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad x \in \mathbb{C}_0.$$

Cherchons les vecteurs propres associés à la valeur propre 2 . On doit trouver les vecteurs

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

non nuls tels que

$$(A - 2I)X = 0;$$

comme

$$(A - 2I)X = \begin{pmatrix} 5 & -4 & 8 \\ -2 & 1 & -2 \\ -5 & 4 & -8 \end{pmatrix} X = \begin{pmatrix} 5x - 4y + 8z \\ -2x + y - 2z \\ -5x + 4y - 8z \end{pmatrix}$$

on a (les première et troisième équations sont les mêmes)

$$(A - 2I)X = 0 \Leftrightarrow \begin{cases} 5x - 4y + 8z = 0 \\ -2x + y - 2z = 0 \end{cases} \Leftrightarrow \begin{cases} -3x = 0 \\ -2x + y - 2z = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = 2z \end{cases} \Leftrightarrow X = z \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}.$$

Il s'ensuit que l'ensemble des vecteurs propres associés à la valeur propre 2 sont les vecteurs

$$X = z \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \quad z \in \mathbb{C}_0.$$

Cherchons les vecteurs propres associés à la valeur propre 3. On doit trouver les vecteurs

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

non nuls tels que

$$(A - 3I)X = 0;$$

comme

$$(A - 3I)X = \begin{pmatrix} 4 & -4 & 8 \\ -2 & 0 & -2 \\ -5 & 4 & -9 \end{pmatrix} X = \begin{pmatrix} 4x - 4y + 8z \\ -2x - 2z \\ -5x + 4y - 9z \end{pmatrix}$$

on a (on fait disparaître le terme en y)

$$(A - 3I)X = 0 \Leftrightarrow \begin{cases} 4x - 4y + 8z = 0 \\ -2x - 2z = 0 \\ -x - z = 0 \end{cases} \Leftrightarrow \begin{cases} x - y + 2z = 0 \\ x = -z \end{cases} \Leftrightarrow \begin{cases} y = z \\ x = -z \end{cases} \Leftrightarrow X = z \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

Il s'ensuit que l'ensemble des vecteurs propres associés à la valeur propre 3 sont les vecteurs

$$X = z \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad z \in \mathbb{C}_0.$$

1.8 Diagonalisation

Comme annoncé précédemment, ici aussi on ne considère que des matrices et des vecteurs de dimension deux ou trois.

Dans l'étude des opérateurs linéaires (entre vecteurs de l'espace, du plan, ou autres...), il est utile de trouver une base dans laquelle la matrice qui représente un opérateur donné est diagonale. En fait, cette matrice est diagonale si et seulement si les vecteurs de la base sont des vecteurs propres pour la matrice, comme on va le voir ci-dessous.

1.8.1 Définition et propriétés

Définition 1.8.1 Une matrice A est diagonalisable s'il existe une matrice inversible S telle que

$$S^{-1}AS = \text{matrice diagonale}.$$

Remarque 1.8.2 Si la matrice inversible S est telle que $S^{-1}AS$ soit une matrice diagonale, on dit que S diagonalise A .

Diagonaliser A consiste à déterminer S et la matrice diagonale correspondante $S^{-1}AS$.

En fait, quand on sait que A est diagonalisable, trouver les éléments diagonaux de $S^{-1}AS$ est simple, comme on va s'en rendre compte tout de suite. Construire S n'est pas difficile non plus mais nécessite davantage de calculs.

Cependant, il n'est pas possible de diagonaliser toutes les matrices, comme nous allons le voir ci-dessous.

Propriété 1.8.3 Si A est diagonalisable par S alors les éléments diagonaux de la matrice $S^{-1}AS$ sont les valeurs propres de A et les colonnes de S sont des vecteurs propres de A .

Par exemple, si $S = (C_1, C_2, C_3)$ est de dimension 3, la preuve donne aussi

$$S^{-1}AS = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \Leftrightarrow AC_i = a_i C_i \quad (i = 1, 2, 3),$$

ce qui signifie que les colonnes de S sont successivement des vecteurs propres associés aux éléments diagonaux, dans le même ordre.

Preuve. Effectuons la preuve pour des matrices de dimension trois. La même chose peut être faite pour des matrices de dimension deux.

Si

$$S^{-1}AS = \text{diag}(a_1, a_2, a_3) = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}$$

on a

$$\det(S^{-1}AS - \lambda I) = \det \begin{pmatrix} a_1 - \lambda & 0 & 0 \\ 0 & a_2 - \lambda & 0 \\ 0 & 0 & a_3 - \lambda \end{pmatrix} = (a_1 - \lambda)(a_2 - \lambda)(a_3 - \lambda).$$

Comme on a aussi

$$S^{-1}AS - \lambda I = S^{-1}AS - \lambda S^{-1}S = S^{-1}(A - \lambda I)S$$

donc

$$\det(S^{-1}AS - \lambda I) = \det S^{-1} \det(A - \lambda I) \det S = \det(A - \lambda I)$$

on obtient

$$\det(A - \lambda I) = (a_1 - \lambda)(a_2 - \lambda)(a_3 - \lambda).$$

Les éléments diagonaux de $S^{-1}AS$ sont donc bien les valeurs propres de A .

Notons C_1, C_2, C_3 les colonnes de S . Les colonnes de AS sont alors les vecteurs

$$AC_1, AC_2, AC_3.$$

Si

$$S^{-1}AS = \text{diag}(a_1, a_2, a_3) = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}$$

on a

$$\begin{aligned} AS &= (AC_1, AC_2, AC_3) \\ &= S \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \\ &= (C_1, C_2, C_3) \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \\ &= (a_1 C_1, a_2 C_2, a_3 C_3) \end{aligned}$$

donc

$$AC_1 = a_1 C_1, \quad AC_2 = a_2 C_2, \quad AC_3 = a_3 C_3.$$

□

Le résultat suivant donne une **condition nécessaire et suffisante pour qu'une matrice soit diagonalisable**.

Théorème 1.8.4 *Une matrice carrée de dimension 2 est diagonalisable si et seulement si elle possède deux vecteurs propres linéairement indépendants.*

Une matrice carrée de dimension 3 est diagonalisable si et seulement si elle possède trois vecteurs propres linéairement indépendants.

Preuve. Effectuons la preuve dans le cas de la dimension trois. L'autre cas est analogue.

Si A est diagonalisable, il existe une matrice S inversible telle que $S^{-1}AS$ soit une matrice diagonale dont les éléments diagonaux sont les valeurs propres de A . On vient de voir aussi que les colonnes de S sont trois vecteurs propres de A . Comme le déterminant de S est non nul, les colonnes de cette matrice sont

des vecteurs linéairement indépendants. La matrice A possède donc trois vecteurs propres linéairement indépendants.

Réciproquement, si $\lambda_1, \lambda_2, \lambda_3$ sont les valeurs propres de A et si les vecteurs X_1, X_2, X_3 sont des vecteurs linéairement indépendants et tels que

$$AX_1 = \lambda_1 X_1, \quad AX_2 = \lambda_2 X_2, \quad AX_3 = \lambda_3 X_3$$

alors la matrice S dont les colonnes sont les vecteurs X_1, X_2, X_3 est inversible et

$$AS = (AX_1, AX_2, AX_3) = (\lambda_1 X_1, \lambda_2 X_2, \lambda_3 X_3) = (X_1, X_2, X_3) \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

donc

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

□

Afin de simplifier au maximum les calculs dans la recherche de la diagonalisabilité et de la diagonalisation d'une matrice, énonçons quelques propriétés relatives aux vecteurs propres.

Propriété 1.8.5 1) Si une matrice possède des valeurs propres simples (c'est-à-dire toutes de multiplicité 1) alors elle est diagonalisable.

2) Si une matrice de dimension deux (resp. trois) possède une seule valeur propre, alors elle est diagonalisable si et seulement s'il existe deux (resp. trois) vecteurs propres linéairement indépendants relatifs à l'unique valeur propre de la matrice.

3) Si une matrice de dimension trois possède une valeur propre simple et une double, alors elle est diagonalisable si et seulement s'il existe deux vecteurs propres linéairement indépendants relatifs à la valeur propre double.

Preuve. 1) On sait que toute valeur propre possède au moins un vecteur propre. Si la matrice est de dimension deux (resp. trois), on prend alors un vecteur propre associé à chacune des deux (resp. trois) valeurs propres de la matrice. Comme des vecteurs propres relatifs à des valeurs propres distinctes sont linéairement indépendants, on obtient ainsi deux (resp. trois) vecteurs propres linéairement indépendants. On conclut en utilisant le théorème 1.8.4.

2) Ce n'est qu'un cas particulier du théorème 1.8.4.

3) En prenant deux vecteurs propres linéairement indépendants relatifs à la valeur propre double et un vecteur non nul relatif à la valeur propre simple, on constitue une famille de trois vecteurs propres linéairement indépendants (cf (1)).

La réciproque est admise (elle nécessite la connaissance du fait que la dimension d'un espace propre ne dépasse pas la multiplicité de la valeur propre comme zéro du polynôme caractéristique). □

1.8.2 Exemples

Reprenons les exemples de matrices donnés précédemment. Pour chaque cas, on demande si la matrice est diagonalisable et si la réponse est oui, on demande de la diagonaliser.

1) La matrice

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

a comme valeurs propres les nombres 1 et -1 . Les deux valeurs propres étant simples, cette matrice est diagonalisable.

Les vecteurs

$$X_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad X_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

sont respectivement des vecteurs propres de valeur propre 1 et -1 . La matrice

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

est telle que

$$S^{-1}AS = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

2) La matrice

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

a comme valeurs propres les complexes i et $-i$. La matrice est donc diagonalisable car elle est de dimension deux et possède deux valeurs propres distinctes.

Les vecteurs

$$X_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}, \quad X_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$$

sont respectivement des vecteurs propres de valeur propre i et $-i$. La matrice

$$S = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}$$

est telle que

$$S^{-1}AS = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

3) La matrice

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

a la valeur propre double 0.

L'ensemble des vecteurs propres de valeur propre 0 sont les vecteurs qui s'écrivent

$$X = \begin{pmatrix} x \\ 0 \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad x \in \mathbb{C}.$$

Ils sont tous multiples du vecteur

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix};$$

deux vecteurs propres sont donc toujours linéairement dépendants. Il s'ensuit que la matrice A n'est pas diagonalisable.

4) La matrice

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

a les valeurs propres 0 et 1 (valeur propre double 0 et la valeur propre simple 1).

Le vecteur

$$X = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

est un vecteur propre relatif à la valeur propre simple 1.

L'ensemble des vecteurs propres associés à la valeur propre double 0 sont les vecteurs

$$X = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad x, z \in \mathbb{C}, \text{ non simultanément nuls.}$$

En particulier, les vecteurs

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sont des vecteurs linéairement indépendants.

Il s'ensuit que la matrice de départ est diagonalisable et que

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

est inversible et telle que

$$S^{-1}AS = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

5) La matrice

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

a 1 comme valeur propre (valeur propre triple).

L'ensemble des vecteurs propres associés à la valeur propre 1 sont les vecteurs

$$X = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad x, z \in \mathbb{C} \quad \text{non simultanément nuls.}$$

Trois vecteurs propres sont donc toujours linéairement dépendants. La matrice de départ n'est donc pas diagonalisable.

6) Les valeurs propres de la matrice

$$\begin{pmatrix} 7 & -4 & 8 \\ -2 & 3 & -2 \\ -5 & 4 & -6 \end{pmatrix}$$

sont $-1, 2, 3$. Ces valeurs propres étant toutes distinctes, la matrice est diagonalisable.

Les vecteurs

$$X_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \quad X_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

sont respectivement des vecteurs propres associés aux valeurs propres $-1, 2, 3$. Il s'ensuit que

$$S = (X_1, X_2, X_3) = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ -1 & 1 & 1 \end{pmatrix}$$

est inversible et telle que

$$S^{-1}AS = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

1.9 Matrices particulières

Une matrice carrée est dite

- *normale* si elle commute avec son adjointe, c'est-à-dire si $AA^* = A^*A$,
- *symétrique* si $\tilde{A} = A$
- *hermitienne* si $A^* = A$
- *unitaire* si $AA^* = A^*A = I$.

Une matrice réelle unitaire est appelée *matrice orthogonale*.

Vu les définitions, une matrice hermitienne réelle est une matrice symétrique.

Les matrices normales, hermitiennes, unitaires, ont une grande importance notamment dans l'étude des opérateurs (ces matrices représentent des opérateurs linéaires particuliers), dans la recherche des extrema de fonctions réelles de plusieurs variables réelles.

Remarquons qu'il y a un lien entre ces différents types de matrices et que des propriétés générales peuvent être directement dégagées.

Propriété 1.9.1 1) Une matrice hermitienne (resp. unitaire) est une matrice normale.

2) Une matrice carrée A telle que $AA^* = I$ (resp. $A^*A = I$) est une matrice unitaire.

3) Une matrice unitaire est toujours inversible. Son inverse est égal à son adjointe et c'est une matrice unitaire.

4) Si une matrice normale (resp. hermitienne) est inversible, son inverse est encore une matrice normale (resp. hermitienne).

Preuve. 1) Si $A = A^*$ alors $AA^* = A^2 = A^*A$. C'est évident pour les matrices unitaires.

2) Vu les propriétés de l'inversion de matrices, la relation $AA^* = I$ (resp. $A^*A = I$) implique que A est inversible et que $A^{-1} = A^*$. On a donc aussi $A^*A = A^{-1}A = I$ (resp. $AA^* = AA^{-1} = I$).

3) Vu les propriétés de l'inverse d'une matrice, de $AA^* = I$ on tire que A est inversible et que $A^{-1} = A^*$.

De plus, $(A^{-1})^*A^{-1} = (A^*)^{-1}A^{-1} = (AA^*)^{-1} = I^{-1} = I$.

4) Si la matrice N est normale et inversible on a

$$(N^{-1})^*N^{-1} = (N^*)^{-1}N^{-1} = (NN^*)^{-1} = (N^*N)^{-1} = N^{-1}(N^*)^{-1} = N^{-1}(N^{-1})^*$$

et, si la matrice H est hermitienne et inversible, on a

$$(H^{-1})^* = (H^*)^{-1} = H^{-1}.$$

□

Remarquons qu'il existe des matrices normales qui ne sont ni hermitiennes ni unitaires. De même, il existe des matrices unitaires (resp. hermitiennes) non hermitiennes (resp. non unitaires) et des matrices qui sont à la fois unitaires et hermitiennes.

Par exemple, toute matrice diagonale est normale mais

$$\begin{pmatrix} 2i & 0 \\ 0 & 3 \end{pmatrix}$$

n'est ni hermitienne ni unitaire. De même la matrice

$$\begin{pmatrix} i & i \\ i & i \end{pmatrix}$$

est normale mais n'est ni hermitienne ni unitaire.

Les matrices

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$$

sont hermitiennes mais pas unitaires.

Les matrices

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

sont unitaires mais pas hermitiennes.

Les matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix}$$

sont unitaires et hermitiennes.

Afin de reconnaître plus aisément certaines de ces matrices, énonçons les propriétés suivantes.

Propriété 1.9.2 1) Une matrice carrée est hermitienne si et seulement si les éléments diagonaux sont réels et les éléments symétriques par rapport à la diagonale principale sont conjugués deux à deux.

2) Soit la matrice réelle

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

Les propriétés suivantes sont équivalentes

(i) la matrice S est orthogonale

(ii) on a

$$\begin{cases} a^2 + b^2 = 1 \\ c^2 + d^2 = 1 \\ ac + bd = 0 \end{cases}$$

(c'est-à-dire les lignes de S sont orthonormées)

(iii) on a

$$\begin{cases} a^2 + c^2 = 1 \\ b^2 + d^2 = 1 \\ ab + cd = 0 \end{cases}$$

(c'est-à-dire les colonnes de S sont orthonormées).

La même propriété peut être énoncée pour des matrices de dimension trois.

3) La matrice réelle

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

est orthogonale si et seulement si elle s'écrit

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$$

avec

$$a, b \in \mathbb{R}, \quad a^2 + b^2 = 1.$$

Preuve. La propriété 1) découle directement de la définition.

2) On sait que la matrice S est orthogonale si et seulement si $S\tilde{S} = I$. Dès lors, par définition du produit matriciel, on obtient que (i) et (ii) sont des propriétés équivalentes.

De même on sait que la matrice S est orthogonale si et seulement si $\tilde{S}S = I$. Dès lors, par définition du produit matriciel, on obtient que (i) et (iii) sont des propriétés équivalentes.

3) On vérifie directement que les matrices proposées sont orthogonales.

Soit alors la matrice orthogonale

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

On a donc

$$\begin{cases} a^2 + b^2 = 1 \\ c^2 + d^2 = 1 \\ ac + bd = 0 \end{cases}$$

A partir de ce système, déterminons les valeurs de c, d en fonction de a, b .

Si $a \neq 0$, on a

$$\begin{cases} a^2 + b^2 = 1 \\ c^2 + d^2 = 1 \\ ac + bd = 0 \end{cases} \Leftrightarrow \begin{cases} a^2 + b^2 = 1 \\ c = -\frac{bd}{a} \\ \frac{b^2 d^2}{a^2} + d^2 = 1 \end{cases} \Leftrightarrow \begin{cases} a^2 + b^2 = 1 \\ c = -\frac{bd}{a} \\ b^2 d^2 + a^2 d^2 = a^2 \end{cases} \Leftrightarrow \begin{cases} a^2 + b^2 = 1 \\ c = -\frac{bd}{a} \\ d = \pm a \end{cases}$$

dès lors

$$U = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{ou} \quad U = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$$

avec $a^2 + b^2 = 1$.

Si $a = 0$, on a

$$\begin{cases} a^2 + b^2 = 1 \\ c^2 + d^2 = 1 \\ ac + bd = 0 \end{cases} \Leftrightarrow \begin{cases} b = \pm 1 \\ d = 0 \\ c^2 = 1 \end{cases} \Leftrightarrow \begin{cases} b = \pm 1 \\ d = 0 \\ c = \pm 1 \end{cases}$$

donc U est égal à l'une des quatre matrices suivantes (qui correspondent toutes à l'une des deux formes canoniques indiquées)

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

□

Voici quelques propriétés spécifiques de ces matrices relatives à la diagonalisation.

Théorème 1.9.3 1) Une matrice normale est toujours diagonalisable par une matrice unitaire. En particulier, toute matrice hermitienne ou unitaire est diagonalisable par une matrice unitaire.

2) Les valeurs propres d'une matrice hermitienne sont toujours réelles.

3) Les valeurs propres d'une matrice unitaire sont toujours des complexes de module 1.

Preuve. Résultat admis. □

1.10 Applications

1.10.1 Utilisation en statistique

En statistique multivariée, la matrice de dispersion Σ remplace l'écart type σ utilisé dans le cas d'une seule variable. Cette matrice permet notamment d'étudier la dispersion des observations autour de la moyenne. C'est une matrice hermitienne réelle (qui est même souvent définie positive, ce qui signifie que ses valeurs propres sont strictement positives). De plus, de nombreux paramètres statistiques sont étudiés à partir des valeurs propres et des vecteurs propres de cette matrice.

En statistique, on utilise aussi beaucoup la trace et le déterminant, notamment de cette matrice de dispersion.

1.10.2 Processus de Markov

UN PROBLÈME À MODÉLISER

On désire étudier des mouvements de population entre la ville et ses faubourgs.

On suppose que, chaque année,

- 95% de la population de la ville y reste,
- 5% de la population de la ville part vers les faubourgs,
- 97% de la population des faubourgs y reste,
- 3% de la population des faubourgs part vers la ville.

Si on donne la population une certaine année, on demande un procédé simple de modélisation de la manière dont elle va évoluer au cours des années suivantes.

UNE MODÉLISATION

Soient V_0, F_0 respectivement les populations de la ville et des faubourgs l'année fixée au départ.

Soient V_1, F_1 respectivement les populations de la ville et des faubourgs un an après. Plus généralement, soient V_k, F_k respectivement les populations de la ville et des faubourgs après k années.

On a donc

$$V_1 = 0.95 V_0 + 0.03 F_0, \quad F_1 = 0.05 V_0 + 0.97 F_0.$$

Cela peut s'écrire sous forme matricielle de la manière suivante

$$\begin{pmatrix} V_1 \\ F_1 \end{pmatrix} = \begin{pmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{pmatrix} \begin{pmatrix} V_0 \\ F_0 \end{pmatrix}.$$

Si on note

$$M = \begin{pmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{pmatrix},$$

la répartition de la population après deux années est

$$\begin{pmatrix} V_2 \\ F_2 \end{pmatrix} = M \begin{pmatrix} V_1 \\ F_1 \end{pmatrix} = M^2 \begin{pmatrix} V_0 \\ F_0 \end{pmatrix}.$$

Plus généralement, après k années, on a

$$\begin{pmatrix} V_k \\ F_k \end{pmatrix} = M^k \begin{pmatrix} V_0 \\ F_0 \end{pmatrix}.$$

Par exemple, si, au départ on a

$$V_0 = 600\,000, \quad F_0 = 400\,000$$

on obtient

$$\begin{pmatrix} V_1 \\ F_1 \end{pmatrix} = M \begin{pmatrix} 600\,000 \\ 400\,000 \end{pmatrix} = \begin{pmatrix} 582\,000 \\ 418\,000 \end{pmatrix}$$

et

$$\begin{pmatrix} V_2 \\ F_2 \end{pmatrix} = M^2 \begin{pmatrix} 600\,000 \\ 400\,000 \end{pmatrix} = \begin{pmatrix} 565\,000 \\ 435\,000 \end{pmatrix}.$$

CHAÎNES DE MARKOV

Définitions

- Une *matrice stochastique* est une matrice (généralement) carrée dont les éléments sont des réels positifs et dont la somme des éléments de chaque colonne est égale à 1.
- Un *vecteur de probabilité* est un vecteur dont les éléments sont des réels positifs et de somme égale à 1. Une matrice stochastique est donc une matrice dont les colonnes sont des vecteurs de probabilité.
- Une matrice stochastique est dite *régulière* lorsque l'une de ses puissances possède des éléments qui sont tous strictement positifs³.

Cela étant, si M est une matrice stochastique et X_0 un vecteur de probabilité, la *chaîne de Markov* associée est la suite de vecteurs de probabilité

$$X_0, X_1 = MX_0, X_2 = MX_1 = M^2X_0, \dots$$

L'intérêt des chaînes de Markov est la prévision à long terme d'une situation, comme on va le voir ci-dessous.

3. On va démontrer que toute puissance d'une matrice stochastique est encore une matrice stochastique

Propriétés générales des matrices stochastiques

1. Si T est une matrice stochastique et X un vecteur de probabilité, alors le vecteur TX est encore un vecteur de probabilité.
2. Si T est une matrice stochastique et si k est un naturel, alors T^k est encore une matrice stochastique.
3. Une matrice stochastique possède toujours la valeur propre 1 et toutes les valeurs propres sont de module inférieur ou égal à 1.

Preuve. Supposons que les matrices (et les vecteurs) soient de dimension n .

1) Quel que soit $j \in \{1, \dots, n\}$, on a $(TX)_j = \sum_{k=1}^n (T)_{j,k} (X)_k$ donc les éléments de TX sont des réels positifs. De plus, leur somme vaut 1 car on a

$$\sum_{j=1}^n (TX)_j = \sum_{j=1}^n \sum_{k=1}^n (T)_{j,k} (X)_k = \sum_{k=1}^n (X)_k \left(\sum_{j=1}^n (T)_{j,k} \right) = \sum_{k=1}^n (X)_k = 1.$$

2) Procédons par récurrence. On suppose que T est une matrice stochastique. Montrons alors que si k est un naturel tel que la matrice T^k soit stochastique, alors la matrice T^{k+1} l'est aussi. On pourra alors conclure.

De fait, si on désigne par $C_1, \dots, C_n$ les colonnes de T^k , la matrice T^{k+1} est

$$T^{k+1} = T(C_1 \dots C_n) = (TC_1 \dots, TC_n)$$

c'est-à-dire que les colonnes de T^{k+1} apparaissent comme résultant du produit de la matrice stochastique T et des vecteurs de probabilité $C_1, \dots, C_n$. Vu ce qui précède, il s'agit donc de vecteurs de probabilité et on conclut.

3) Si X est le vecteur colonne dont tous les éléments sont égaux à 1, alors on a $\tilde{T}X = X$, par définition des matrices stochastiques. Le nombre 1 est donc valeur propre de $\tilde{T}$, donc de T puisque ces matrices possèdent les mêmes valeurs propres.

Soit maintenant λ une valeur propre quelconque de T . Soit alors X , vecteur propre de $\tilde{T}$ relatif à celle-ci. Si $k \in \{1, \dots, n\}$ est tel que $|(X)_k| = \sup \{|(X)_j| : 1 \leq j \leq n\}$ alors $|(X)_k| \neq 0$ et

$$\begin{aligned} |\lambda| &= \frac{1}{|(X)_k|} |(\lambda X)_k| = \frac{1}{|(X)_k|} \left| (\tilde{T}X)_k \right| = \frac{1}{|(X)_k|} \left| \sum_{j=1}^n (\tilde{T})_{k,j} (X)_j \right| \\ &\leq \frac{1}{|(X)_k|} \sum_{j=1}^n (\tilde{T})_{k,j} |(X)_j| \\ &\leq \sum_{j=1}^n (\tilde{T})_{k,j} = 1 \end{aligned}$$

Propriétés plus spécifiques relatives aux chaînes de Markov

1. Exemple des matrices de dimension 2. Soit T une matrice stochastique, à savoir

$$T = \begin{pmatrix} a & b \\ 1-a & 1-b \end{pmatrix}$$

avec $a, b \in [0, 1]$. Alors

- (1) les valeurs propres de T sont les réels 1 et $a - b$
- (2) si $a - b \neq 1$, il existe un unique vecteur de probabilité qui soit un vecteur propre relatif à la valeur propre 1
- (3) si $|a - b| < 1$, pour toute condition initiale X (vecteur de probabilité), la suite $T^k X$ ($k \in \mathbb{N}_0$) converge vers l'unique vecteur de probabilité qui soit un vecteur propre relatif à la valeur propre 1.

2. Soit T une matrice stochastique régulière. Alors

- (1) la valeur propre 1 est de multiplicité 1 et il existe un unique vecteur de probabilité qui soit un vecteur propre de valeur propre 1
 (2) pour toute condition initiale X (vecteur de probabilité), la suite $T^k X$ ($k \in \mathbb{N}_0$) converge vers l'unique vecteur de probabilité qui soit un vecteur propre de valeur propre 1.

Preuve. 1) Vu la définition, toute matrice stochastique peut effectivement s'écrire de manière annoncée.

(1) Par un calcul direct, on obtient

$$\det(T - \lambda I) = \det \begin{pmatrix} a - \lambda & b \\ 1 - a & 1 - b - \lambda \end{pmatrix} = (\lambda - 1)(\lambda - (a - b)).$$

D'où la conclusion du premier point.

(2) Les vecteurs propres relatifs à la valeur propre 1 sont les vecteurs non nuls $\begin{pmatrix} x \\ y \end{pmatrix}$ tels que

$$\begin{pmatrix} a - 1 & b \\ 1 - a & -b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

ou encore tels que

$$(a - 1)x + by = 0.$$

Comme les coefficients $a - 1$ et b ne sont pas simultanément nuls, il s'agit donc des vecteurs

$$\begin{pmatrix} x \\ y \end{pmatrix} = r \begin{pmatrix} b \\ 1 - a \end{pmatrix} \quad r \in \mathbb{C}_0.$$

Si un tel vecteur est de probabilité, alors on a $rb + r(1 - a) = 1$ donc $r = 1/(1 - a + b)$, donc on a l'unicité. Par ailleurs, le vecteur propre obtenu en prenant cette valeur de r est bien un vecteur de probabilité. Donc on conclut pour le point (2).

(3) Désignons par X_0 le vecteur propre de probabilité relatif à la valeur propre 1 et désignons par Y un vecteur propre relatif à la valeur propre $a - b$. Si X est un vecteur quelconque de probabilité, alors il existe des complexes c, c' tels que

$$X = cX_0 + c'Y.$$

On applique alors successivement la matrice T aux deux membres de l'égalité et on obtient ainsi

$$T^k X = cX_0 + c'(a - b)^k Y$$

quel que soit le naturel k . Comme $|a - b| < 1$, on en déduit que la suite $T^k X$ ($k \in \mathbb{N}_0$) converge vers cX_0 . Pour conclure, il reste donc à montrer que $c = 1$. De fait, si on désigne par α, β les éléments de Y et par x_0, y_0 les éléments de X_0 , comme $T^k X$ est un vecteur de probabilité quel que soit k , on a

$$c(x_0 + y_0) + c'(a - b)^k(\alpha + \beta) = 1, \quad k \in \mathbb{N}_0.$$

Comme la suite $(a - b)^k$ ($k \in \mathbb{N}_0$) converge vers 0 et comme $x_0 + y_0 = 1$, on en déduit que $c = 1$ et on conclut.

2) Admis.

Le résultat précédent signifie donc que si l'on veut savoir ce que devient la population à long terme, on calcule le vecteur propre de valeur propre 1 dont la somme des éléments est 1 (il s'agit de pourcentages!). Bien sûr, la situation décrite par le vecteur ne se réalise en général pas (il s'agit d'une limite!) mais les vecteurs $X_k = M^k X_0$ s'approchent de plus en plus de cette situation.

Reprenons l'exemple introduit au début de cette section. Après calcul, on trouve que le vecteur de probabilité qui soit un vecteur propre de valeur propre 1 est le vecteur

$$X = \begin{pmatrix} 0.375 \\ 0.625 \end{pmatrix}.$$

A partir d'une population de 600 000 habitants en ville et de 400 000 habitants dans les faubourgs, c'est-à-dire une

population totale de 1 000 000 d'habitants,
on évolue vers une population de
375 000 habitants en ville et 625 000 habitants dans les faubourgs.

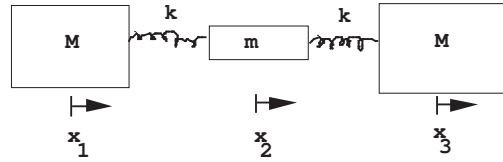
1.10.3 Modes normaux de vibration

Nous allons expliquer comment la diagonalisation peut être utile dans la résolution d'équations différentielles couplées.

INTRODUCTION

Considérons 3 masses vibrantes situées sur l'axe X (on peut voir cette situation comme celle de trois atomes vibrant entre eux). On repère les masses par leur déplacement par rapport à leur position d'équilibre, x_1, x_2, x_3 . La physique indique que ce système est régi par les équations différentielles suivantes (t =variable temporelle, m, M, k constantes strictement positives)

$$\begin{cases} D_t^2 x_1 &= -\frac{k}{M}(x_1 - x_2) \\ D_t^2 x_2 &= -\frac{k}{m}(x_2 - x_1) - \frac{k}{m}(x_2 - x_3) \\ D_t^2 x_3 &= -\frac{k}{M}(x_3 - x_2). \end{cases}$$



Le but est de déterminer la manière dont les masses vibrent entre elles, c'est-à-dire la forme des solutions $x_1(t), x_2(t), x_3(t)$, t étant la variable temporelle.

RÉSOLUTION

Le système précédent peut s'écrire sous la forme

$$D_t^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -\frac{k}{M} & \frac{k}{M} & 0 \\ \frac{k}{m} & -\frac{2k}{m} & \frac{k}{m} \\ 0 & \frac{k}{M} & -\frac{k}{M} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \quad (1.1)$$

Posons

$$A = \begin{pmatrix} -\frac{k}{M} & \frac{k}{M} & 0 \\ \frac{k}{m} & -\frac{2k}{m} & \frac{k}{m} \\ 0 & \frac{k}{M} & -\frac{k}{M} \end{pmatrix}.$$

Essayons de diagonaliser cette matrice (cela conduira à un découplage des équations).

On a

$$\begin{aligned} \begin{vmatrix} -\frac{k}{M} - \lambda & \frac{k}{M} & 0 \\ \frac{k}{m} & -\frac{2k}{m} - \lambda & \frac{k}{m} \\ 0 & \frac{k}{M} & -\frac{k}{M} - \lambda \end{vmatrix} &= \begin{vmatrix} -\lambda & \frac{k}{M} & 0 \\ -\lambda & -\frac{2k}{m} - \lambda & \frac{k}{m} \\ -\lambda & \frac{k}{M} & -\frac{k}{M} - \lambda \end{vmatrix} \quad (C'_1 = C_1 + C_2 + C_3) \\ &= -\lambda \begin{vmatrix} 1 & \frac{k}{M} & 0 \\ 1 & -\frac{2k}{m} - \lambda & \frac{k}{m} \\ 1 & \frac{k}{M} & -\frac{k}{M} - \lambda \end{vmatrix} \\ &= -\lambda \begin{vmatrix} 1 & \frac{k}{M} & 0 \\ 0 & -\frac{2k}{m} - \frac{k}{M} - \lambda & \frac{k}{m} \\ 0 & 0 & -\frac{k}{M} - \lambda \end{vmatrix} \\ &= -\lambda \left(\frac{k}{M} + \lambda \right) \left(\lambda + \frac{2k}{m} + \frac{k}{M} \right). \end{aligned}$$

Les valeurs propres de cette matrice sont donc

$$0, -\frac{k}{M}, -\frac{k}{M} - \frac{2k}{m}.$$

Comme ces valeurs propres sont distinctes, la matrice est diagonalisable.

Les vecteurs propres X relatifs à la valeur propre 0 vérifient

$$\begin{pmatrix} -\frac{k}{M} & \frac{k}{M} & 0 \\ \frac{k}{m} & -\frac{2k}{m} & \frac{k}{m} \\ 0 & \frac{k}{M} & -\frac{k}{M} \end{pmatrix} X = 0.$$

Si

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

cette équation est équivalente au système

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

ou encore au système

$$\begin{cases} x = y \\ y = z \end{cases}$$

Les vecteurs propres relatifs à la valeur propre 0 s'écrivent donc

$$c \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad c = \text{constante} \neq 0.$$

Les vecteurs propres X relatifs à la valeur propre $-\frac{k}{M}$ vérifient

$$\begin{pmatrix} 0 & \frac{k}{M} & 0 \\ \frac{k}{m} & -\frac{2k}{m} + \frac{k}{M} & \frac{k}{m} \\ 0 & \frac{k}{M} & 0 \end{pmatrix} X = 0.$$

Si

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

cette équation est équivalente au système

$$\begin{cases} y = 0 \\ \frac{k}{m}x + (-\frac{2k}{m} + \frac{k}{M})y + \frac{k}{m}z = 0 \end{cases}$$

ou encore au système

$$\begin{cases} y = 0 \\ z = -x \end{cases}$$

Les vecteurs propres s'écrivent donc

$$c \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad c = \text{constante} \neq 0.$$

Les vecteurs propres X relatifs à la valeur propre $-\frac{k}{M} - \frac{2k}{m}$ vérifient

$$\begin{pmatrix} \frac{2k}{m} & \frac{k}{M} & 0 \\ \frac{k}{m} & \frac{M}{k} & \frac{k}{m} \\ 0 & \frac{M}{k} & \frac{2k}{m} \end{pmatrix} X = 0.$$

Si

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

cette équation est équivalente au système

$$\begin{cases} \frac{2k}{m}x + \frac{k}{M}y = 0 \\ \frac{k}{m}x + \frac{M}{k}y + \frac{k}{m}z = 0 \\ \frac{k}{M}y + \frac{2k}{m}z = 0 \end{cases}$$

ou encore au système

$$\begin{cases} x = z \\ y = -\frac{2M}{m}x \end{cases}$$

Les vecteurs propres s'écrivent donc

$$c \begin{pmatrix} 1 \\ -\frac{2M}{m} \\ 1 \end{pmatrix}, \quad c = \text{constante} \neq 0.$$

On a donc

$$\Delta := S^{-1}AS = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\frac{k}{M} & 0 \\ 0 & 0 & -\frac{k}{M} - \frac{2k}{m} \end{pmatrix}$$

avec

$$S = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -\frac{2M}{m} \\ 1 & -1 & 1 \end{pmatrix}.$$

Si on pose

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = S^{-1} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

le système (1.1) est alors équivalent au système

$$D_t^2 \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \Delta \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

lequel s'écrit encore

$$\begin{cases} D_t^2 y_1 &= 0 \\ D_t^2 y_2 &= -\frac{k}{M} y_2 \\ D_t^2 y_3 &= -(\frac{k}{M} + \frac{2k}{m}) y_3. \end{cases}$$

On constate donc que les équations sont maintenant découplées et que chacune d'entre elles se résout aisément (puisque c'est une équation différentielle linéaire à coefficients constants homogène d'ordre deux).

Résolution de $D_t^2 y_1 = 0$.

L'ensemble des solutions réelles de cette équation est l'ensemble des fonctions qui s'écrivent

$$y_1(t) = r_1 t + r_2, \quad r_1, r_2 \in \mathbb{R}.$$

Résolution de $D_t^2 y_2 = -\frac{k}{M} y_2$.

L'ensemble des solutions de cette équation est l'ensemble des fonctions qui s'écrivent

$$y_2(t) = c_1 e^{i\sqrt{k/M}t} + c_2 e^{-i\sqrt{k/M}t}, \quad c_1, c_2 \in \mathbb{C}$$

et l'ensemble des solutions réelles est l'ensemble des fonctions qui s'écrivent

$$y_2(t) = r_1 \cos(\sqrt{k/M}t) + r_2 \sin(\sqrt{k/M}t), \quad r_1, r_2 \in \mathbb{R}.$$

Résolution de $D_t^2 y_3 = -(\frac{k}{M} + \frac{2k}{m})y_3$.

Ce cas est analogue au précédent. L'ensemble des solutions réelles est l'ensemble des fonctions qui s'écrivent

$$y_3(t) = r_1 \cos(\sqrt{k/M + 2k/m}t) + r_2 \sin(\sqrt{k/M + 2k/m}t), \quad r_1, r_2 \in \mathbb{R}.$$

Finalement, les solutions x_1, x_2, x_3 sont données par

$$\begin{aligned} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= S \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & \frac{-2M}{m} \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ &= y_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + y_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + y_3 \begin{pmatrix} 1 \\ -2M/m \\ 1 \end{pmatrix} \end{aligned}$$

où y_1 ne fait intervenir que la fréquence nulle, y_2 ne fait intervenir que la fréquence $\sqrt{k/M}$ et y_3 ne fait intervenir que la fréquence $\sqrt{k/M + 2k/m}$. Ces fréquences sont appelées **modes normaux de vibration**.

CONCLUSION

Le mouvement des trois masses, déterminé par le vecteur de fonctions (du temps)

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

apparaît donc comme une superposition de trois “mouvements fondamentaux”, chacun faisant intervenir un mode normal :

$$(a) y_1(t) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad (b) y_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad (c) y_3 \begin{pmatrix} 1 \\ -2M/m \\ 1 \end{pmatrix}$$

qui correspondent respectivement aux mouvements suivants

$$(a) x_1(t) = x_2(t) = x_3(t) :$$

les masses bougent sans vibration entre elles

$$(b) x_1(t) = -x_3(t), \quad x_2(t) = 0 :$$

immobilité de la masse centrale et mouvement opposé des masses extrêmes

$$(c) x_1(t) = x_3(t), \quad x_2(t) = -2M/m x_3(t) :$$

même mouvement pour les masses extrêmes et mouvement opposé pour la masse centrale.

1.10.4 Changement de repère orthonormé

1.10.5 Ecriture matricielle des transformations linéaires

Chapitre 2

Fonctions de plusieurs variables réelles

Jusqu'à présent, nous avons considéré des fonctions d'une variable réelle, à valeurs réelles surtout. Mais nous avons aussi rencontré des fonctions d'une variable réelle à valeurs complexes et aussi vectorielles.

Nous allons à présent introduire et étudier les fonctions qui dépendent non plus d'une seule, mais de plusieurs variables réelles. Dans un premier temps, nous n'envisagerons que le cas où ces fonctions sont à valeurs réelles et dépendent de deux ou trois variables réelles.

2.1 Définitions, représentations

2.1.1 Définitions et premiers exemples

Définition 2.1.1 Une fonction de deux variables réelles à valeurs réelles est une loi qui, à tout point d'un sous ensemble de $\mathbb{R}^2$ (c'est-à-dire du plan) associe un nombre réel. On écrit

$$f : A \rightarrow \mathbb{R} \quad (x, y) \mapsto f(x, y).$$

L'ensemble des points du plan où f est défini est appelé domaine de définition de f (et est noté $\text{dom}(f)$) et l'ensemble des valeurs de f , c'est-à-dire le sous ensemble de $\mathbb{R}$ $\{f(x, y) : (x, y) \in \text{dom}(f)\}$, est appelé image de f .

Une fonction de trois variables réelles à valeurs réelles est une loi qui, à tout point d'un sous ensemble de $\mathbb{R}^3$ (c'est-à-dire de l'espace) associe un nombre réel. On écrit

$$f : A \rightarrow \mathbb{R} \quad (x, y, z) \mapsto f(x, y, z).$$

L'ensemble des points de l'espace où f est défini est appelé domaine de définition de f (et est noté $\text{dom}(f)$) et l'ensemble des valeurs de f , c'est-à-dire le sous ensemble de $\mathbb{R}$ $\{f(x, y, z) : (x, y, z) \in \text{dom}(f)\}$, est appelé image de f .

Voici quelques exemples de fonctions de plusieurs variables rencontrées de façon usuelle.

- Aire d'un rectangle de côtés x, y :

$$A(x, y) = xy$$

- Volume d'un parallélépipède de base rectangulaire (de côtés x, y et de hauteur z) :

$$V(x, y, z) = xyz$$

- Distance entre un point P de coordonnées (x, y, z) et l'origine des axes :

$$d(x, y, z) = \|\vec{OP}\| = \sqrt{x^2 + y^2 + z^2}$$

- Force gravitationnelle exercée par le soleil (supposé être à l'origine des axes) sur une masse unitaire située au point de coordonnées (x, y, z) :

$$F(x, y, z) = \frac{c}{x^2 + y^2 + z^2}$$

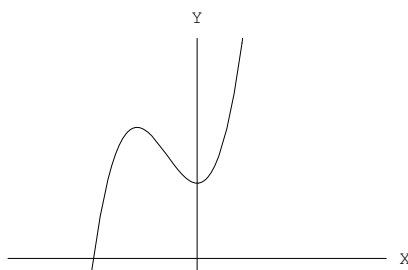
(c est une constante strictement positive)

Représenter une fonction d'une variable réelle, à valeurs réelles, c'est représenter le graphe

$$\{(x, f(x)) : x \in \text{dom}(f)\}.$$

Pour cela, on utilise un repère orthonormé du plan et on représente les points de coordonnées cartésiennes $(x, f(x))$ lorsque x varie dans le domaine de f .

La représentation graphique de f s'appelle une courbe. Il s'agit de la courbe d'équation cartésienne $y = f(x)$. On utilise le terme “courbe” car, par définition, *une courbe est un ensemble de points décrits à l'aide d'une seule variable réelle*.

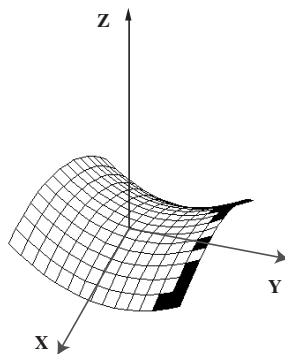


De façon analogue, représenter une fonction de deux variables réelles, à valeurs réelles, c'est représenter l'ensemble

$$\{(x, y, f(x, y)) : (x, y) \in \text{dom}(f)\}.$$

Pour cela, on utilise un repère orthonormé de l'espace et on représente les points de coordonnées cartésiennes $(x, y, f(x, y))$ lorsque (x, y) varie dans le domaine de f .

La représentation graphique de f s'appelle une surface. Il s'agit de la surface d'équation cartésienne $z = f(x, y)$. On utilise le terme “surface” car, par définition, *une surface est un ensemble de points décrits à l'aide de deux variables réelles*.



Si on désire suivre le même processus, représenter une fonction de trois variables réelles à valeurs réelles n'est plus possible : on devrait pouvoir représenter un ensemble de “quadruplets” $(x, y, z, f(x, y, z))$!

2.1.2 Courbes de niveau

Pour donner une idée de f , fonction de deux variables réelles, pour en utiliser certains éléments, en d'autres termes pour examiner des propriétés de la surface qui représente f , on introduit la notion de courbe de niveau.

Si f est une fonction de deux variables et si r appartient à l'image de f , l'ensemble

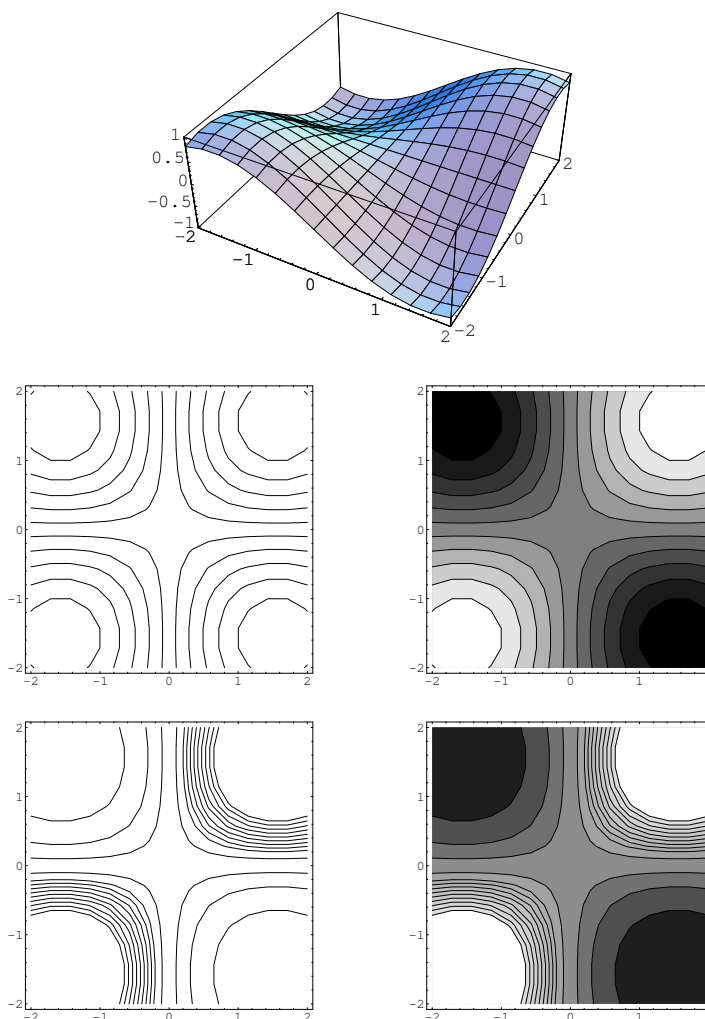
$$\{(x, y) \in \text{dom } f : f(x, y) = r\}$$

s'appelle une courbe de niveau de f . Le terme courbe est encore employé ici car il s'agit de points qui sont décrits à l'aide d'une seule variable réelle.

Sur une carte, ce que l'on appelle “courbe de niveau” est l'ensemble des points de la carte qui sont situés à une même altitude. Si $f(x, y)$ désigne l'altitude au point de coordonnées (x, y) , si h est une

constante positive (une altitude donnée), il s'agit bien de l'ensemble des points où $f(x, y) = h$, ce qui justifie l'appellation. De même, les isothermes d'une région (resp. les isobares), sont des courbes de niveau : ce sont les points (du sol, par exemple) qui ont une même température (resp. pression atmosphérique).

Voici la représentation graphique d'une fonction de deux variables et des représentations de courbes de niveau (la représentation des courbes avec divers niveaux de gris est la même que celle de la voisine mais les niveaux de gris permettent d'augmenter l'information : plus la valeur de la fonction est grande, plus le gris est clair). Sur les premières courbes, la différence entre les divers niveaux représentés est constante ; ce n'est pas le cas sur les secondes.



Remarquons que la représentation graphique d'une fonction f d'une variable est une courbe de niveau. En effet, si on pose $g(x, y) = y - f(x)$, on a

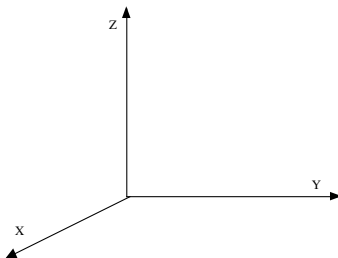
$$g(x, y) = 0 \quad \Leftrightarrow \quad y = f(x)$$

donc la représentation de f est une courbe de niveau de $g(x, y) = y - f(x)$.

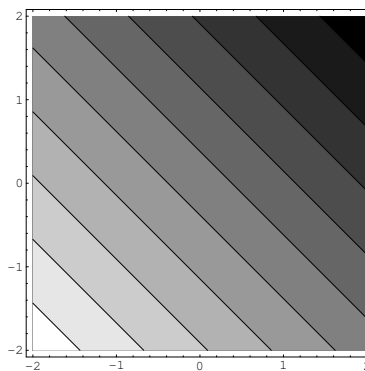
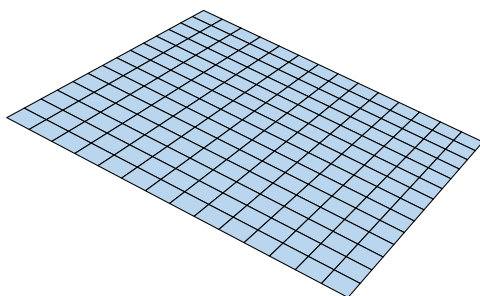
En géologie, il est fréquent de trouver des courbes de ce type sur des cartes (voir votre cours). Les étudiants en géologie rencontreront sans doute ces cartes à de nombreuses reprises avec Monsieur Poty !

2.1.3 Exemples

Dans ce qui suit, les représentations à trois dimensions sont toujours faites dans un système du type représenté ci-dessous. Nous avons omis les axes afin de ne pas alourdir les figures.

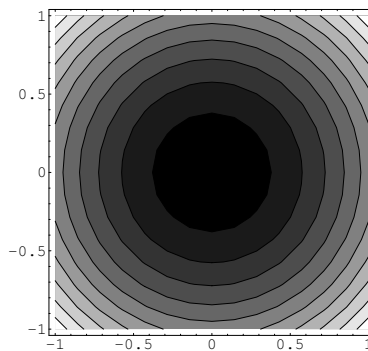
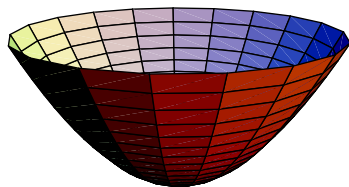


1) Représentation de $f(x, y) = 1 - x - y$ et de courbes de niveau (avec même différence entre les niveaux).

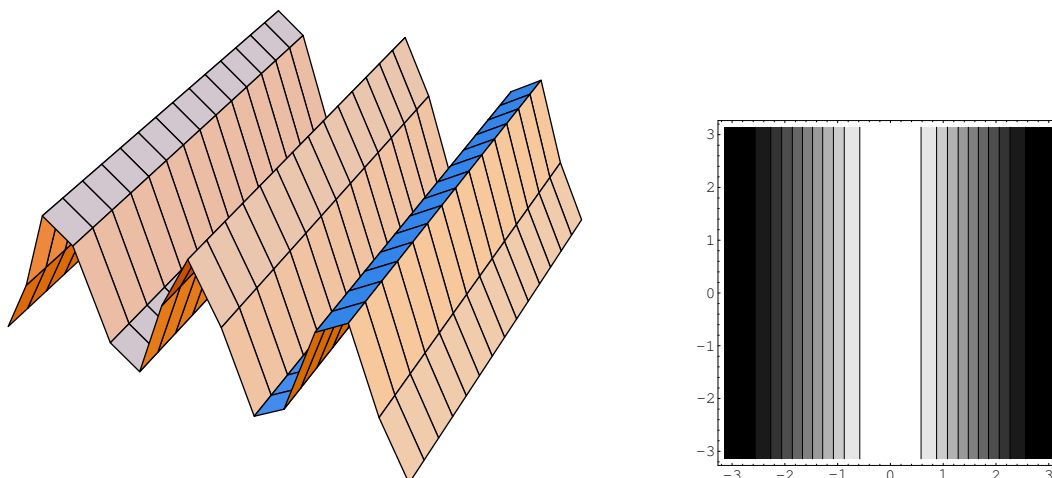


La représentation de f , c'est-à-dire la surface d'équation $z + x + y = 1$ est un plan. En toute généralité, dans l'espace, un plan a une équation du type $ax + by + cz + d = 0$ où a, b, c, d sont réels et où a, b, c ne sont pas simultanément nuls. Réciproquement, une équation de ce type est toujours celle d'un plan.

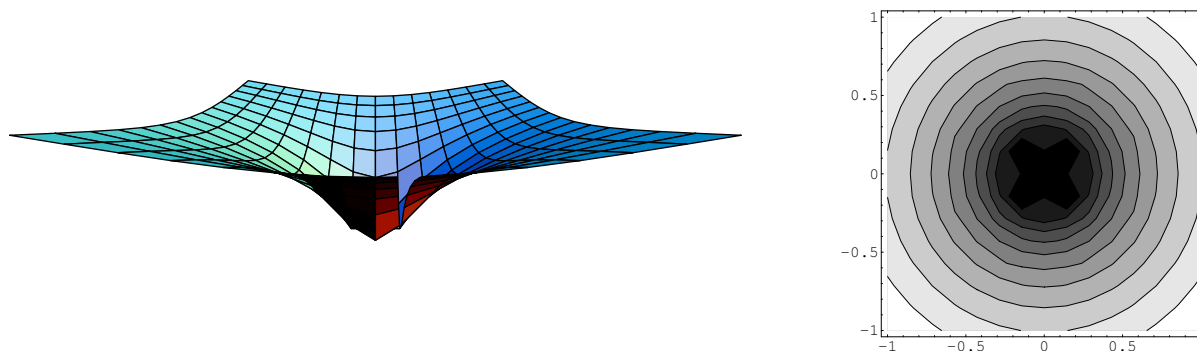
2) Représentation de $f(x, y) = x^2 + y^2$ et de courbes de niveau (avec même différence entre les niveaux).



3) Représentation graphique de $f(x, y) = \cos y$ et de courbes de niveau (avec même différence entre les niveaux).



4) Représentation graphique de $f(x, y) = \ln(x^2 + y^2)$ et de courbes de niveau (avec même différence entre les niveaux).



2.1.4 Surfaces de niveau

Pour donner une idée de f , fonction de trois variables réelles, on introduit la notion de surface de niveau.

Si f est une fonction de trois variables et si r appartient à l'image de f , l'ensemble

$$\{(x, y, z) \in \text{dom } f : f(x, y, z) = r\}$$

s'appelle une surface de niveau de f . Le terme surface est encore employé ici car il s'agit de points qui sont décrits à l'aide de deux variables réelles. Pour donner une idée de la représentation d'une surface de niveau, on effectue souvent l'intersection de cette surface avec des plans orthogonaux aux axes X, Y, Z . Ces intersections sont appelées traces de la surface sur les plans.

Par exemple, si $f(x, y, z)$ désigne la température au point de coordonnées (x, y, z) d'une région de l'espace, la surface de niveau d'équation $f(x, y, z) = c$ est la surface sur laquelle la température est constante et vaut c ; elle est appelée surface isotherme. De même, si $V(x, y, z)$ désigne le potentiel (électrique) au point de coordonnées (x, y, z) , la surface de niveau d'équation $V(x, y, z) = c$ est appelée surface équipotentielle.

Remarquons que la représentation graphique d'une fonction f de deux variables est une surface de niveau. En effet, si on pose $g(x, y, z) = z - f(x, y)$ alors

$$g(x, y, z) = 0 \Leftrightarrow z = f(x, y);$$

dès lors la représentation graphique de f est une surface de niveau de $g(x, y, z) = z - f(x, y)$.

2.1.5 Surfaces quadriques

Dans l'espace, les surfaces de niveau d'une fonction polynomiale du second degré (en toutes les variables)

$$f(x, y, z) = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + a_{12}xy + a_{13}xz + a_{23}yz + b_1x + b_2y + b_3z + c$$

où les a_{ij}, b_i, c sont des réels tels que les a_{ij} ne soient pas tous nuls s'appellent *surfaces quadriques*.

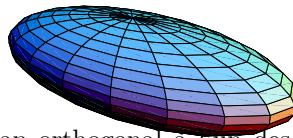
Tout comme dans le cas des coniques du plan, on montre qu'un changement de repère adéquat permet de ramener l'équation cartésienne $f(x, y, z) = 0$ à une forme canonique (en fait il s'agit de changer de repère de telle sorte que la matrice qui représente la quadrique ait une forme diagonale ; dans ce cas, dans l'équation cartésienne obtenue, il n'y a plus de terme du second degré faisant intervenir deux variables distinctes). Les différents types d'équations canoniques figurent dans la liste exhaustive suivante. On classe alors les quadriques en quadriques du type elliptique, hyperbolique, parabolique ; si l'on excepte les cas où la quadrique dégénère en plans, il existe 9 types de quadriques.

Une aide à la représentation d'une quadrique consiste à considérer les traces de cette surface dans des plans orthogonaux aux axes (c'est-à-dire les intersections de la surface avec des plans orthogonaux aux axes)¹.

- Ellipsoïde : surface d'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

où a, b, c sont des réels strictement positifs.



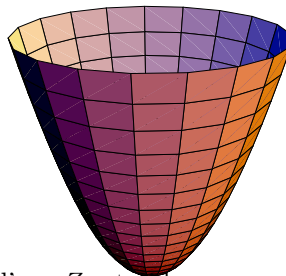
La trace de cette surface sur un plan orthogonal à l'un des axes de coordonnées est vide ou est une ellipse. Par exemple, si $z = r$ avec $r \in]-c, c[$ alors la trace correspondante a pour équation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 - \frac{r^2}{c^2}$; ceci est bien l'équation d'une ellipse.

Lorsque $a = b = c$, l'ellipsoïde est une sphère de rayon a (ensemble des points qui sont situés à une distance a de l'origine).

- Paraboloïde elliptique : surface d'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = pz$$

où $a > 0, b > 0, p \in \mathbb{R}_0$.



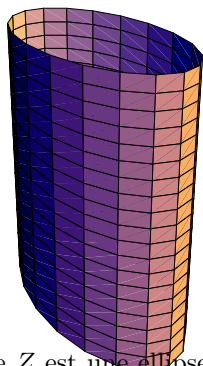
La trace sur un plan orthogonal à l'axe Z est vide ou une ellipse. La trace sur un plan orthogonal à l'axe X (ou Y) est une parabole.

- Cylindre elliptique : surface d'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

où a, b sont des réels strictement positifs.

1. Il s'agit en fait aussi de courbes de niveaux (mais les différents niveaux ne se prennent plus uniquement selon Z , mais aussi selon X, Y).

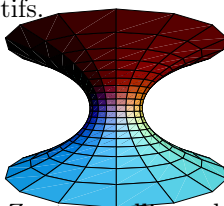


La trace sur un plan orthogonal à l'axe Z est une ellipse. La trace sur un plan orthogonal à l'axe X (ou Y) est vide ou formée de deux droites parallèles.

- Hyperboloïde à une nappe : surface d'équation cartésienne

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

où a, b, c sont des réels strictement positifs.

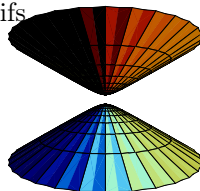


La trace sur un plan orthogonal à l'axe Z est une ellipse ; la trace sur un plan orthogonal à l'axe X (ou à l'axe Y) est une hyperbole.

- Hyperboloïde à deux nappes : surface d'équation cartésienne

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$$

où a, b, c sont des réels strictement positifs.

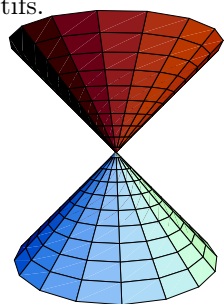


La trace sur un plan orthogonal à l'axe Z est vide ou une ellipse ; la trace sur un plan orthogonal à l'axe X (ou à l'axe Y) est une hyperbole.

- Cône : surface d'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$$

où a, b, c sont des réels strictement positifs.

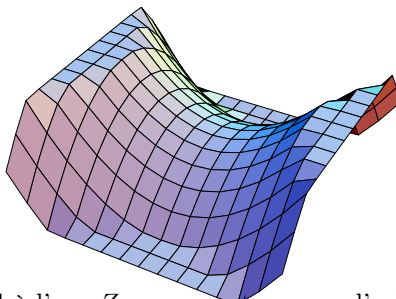


La trace sur un plan orthogonal à l'axe Z ne passant pas par l'origine est une ellipse ; la trace sur le plan d'équation $z = 0$ est l'origine. La trace sur un plan orthogonal à l'axe X (ou Y) est formée par deux droites sécantes.

- Paraboloïde hyperbolique : surface d'équation cartésienne

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2pz$$

où $a, b > 0$ et $p \in \mathbb{R}_0$.

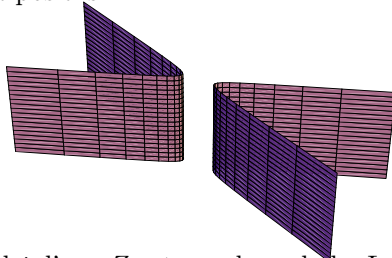


La trace sur un plan orthogonal à l'axe Z ne passant pas par l'origine est une hyperbole ; la trace sur le plan d'équation $z = 0$ est l'union de deux droites sécantes. La trace sur un plan orthogonal à l'axe X (ou Y) est une parabole.

- Cylindre hyperbolique : surface d'équation cartésienne

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

où a, b sont des réels strictement positifs.

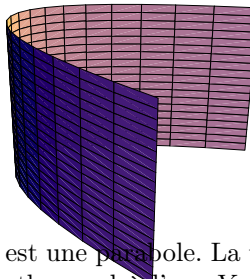


La trace sur un plan orthogonal à l'axe Z est une hyperbole. La trace sur un plan orthogonal à l'axe X (ou Y) est vide ou formée de deux droites parallèles à l'axe Z .

- Cylindre parabolique : surface d'équation cartésienne

$$x^2 = py$$

où p est un réel non nul.



La trace sur un plan orthogonal à l'axe Z est une parabole. La trace sur un plan orthogonal à l'axe X est une droite. La trace sur un plan orthogonal à l'axe Y est vide ou formée de deux droites parallèles à l'axe Z .

Remarque sur les termes “elliptique, hyperbolique, parabolique” : le terme elliptique (resp. hyperbolique) est employé lorsque, dans l'équation canonique, les coefficients devant les termes du second degré sont de même signe (resp. de signe différents). Le terme parabolique est utilisé lorsque, dans l'équation canonique, subsiste un terme du premier degré. Le terme cylindrique est utilisé lorsque, dans l'équation canonique, une variable est manquante.

2.1.6 Opérations entre fonctions

Comme dans le cadre des fonctions d'une variable réelle, on définit la somme, le produit, le quotient de deux fonctions de plusieurs variables à valeurs réelles (ou complexes).

Quant à la composition de fonctions (terme plus général que “fonction de fonction”), elle se définit aussi de manière analogue. Accordons une attention particulière à ce point car il est important lorsque l'on parle de “dérivée partielle, dérivée totale, règle de dérivation en chaîne (‘chain rule’)”. Les cours de sciences font abondamment appel à ce genre de technique.

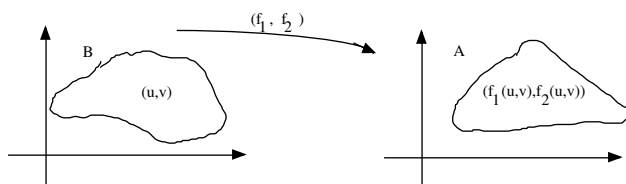
Par exemple, si f est de domaine $A \subset \mathbb{R}^2$, si f_1, f_2 sont à valeurs réelles et définies dans $B \subset \mathbb{R}^2$ et si

$$\{(f_1(u, v), f_2(u, v)) : (u, v) \in B\} \subset A$$

alors la fonction composée donnée par

$$F(u, v) = f(f_1(u, v), f_2(u, v))$$

est définie dans B .



De même si f est de domaine $A \subset \mathbb{R}^3$, si f_1, f_2, f_3 sont définies dans $B \subset \mathbb{R}$ et si

$$\{(f_1(t), f_2(t), f_3(t)) : t \in B\} \subset A$$

alors la fonction donnée par

$$F(t) = f(f_1(t), f_2(t), f_3(t))$$

est définie dans B .

2.2 Limites et continuité

2.2.1 Limites

DÉFINITIONS ET PROPRIÉTÉS

La notion de limite s'introduit de manière analogue au cas d'une variable. Ici, nous présenterons seulement le cas d'une limite finie en un point de $\mathbb{R}^2$ ou de $\mathbb{R}^3$. La définition d'une limite en l'infini s'introduit de manière analogue.

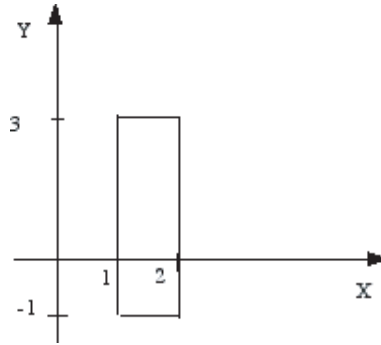
Considérons le cas de deux variables. Le même développement peut être fait dans le cas de trois variables.

Un rectangle borné de $\mathbb{R}^2$ est un sous-ensemble de $\mathbb{R}^2$ qui est le produit (on parle de produit cartésien) de deux intervalles bornés I, J de $\mathbb{R}$. On le désigne par

$$I \times J.$$

Ainsi, par exemple, si $I = [1, 2]$, $J = [-1, 3]$, on a

$$I \times J = [1, 2] \times [-1, 3] = \{(x, y) \in \mathbb{R}^2 : x \in [1, 2], y \in [-1, 3]\}.$$



Plus généralement, un rectangle de $\mathbb{R}^2$ est un sous-ensemble de $\mathbb{R}^2$ qui est le produit de deux intervalles I, J de $\mathbb{R}$.

Le rectangle est dit *ouvert* (resp. *fermé*) s'il est le produit de deux intervalles ouverts (resp. fermés) de $\mathbb{R}$.

Hypothèse de travail. Dans ce qui suit, lorsqu'on envisagera la limite des valeurs de f en (x_0, y_0) , on supposera toujours que le point (x_0, y_0) est tel que tout rectangle ouvert le contenant est d'intersection non vide avec le domaine de définition de f .

Comme dans le cas d'une variable, nous adoptons la définition des limites par les suites.

Définition 2.2.1 Soit $A \subset \mathbb{R}^2$ le domaine de définition de f .

La fonction f admet une limite finie en (x_0, y_0) s'il existe $L \in \mathbb{R}$ tel que, pour toutes suites x_m ($m \in \mathbb{N}_0$), y_m ($m \in \mathbb{N}_0$) qui convergent respectivement vers x_0, y_0 et telles que $(x_m, y_m) \in A$ pour tout m , la suite de réels $f(x_m, y_m)$ ($m \in \mathbb{N}_0$) converge vers L .

La notation employée est alors

$$\lim_{(x,y) \rightarrow (x_0, y_0)} f(x, y) = L.$$

La fonction f admet une limite infinie en (x_0, y_0) si pour toutes suites x_m ($m \in \mathbb{N}_0$), y_m ($m \in \mathbb{N}_0$) qui convergent respectivement vers x_0, y_0 et telles que $(x_m, y_m) \in A$ pour tout m , la suite de réels $f(x_m, y_m)$ ($m \in \mathbb{N}_0$) converge vers l'infini.

La notation employée est alors

$$\lim_{(x,y) \rightarrow (x_0, y_0)} f(x, y) = \infty.$$

Comme dans le cas d'une variable, on a une propriété d'unicité de la limite. On a aussi la propriété suivante, dite "en ε, η ", donnant en fait une définition équivalente de la notion de limite des valeurs d'une fonction.

Propriété 2.2.2 Soit $A \subset \mathbb{R}^2$ le domaine de définition de f .

On a

$$\lim_{(x,y) \rightarrow (x_0, y_0)} f(x, y) = L$$

si et seulement si

$$\forall \varepsilon > 0, \exists \eta > 0 : \left\{ \frac{(x, y) \in A}{\sqrt{(x - x_0)^2 + (y - y_0)^2} \leq \eta} \right\} \Rightarrow |f(x, y) - L| \leq \varepsilon.$$

On a

$$\lim_{(x,y) \rightarrow (x_0, y_0)} f(x, y) = \infty$$

si et seulement si

$$\forall \varepsilon > 0, \exists \eta > 0 : \left\{ \frac{(x, y) \in A}{\sqrt{(x - x_0)^2 + (y - y_0)^2} \leq \eta} \right\} \Rightarrow |f(x, y)| \geq \varepsilon.$$

Dans ce qui précède et de manière analogue à ce qui se définit pour les fonctions d'une variable, on peut préciser parfois que la limite est L^+ ou L^- ; de même $+\infty$ ou $-\infty$.

Les propriétés concernant les combinaisons linéaires, produits, quotients, fonctions composées, sont analogues à celles rencontrées dans le cadre des fonctions d'une variable réelle, de même que les propriétés concernant les inégalités (notamment le théorème de l'étau).

EXEMPLES

1) On a

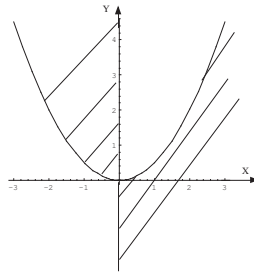
$$\lim_{(x,y) \rightarrow (-1,2)} \frac{\ln(x^2 + y^2)}{\sqrt{x^3 - 2xy}} = \frac{\ln 5}{\sqrt{3}}, \quad \lim_{(x,y) \rightarrow (-1,1/2)} \frac{\ln(x^2 + y^2)}{\sqrt{x^3 - 2xy}} = +\infty,$$

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\ln(x^2 + y^2)}{\sqrt{x^3 - 2xy}} = -\infty.$$

Solution. La fonction

$$f(x, y) = \frac{\ln(x^2 + y^2)}{\sqrt{x^3 - 2xy}}$$

est définie sur $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 > 0, x^3 - 2xy > 0\}$ qui est l'ensemble hachuré ci-dessous (dont on ne prend pas les bords).



Le point du plan de coordonnées $(-1, 2)$ appartient à cet ensemble, donc tout voisinage de ce point rencontre A ; de même le point du plan de coordonnées $(-1, 1/2)$ (*resp.* $(0, 0)$) n'appartient pas à A mais est tel que tout voisinage de celui-ci rencontre A également. Les questions posées ont donc un sens.

On a $\lim_{(x,y) \rightarrow (-1,2)} x^2 = (-1)^2 = 1$, $\lim_{(x,y) \rightarrow (-1,2)} y^2 = 2^2 = 4$ donc

$$\lim_{(x,y) \rightarrow (-1,2)} \ln(x^2 + y^2) = \ln(1 + 4) = \ln 5.$$

De même on a

$$\lim_{(x,y) \rightarrow (-1,2)} \sqrt{x^3 - 2xy} = \sqrt{(-1)^3 - 2 \cdot (-1) \cdot 2} = \sqrt{3} \neq 0.$$

Dès lors

$$\lim_{(x,y) \rightarrow (-1,2)} \frac{\ln(x^2 + y^2)}{\sqrt{x^3 - 2xy}} = \frac{\lim_{(x,y) \rightarrow (-1,2)} \ln(x^2 + y^2)}{\lim_{(x,y) \rightarrow (-1,2)} \sqrt{x^3 - 2xy}} = \frac{\ln 5}{\sqrt{3}}.$$

On a $\lim_{(x,y) \rightarrow (-1,1/2)} x^2 = (-1)^2 = 1$, $\lim_{(x,y) \rightarrow (-1,1/2)} y^2 = (1/2)^2 = 1/4$ donc

$$\lim_{(x,y) \rightarrow (-1,1/2)} \ln(x^2 + y^2) = \ln(1 + 1/4) = \ln(5/4) > 0.$$

De même on a

$$\lim_{(x,y) \rightarrow (-1,1/2)} \sqrt{x^3 - 2xy} = \sqrt{(-1)^3 - 2 \cdot (-1) \cdot (1/2)} = 0.$$

Dès lors

$$\lim_{(x,y) \rightarrow (-1,1/2)} \frac{\ln(x^2 + y^2)}{\sqrt{x^3 - 2xy}} = \frac{\ln(5/4)}{0^+} = +\infty.$$

On a $\lim_{(x,y) \rightarrow (0,0)} x^2 = 0$, $\lim_{(x,y) \rightarrow (0,0)} y^2 = 0$ donc

$$\lim_{(x,y) \rightarrow (0,0)} \ln(x^2 + y^2) = \lim_{z \rightarrow 0^+} \ln z = -\infty.$$

De même on a

$$\lim_{(x,y) \rightarrow (0,0)} \sqrt{x^3 - 2xy} = 0.$$

Dès lors

$$\lim_{(x,y) \rightarrow (-1,2)} \frac{\ln(x^2 + y^2)}{\sqrt{x^3 - 2xy}} = \frac{-\infty}{0^+} = -\infty.$$

2) On a

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^3 + y^3}{x^2 + y^2} = 0.$$

Solution. Le domaine de définition de la fonction

$$f(x, y) = \frac{x^3 + y^3}{x^2 + y^2}$$

est le complémentaire de l'origine. La question posée a donc un sens.

On a $\lim_{(x,y) \rightarrow (0,0)} (x^3 + y^3) = 0 + 0 = 0$ et $\lim_{(x,y) \rightarrow (0,0)} (x^2 + y^2) = 0 + 0 = 0$ si bien que l'on se retrouve avec le cas

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^3 + y^3}{x^2 + y^2} = \frac{0}{0}.$$

Les propriétés standards pour trouver cette limite ne s'appliquent plus. On doit procéder autrement.

Pour tout $x, y \in \mathbb{R}$ et $x \neq 0, y \neq 0$, on a

$$\left| \frac{x^3 + y^3}{x^2 + y^2} \right| = \left| \frac{x^3}{x^2 + y^2} + \frac{y^3}{x^2 + y^2} \right| \leq \frac{|x|^3}{x^2} + \frac{|y|^3}{y^2} \leq |x| + |y|$$

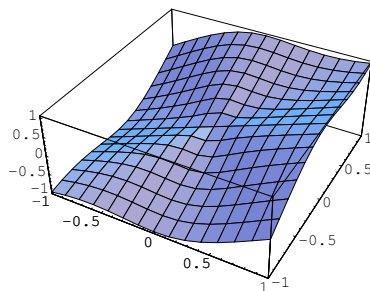
donc

$$0 \leq |f(x, y)| = \left| \frac{x^3 + y^3}{x^2 + y^2} \right| \leq |x| + |y|, \quad \forall (x, y) \in \mathbb{R}^2, (x, y) \neq (0, 0).$$

Comme

$$\lim_{(x,y) \rightarrow (0,0)} (|x| + |y|) = 0$$

le théorème de l'étau fournit la conclusion annoncée.



Représentation graphique de $f(x, y) = \frac{x^3 + y^3}{x^2 + y^2}$.

3) La limite

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 - y^2}{x^2 + y^2}$$

n'existe pas.

Solution. Le domaine de définition de la fonction

$$f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$$

est le complémentaire de l'origine. La question posée a donc un sens.

On a $\lim_{(x,y) \rightarrow (0,0)} (x^2 - y^2) = 0$ et $\lim_{(x,y) \rightarrow (0,0)} (x^2 + y^2) = 0 + 0 = 0$ si bien que l'on se retrouve avec le cas

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 - y^2}{x^2 + y^2} = \frac{0}{0}.$$

Les propriétés standards pour trouver cette limite ne s'appliquent plus. On doit procéder à l'aide de la définition.

On remarque alors que, si $y \neq 0$, on a

$$f(0, y) = \frac{-y^2}{y^2} = -1$$

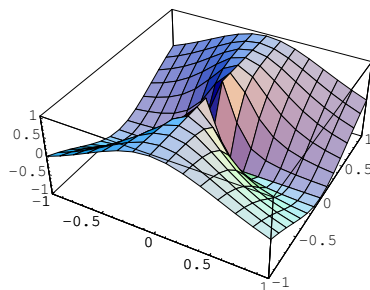
et si $x \neq 0$

$$f(x, 0) = \frac{x^2}{x^2} = 1$$

Dès lors,

$$\lim_{m \rightarrow +\infty} f(0, \frac{1}{m}) = -1, \quad \lim_{m \rightarrow +\infty} f(\frac{1}{m}, 0) = 1$$

ce qui permet de conclure.



Représentation graphique de $f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$.

2.2.2 Continuité

Définition et propriétés

La continuité s'introduit de la même manière que dans le cas des fonctions d'une variable. Les propriétés sont aussi analogues (opérations entre fonctions : combinaisons linéaires, produits, quotients, fonctions composées).

Définition 2.2.3 Soit f une fonction définie sur une partie A de $\mathbb{R}^2$ et soit $(x_0, y_0) \in A$. On dit que la fonction est continue au point (x_0, y_0) si

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) \text{ existe.}$$

Dans ce cas, elle vaut nécessairement $f(x_0, y_0)$.

On dit que f est continu sur A si cette fonction est continue en tout point de A .

Si f est une fonction définie sur A , l'ensemble des points de A où elle est continue s'appelle le domaine de continuité de f .

EXEMPLES

Déterminer le domaine de définition et de continuité des fonctions données explicitement ci-dessous. Représenter ces domaines.

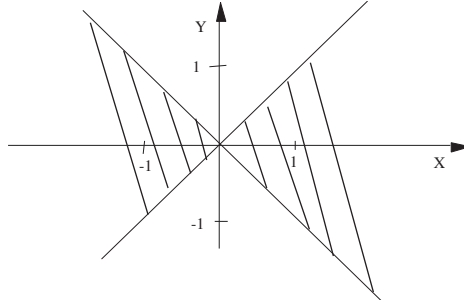
- $f_1(x, y) = \sin(xy)$
- $f_2(x, y) = \ln(x^2 + y^2)$
- $f_3(x, y) = e^{\frac{1}{x+y}}$
- $f_4(x, y) = \frac{1}{\sqrt{x^2 - y^2}}$
- $f_5(x) = \begin{cases} \frac{y^2 - x^2}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$

Suggestion. La fonction f_1 est définie et continue sur $\mathbb{R}^2$.

La fonction f_2 est définie et continue sur $\mathbb{R}^2 \setminus \{(0, 0)\}$.

La fonction f_3 est définie et continue sur $\{(x, y) \in \mathbb{R}^2 : x + y \neq 0\}$ c'est-à-dire sur le complémentaire de la droite d'équation $x + y = 0$ (qui est la seconde bissectrice des axes).

La fonction f_4 est définie et continue sur $\{(x, y) \in \mathbb{R}^2 : x^2 - y^2 > 0\}$; la représentation graphique de cet ensemble est la partie hachurée suivante (les droites d'équation $x - y = 0$ et $x + y = 0$ ne font pas partie de l'ensemble).



La fonction f_5 est définie sur $\mathbb{R}^2$ et continue sur $\mathbb{R}^2 \setminus \{(0,0)\}$

2.3 Dérivation

2.3.1 Définitions des dérivées partielles et premières propriétés

DÉFINITIONS

Comme dans le cas des fonctions d'une variable, on travaille dans des ensembles particuliers, appelés ouverts : une partie A de $\mathbb{R}^2$ est appelée un ouvert si tout point de A est le centre d'un rectangle qui reste inclus dans A .

Définition 2.3.1 Soit f une fonction définie sur un ouvert A de $\mathbb{R}^2$ et soit $(x_0, y_0) \in A$.

a) La fonction f est dérivable par rapport à sa première variable x en (x_0, y_0) si la limite

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} = \lim_{x \rightarrow x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0}$$

existe et est finie. Dans ce cas, cette limite est appelée dérivée partielle de f par rapport à x au point (x_0, y_0) et est notée

$$D_x f(x_0, y_0) \quad \text{ou encore} \quad \frac{\partial f}{\partial x}(x_0, y_0).$$

b) La fonction f est dérivable par rapport à sa deuxième variable y en (x_0, y_0) si la limite

$$\lim_{h \rightarrow 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h} = \lim_{y \rightarrow y_0} \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0}$$

existe et est finie. Dans ce cas, cette limite est appelée dérivée partielle de f par rapport à y au point (x_0, y_0) et est notée

$$D_y f(x_0, y_0) \quad \text{ou encore} \quad \frac{\partial f}{\partial y}(x_0, y_0).$$

c) la fonction f est dérivable sur (ou dans) A si elle est dérivable par rapport à x et par rapport à y en tout point de A .

Définition 2.3.2 Soit f défini sur $A \subset \mathbb{R}^2$. Le domaine de dérivabilité de f est le plus grand ouvert inclus dans A tel que f soit dérivable en tout point de cet ouvert.

Si A est un ouvert de $\mathbb{R}^2$, la notation

$$C_1(A)$$

désigne l'ensemble des fonctions dérivables dans A dont les dérivées partielles sont continues dans A . Un élément de cet ensemble s'appelle une fonction continûment dérivable dans (sur) A .

PROPRIÉTÉS

Les propriétés relatives aux combinaisons linéaires, produits, quotients, sont analogues à celles rencontrées dans le cadre des fonctions d'une variable réelle. La propriété concernant la dérivabilité et l'expression des dérivées partielles des fonctions composées est un peu plus complexe. Nous admettrons ce résultat sans démonstration.

Proposition 2.3.3 a) Soient

- f une fonction de deux variables réelles continûment dérivable sur un ouvert U de $\mathbb{R}^2$,
- f_1, f_2 deux fonctions de deux variables réelles dérivables sur l'ouvert Ω de $\mathbb{R}^2$ et telles que

$$\{(f_1(x, y), f_2(x, y)) : (x, y) \in \Omega\} \subset U.$$

Alors la fonction définie par

$$F(x, y) = f(f_1(x, y), f_2(x, y))$$

est dérivable sur Ω et on a (*)

$$D_x F = (D_1 f)_{(f_1, f_2)} D_x f_1 + (D_2 f)_{(f_1, f_2)} D_x f_2, \quad D_y F = (D_1 f)_{(f_1, f_2)} D_y f_1 + (D_2 f)_{(f_1, f_2)} D_y f_2$$

b) Soient

- f une fonction d'une variable réelle dérivable sur un ouvert I de $\mathbb{R}$,
- f_1 une fonction de deux variables réelles dérivable sur l'ouvert Ω de $\mathbb{R}^2$ et telles que

$$\{f_1(x, y) : (x, y) \in \Omega\} \subset I.$$

Alors la fonction définie par

$$F(x, y) = f(f_1(x, y))$$

est dérivable sur Ω et on a

$$D_x F = (Df)_{f_1} D_x f_1, \quad D_y F = (Df)_{f_1} D_y f_1$$

c) Soient

- f une fonction de trois variables réelles continûment dérivable sur un ouvert U de $\mathbb{R}^3$,
- f_1, f_2, f_3 trois fonctions d'une variable réelle, dérivables sur l'ouvert I de $\mathbb{R}$ et telles que

$$\{(f_1(t), f_2(t), f_3(t)) : t \in I\} \subset U.$$

Alors la fonction définie par

$$F(t) = f(f_1(t), f_2(t), f_3(t))$$

est dérivable sur I et on a

$$DF = (D_x f)_{(f_1, f_2, f_3)} Df_1 + (D_y f)_{(f_1, f_2, f_3)} Df_2 + (D_z f)_{(f_1, f_2, f_3)} Df_3$$

d) Si, dans les énoncés ci-dessus, les fonctions sont p fois continûment dérivables, alors la fonction composée est aussi p fois continûment dérivable.

Les relations (*) (et les autres du même type) doivent bien sûr être prises au point de coordonnées (x, y) . Par exemple, avec précision, la première s'écrit en fait

$$D_x F(x, y) = (D_u f)_{(f_1(x, y), f_2(x, y))} D_x f_1(x, y) + (D_v f)_{(f_1(x, y), f_2(x, y))} D_x f_2(x, y).$$

Le résultat ci-dessus s'étend de manière naturelle à d'autres cas, par exemple $f(u, v)$ et $f_1(x, y, z), f_2(x, y, z)$ (qui donne $F(x, y, z) = f(f_1(x, y, z), f_2(x, y, z))$), $f(t)$ et $f_1(x, y)$ (qui donne $F(x, y) = f(f_1(x, y))$), ...

Il faut aussi remarquer la différence d'hypothèse. En effet, lorsque la fonction f est une fonction d'une variable réelle, on obtient un résultat beaucoup plus fort : la dérivabilité de la fonction composée est obtenue en supposant f seulement dérivable alors qu'on suppose f continûment dérivable dans le cas où il s'agit d'une fonction de plusieurs variables. Ceci n'est pas une faiblesse de la démonstration comme le montrent les exemples suivants.

EXEMPLE 1

- La fonction

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

est définie et dérivable sur $\mathbb{R}^2$ mais n'est pas continûment dérivable sur $\mathbb{R}^2$.

En effet, en un point $(x, y) \neq (0, 0)$ la fonction est dérivable et on a

$$D_x f(x, y) = \frac{y(x^2 + y^2) - (xy) \cdot 2x}{(x^2 + y^2)^2} = \frac{y(y^2 - x^2)}{(x^2 + y^2)^2}, \quad D_y f(x, y) = \frac{x(x^2 - y^2)}{(x^2 + y^2)^2}.$$

La fonction est aussi dérivable en $(0, 0)$ car

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{f(0 + h, 0) - f(0, 0)}{h} &= \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0 \\ \lim_{h \rightarrow 0} \frac{f(0, 0 + h) - f(0, 0)}{h} &= \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0. \end{aligned}$$

Dès lors

$$D_x f(x, y) = \begin{cases} \frac{y(y^2 - x^2)}{(x^2 + y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

et

$$D_y f(x, y) = \begin{cases} \frac{x(x^2 - y^2)}{(x^2 + y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0). \end{cases}$$

La dérivée partielle $D_x f(x, y)$ (de même pour celle par rapport à y) n'est pas continue en $(0, 0)$. En effet, on a

$$D_x f(0, \frac{1}{m}) = m \quad \forall m \in \mathbb{N}_0;$$

comme la suite $D_x f(0, \frac{1}{m})$ ($m \in \mathbb{N}_0$) ne converge pas vers $D_x f(0, 0) = 0$, on conclut par le critère liant la limite des valeurs d'une fonction et la convergence de suites.

- Cela étant, on considère les fonctions

$$f_1(t) = t, \quad f_2(t) = t\sqrt{|t|}.$$

Ces fonctions sont dérivables sur $\mathbb{R}$. On a

$$F(t) = f(f_1(t), f_2(t)) = \frac{\sqrt{|t|}}{1 + |t|}, \quad t \in \mathbb{R}.$$

Cette fonction n'est pas dérivable en 0 car

$$\lim_{t \rightarrow 0^+} \frac{F(t) - F(0)}{t} = \lim_{t \rightarrow 0^+} \frac{\sqrt{t}}{(1+t)t} = +\infty.$$

EXEMPLE 2

De même, il se peut que la fonction composée soit dérivable mais qu'on ne puisse pas utiliser la « formule » de dérivation des fonctions composées pour trouver la dérivée. L'exemple suivant illustre ce résultat.

La fonction

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

est définie et dérivable sur $\mathbb{R}^2$ mais n'est pas continûment dérivable sur $\mathbb{R}^2$. La composée de f et de $f_1(t) = f_2(t) = t$ fournit une fonction F qui est dérivable sur $\mathbb{R}$ mais dont la dérivée en 0 ne s'obtient pas via la dérivation des fonctions composées.

REMARQUE À PROPOS DE LA NOTION DE “DÉRIVÉE TOTALE”

Supposons que f soit fonction de deux ou trois variables et que ces variables évoluent au cours du temps (par exemple). Sous de faibles hypothèses, la nouvelle fonction du temps ainsi construite, fonction d’une seule variable réelle, est dérivable par rapport au temps. Sa dérivée est souvent appelée *dérivée totale* de f et on rencontre la formule

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

Il ne s’agit pas de quelque chose de nouveau : il s’agit en fait de la dérivée de la fonction composée

$$F(t) = f(x(t), y(t), z(t))$$

utilisant le résultat (2.3.3) et les notations de dérivation utilisées dans certaines présentations de cours de sciences.

EXEMPLES

Déterminer l’ensemble où les fonctions suivantes sont dérivables ; représenter ces ensembles. Calculer ensuite leurs dérivées partielles.

$$f_1(x, y) = \sin(xy), \quad f_2(x, y) = \exp(\sqrt{x^2 + y^2}), \quad f_3(x, y) = \frac{1}{x^2 - y^2},$$

$$f_4(x, y) = \ln(\sqrt{x^2 + y^2}), \quad f_5(x, y) = \arctg \frac{x}{y}, \quad f_6(x, y) = f(g_1(x, y), g_2(x, y))$$

avec

$$g_1(x, y) = x^2 - y^2, \quad g_2(x, y) = x - y, \quad f \text{ continûment dérivable dans } U = \{(u, v) \in \mathbb{R}^2 : |u| \neq |v|\}.$$

Suggestion. $\boxed{f_1}$ On a $f_1(x, y) = \sin(g(x, y))$ avec $g(x, y) = xy$.

Comme g est dérivable sur $\Omega = \mathbb{R}^2$ et comme la fonction $\sin$ est dérivable sur $U = \mathbb{R}$, la fonction $f_1(x, y) = \sin(g(x, y))$ est dérivable sur

$$\{(x, y) \in \Omega = \mathbb{R}^2 : xy \in U = \mathbb{R}\} = \mathbb{R}^2$$

et on a

$$D_x \sin(xy) = (D \sin)(xy) \cdot D_x(g(x, y)) = y \cos(xy), \quad D_y \sin(xy) = (D \sin)(xy) \cdot D_y(g(x, y)) = x \cos(xy).$$

$\boxed{f_2}$ On a $f_2(x, y) = \exp g(x, y)$ avec $g(x, y) = \sqrt{x^2 + y^2}$.

La fonction g s’écrit $g(x, y) = \sqrt{h(x, y)}$ avec $h(x, y) = x^2 + y^2$. Comme h est dérivable sur $A = \mathbb{R}^2$ et comme la fonction $H(X) = \sqrt{X}$ est dérivable sur $]0, +\infty[$, la fonction $g(x, y) = \sqrt{h(x, y)}$ est dérivable sur

$$\{(x, y) \in A = \mathbb{R}^2 : h(x, y) = x^2 + y^2 \in]0, +\infty[\} = \mathbb{R}^2 \setminus \{(0, 0)\}.$$

Cela étant, comme la fonction $\exp$ est continûment dérivable sur $U = \mathbb{R}$, la fonction

$$f_2(x, y) = \exp(g(x, y))$$

est dérivable sur

$$\{(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\} : \sqrt{x^2 + y^2} \in U = \mathbb{R}\} = \mathbb{R}^2 \setminus \{(0, 0)\}$$

et on a

$$\begin{aligned} D_x \exp(\sqrt{x^2 + y^2}) &= (D \exp)(g(x, y)) \cdot D_x g(x, y) = \frac{x \exp(\sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}} \\ D_y \exp(\sqrt{x^2 + y^2}) &= (D \exp)(g(x, y)) \cdot D_y g(x, y) = \frac{y \exp(\sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}}. \end{aligned}$$

$\boxed{f_3}$ On a $f_3(x, y) = H((g(x, y)))$ avec $g(x, y) = x^2 - y^2$ et $H(X) = \frac{1}{X}$.

Comme la fonction g est dérivable sur $\mathbb{R}^2$ et comme la fonction H est dérivable sur $\mathbb{R} \setminus \{0\}$, la fonction f_3 est dérivable sur

$$\{(x, y) \in \mathbb{R}^2 : g(x, y) \neq 0\} = \{(x, y) \in \mathbb{R}^2 : (x - y)(x + y) \neq 0\}$$

c'est-à-dire sur le complémentaire de l'union des deux bissectrices des axes du repère. De plus,

$$D_x f_3(x, y) = -\frac{2x}{(x^2 - y^2)^2}, \quad D_y f_3(x, y) = \frac{2y}{(x^2 - y^2)^2}.$$

$\boxed{f_4}$ Si $x^2 + y^2 > 0$, on a $\ln(\sqrt{x^2 + y^2}) = \frac{1}{2} \ln(x^2 + y^2)$. Dès lors, $f_4(x, y) = H(g(x, y))$ avec $H(X) = \frac{1}{2} \ln X$, $g(x, y) = x^2 + y^2$. Comme la fonction H est dérivable sur $]0, +\infty[$ et comme la fonction g est dérivable sur $\mathbb{R}^2$, la fonction f_4 est dérivable sur

$$\{(x, y) \in \mathbb{R}^2 : g(x, y) \in]0, +\infty[\} = \mathbb{R}^2 \setminus \{(0, 0)\}$$

et on a

$$D_x f_4(x, y) = \frac{x}{x^2 + y^2}, \quad D_y f_4(x, y) = \frac{y}{x^2 + y^2}.$$

$\boxed{f_5}$ On a $f_5(x, y) = H(g(x, y))$ avec $H(X) = \arctg(X)$ et $g(x, y) = \frac{x}{y}$. Comme la fonction H est dérivable sur $\mathbb{R}$ et comme la fonction g est dérivable sur $\{(x, y) \in \mathbb{R}^2 : y \neq 0\}$, la fonction f_5 est dérivable sur

$$\{(x, y) \in \mathbb{R}^2 : y \neq 0\}$$

et on a

$$D_x f_5(x, y) = (DH(X))_{X=g(x, y)} \cdot D_x g(x, y) = \frac{1}{1 + \frac{x^2}{y^2}} \cdot \frac{1}{y} = \frac{y}{x^2 + y^2}$$

$$D_y f_5(x, y) = (DH(X))_{X=g(x, y)} \cdot D_y g(x, y) = \frac{1}{1 + \frac{x^2}{y^2}} \cdot \frac{-x}{y^2} = \frac{-x}{x^2 + y^2}.$$

$\boxed{f_6}$ Comme les fonctions g_1 et g_2 sont dérivables dans $\mathbb{R}^2$ et comme la fonction f est continûment dérivable sur U , la fonction composée $f_6 = f(g_1, g_2)$ est dérivable sur $\Omega = \{(x, y) \in \mathbb{R}^2 : |g_1(x, y)| \neq |g_2(x, y)|\}$ et on a

$$\begin{aligned} D_x f_6(x, y) &= (D_u f)(x^2 - y^2, x - y) \cdot D_x g_1(x, y) + (D_v f)(x^2 - y^2, x - y) \cdot D_x g_2(x, y) \\ &= 2x(D_u f)(x^2 - y^2, x - y) + (D_v f)(x^2 - y^2, x - y) \\ D_y f_6(x, y) &= (D_u f)(x^2 - y^2, x - y) \cdot D_y g_1(x, y) + (D_v f)(x^2 - y^2, x - y) \cdot D_y g_2(x, y) \\ &= -2y(D_u f)(x^2 - y^2, x - y) - (D_v f)(x^2 - y^2, x - y). \end{aligned}$$

Explicitons l'ensemble Ω : on a successivement

$$\begin{aligned} |g_1(x, y)| = |g_2(x, y)| &\Leftrightarrow |x - y| |x + y| = |x - y| \Leftrightarrow |x - y| = 0 \text{ ou } |x + y| = 1 \\ &\Leftrightarrow x = y \text{ ou } (x + y)^2 = 1 \\ &\Leftrightarrow x = y \text{ ou } (x + y - 1)(x + y + 1) = 0 \\ &\Leftrightarrow x = y \text{ ou } x + y - 1 = 0 \text{ ou } x + y + 1 = 0. \end{aligned}$$

Le complémentaire de Ω est donc l'union de trois droites ; dès lors, Ω est le complémentaire de l'union de ces trois droites (respectivement d'équation $x - y = 0$, $x + y - 1 = 0$, $x + y + 1 = 0$).

2.3.2 Lien entre dérivabilité et continuité

Dans le cas des fonctions d'une variable réelle, on a démontré qu'une fonction dérivable est toujours continue. Ce résultat n'est plus vrai dans le cas des fonctions de plusieurs variables réelles comme le montre l'exemple suivant.

La fonction

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

est définie et dérivable sur $\mathbb{R}^2$ mais n'est pas continue sur $\mathbb{R}^2$.

En effet, nous avons déjà démontré plus haut que cette fonction était dérivable dans $\mathbb{R}^2$. Montrons qu'elle n'est pas continue en $(0, 0)$.

Les suites $x_m = \frac{1}{m}$, $y_m = \frac{1}{m}$ ($m \in \mathbb{N}_0$) convergent vers 0. Cependant, $f(x_m, y_m) = \frac{1}{2}$ pour tout m . La suite $f(x_m, y_m)$ ($m \in \mathbb{N}_0$) ne converge donc pas vers $f(0, 0) = 0$; dès lors la fonction f n'est pas continue en $(0, 0)$.

Cependant, il existe un résultat permettant d'obtenir la continuité à partir de la dérivabilité. Il s'énonce comme suit. Nous l'admettrons sans faire la démonstration.

Proposition 2.3.4 *Si f est continûment dérivable sur un ouvert Ω de $\mathbb{R}^2$ alors f y est continu.*

2.3.3 Dérivées multiples

Comme dans le cas des fonctions d'une variable réelle, on peut se demander si les dérivées d'ordre 1 (c'est-à-dire les dérivées partielles) sont encore dérivables. Comme on est dans le cadre de plusieurs variables, plusieurs dérivées secondes vont apparaître. Par exemple, si f est une fonction dérivable de deux variables réelles, on se pose la question de savoir si les fonctions (de deux variables réelles)

$$D_x f(x, y), \quad D_y f(x, y)$$

sont encore dérivables, par rapport à x et par rapport à y . Si c'est le cas, on introduit alors les fonctions dérivées d'ordre deux

$$D_x D_x f(x, y) = D_x^2 f(x, y), \quad D_y D_x f(x, y), \quad D_y D_y f(x, y) = D_y^2 f(x, y), \quad D_x D_y f(x, y).$$

Dans beaucoup de cas, on a $D_x D_y f(x, y) = D_y D_x f(x, y)$ mais pas toujours !

Par exemple, la fonction $f(x, y) = \begin{cases} \frac{x^3 y - x y^3}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$ est deux fois dérivable dans $\mathbb{R}^2$ mais

$$D_x D_y f(0, 0) \neq D_y D_x f(0, 0).$$

On a cependant le résultat suivant, que nous admettrons sans démonstration.

Proposition 2.3.5 *Soit f une fonction définie sur un ouvert Ω de $\mathbb{R}^2$ telle que $D_x f, D_y f, D_x D_y f$ existent dans Ω (resp. $D_x f, D_y f, D_y D_x f$ existent dans Ω) et $D_x D_y f$ est continu dans Ω . Alors la fonction $D_y D_x f$ (resp. $D_x D_y f$) existe dans Ω et on a $D_x D_y f = D_y D_x f$ dans Ω .*

On introduit de manière analogue les dérivées d'ordre p pour tout naturel p .

2.3.4 Des opérateurs de dérivation fort utiles

En physique (notamment), les opérateurs de dérivation appelés

gradient, divergence, rotationnel

apparaissent à de multiples endroits, comme outils de modélisation de divers phénomènes (flux au travers de parois, équations de Maxwell en électromagnétisme, équation de la chaleur, des ondes, ...). Il s'agit d'*opérateurs faisant intervenir des dérivées partielles* et les équations impliquant celles-ci sont des cas particuliers d'*équations aux dérivées partielles*, lesquelles sont aux fonctions de plusieurs variables ce que sont les équations différentielles aux fonctions d'une variable réelle.

A COMPLETER (voir aussi cours)

2.4 Intégration

Comme la grande majorité des cas d'applications concerne des fonctions continues et vu les difficultés techniques que l'on rencontre dans le cas de fonctions non continues, dans ce chapitre nous n'envisagerons que l'intégration de fonctions continues.

2.4.1 Intégration sur des ensembles bornés fermés

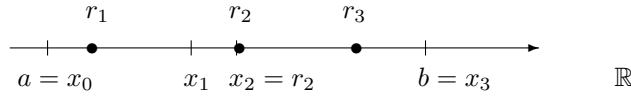
Dans le cadre des fonctions d'une variable réelle, on introduit la notion d'intégrale sur un intervalle borné fermé comme limite de sommes de Riemann obtenues à partir de découpages de l'intervalle. On étend ensuite cette notion au cas des autres intervalles.

Dans $\mathbb{R}^2, \mathbb{R}^3$, ces notions sont plus délicates à manipuler car les ensembles sur lesquels on voudrait intégrer se présentent géométriquement de manière beaucoup plus complexe.

Afin de ne pas alourdir la présentation, nous ne considérerons (dans un premier temps) que le cas de fonctions de deux variables. Un développement analogue pourrait être fait dans le cas de fonctions de plus de deux variables mais les représentations géométriques sont alors beaucoup plus lourdes à manipuler.

Intégration sur des rectangles

Pour rappel, à une variable, on considèrerait un découpage de $[a, b]$ et, dans chacun des sous-intervalles ainsi déterminés, on prenait un réel. On construisait de la sorte des “Sommes de Riemann”.



Sous-intervalles : $[a, x_1]$, $[x_1, x_2]$, $[x_2, b]$

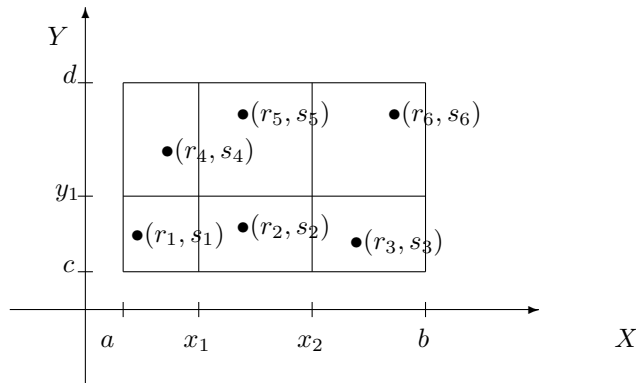
réels dans les sous-intervalles : $r_1 \in [a, x_1]$, $r_2 \in [x_1, x_2]$, $r_3 \in [x_2, b]$.

Somme de Riemann associée à ce découpage :

$$\sum_{k=1}^3 f(r_k)(x_k - x_{k-1}) = f(r_1)(x_1 - a) + f(r_2)(x_2 - x_1) + f(r_3)(b - x_2).$$

(Lorsque f est à valeurs positives, cette somme représente une somme d'aires de rectangles.) On considèrerait alors des suites de découpages dont la largeur tendait vers 0 (cf le chapitre consacré au calcul intégral à une variable).

Dans le cas de deux variables, on procède comme dans le cas d'une variable mais à partir d'un rectangle R borné fermé $[a, b] \times [c, d]$. On considère aussi des sous-rectangles, construits à partir de découpages des intervalles $[a, b]$ et $[c, d]$, et dans chacun d'eux, on considère un point.



Dans cet exemple, on a donc découpé le rectangle R en six sous-rectangles $R_1, R_2, R_3, R_4, R_5, R_6$ dans lesquels on a choisi chaque fois un point. Si on désigne par Δ_k l'aire du rectangle R_k , la somme de Riemann associée à ce découpage est

$$\sum_{k=1}^6 f(r_k, s_k) \Delta_k.$$

Lorsque la fonction f est à valeurs positives, cette somme représente une somme de volumes de parallélépipèdes.

On note d la borne supérieure des longueurs des diagonales des sous-rectangles.

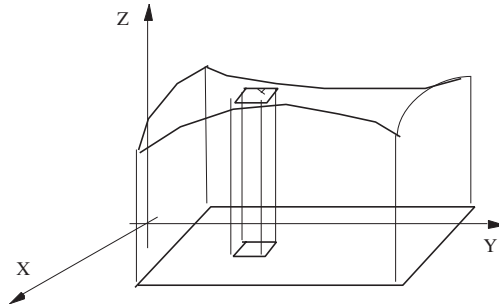
Comme dans le cas d'une variable, on considère des suites de découpages. Lorsque l'on construit une suite de découpages en rectangles telle que la suite associée d_n ($n \in \mathbb{N}_0$) converge vers 0, on examine ici encore le comportement de la suite S_n ($n \in \mathbb{N}_0$) des sommes de Riemann. **On peut démontrer que la suite S_n ($n \in \mathbb{N}_0$) converge vers une limite finie et que cette limite est indépendante de la suite de découpages qui a servi à la définir.**

On dit que f est intégrable sur R et que son intégrale sur R est la limite de la suite S_n ($n \in \mathbb{N}_0$). L'intégrale de f sur R est notée

$$\int \int_R f(x, y) \, dx dy \quad \text{ou} \quad \int \int_R f(x, y) \, dy dx.$$

Tout comme dans le cas d'une variable, vu l'interprétation géométrique de ces sommes, on définit le volume du corps compris sous la surface d'équation $z = f(x, y)$, $(x, y) \in R$ (où f est supposé à valeurs positives) par

$$V = \int \int_R f(x, y) \, dx dy.$$



Une propriété très utile pour le calcul des intégrales multiples est la suivante (elle permet de se ramener à des calculs d'intégrales d'une variable). Nous l'admettons sans démonstration.

Géométriquement, elle a une interprétation claire : le volume du corps situé "sous" la surface d'équation $z = f(x, y)$ est la superposition de l'aire des surfaces obtenues en coupant le volume par une succession de plans orthogonaux à X (ou à Y).

Proposition 2.4.1 Soit f une fonction continue sur le rectangle $A = [a, b] \times [c, d]$. On a

$$\int \int_A f(x, y) \, dx \, dy = \int_a^b \left(\int_c^d f(x, y) \, dy \right) dx$$

et

$$\int \int_A f(x, y) \, dx \, dy = \int_c^d \left(\int_a^b f(x, y) \, dx \right) dy.$$

Remarque 2.4.2 On utilise souvent la notation suivante pour les intégrales dont il est question dans la propriété précédente :

$$\int_a^b \left(\int_c^d f(x, y) \, dy \right) dx = \int_a^b dx \int_c^d f(x, y) \, dy$$

et

$$\int_c^d \left(\int_a^b f(x, y) \, dx \right) dy = \int_c^d dy \int_a^b f(x, y) \, dx$$

Intégration sur certains ensembles bornés fermés

Une définition analogue de l'intégrale d'une fonction continue sur des ensembles qui sont seulement bornés fermés peut aussi être faite. Nous nous contenterons ici de donner des résultats pratiques concernant l'intégration sur des ensembles particuliers mais qui recouvrent la plupart des cas pratiques.

Définition 2.4.3 Soit A un sous-ensemble borné fermé du plan.

a) On dit que A est parallèle à l'axe Y s'il existe deux fonctions f_1, f_2 continues sur un intervalle $[a, b]$ de $\mathbb{R}$ telles que $f_1 \leq f_2$ sur $[a, b]$ et telles que

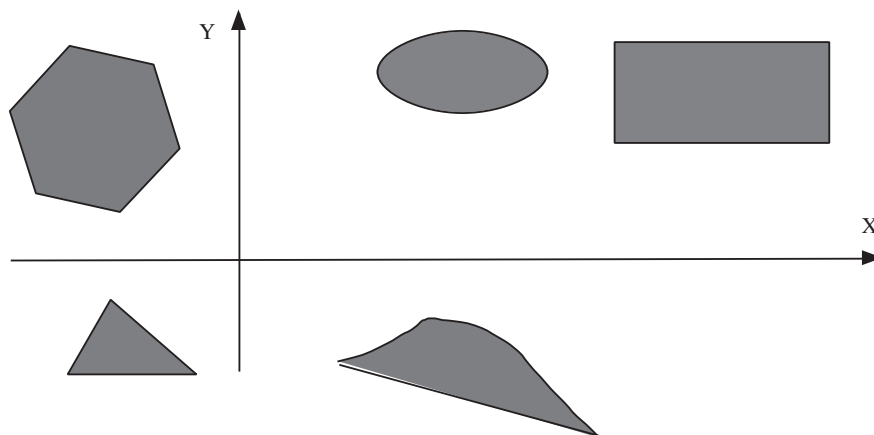
$$A = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, f_1(x) \leq y \leq f_2(x)\}.$$

b) On dit que A est parallèle à l'axe X s'il existe deux fonctions g_1, g_2 continues sur un intervalle $[c, d]$ de $\mathbb{R}$ telles que $g_1 \leq g_2$ sur $[c, d]$ et telles que

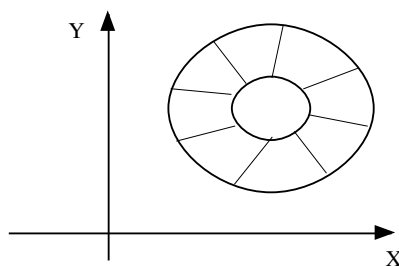
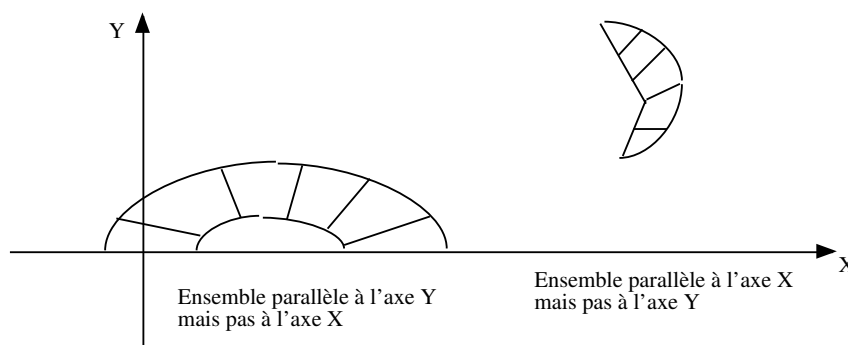
$$A = \{(x, y) \in \mathbb{R}^2 : g_1(y) \leq x \leq g_2(y), c \leq y \leq d\}.$$

Comment reconnaître géométriquement de tels ensembles dans les cas usuels ?

- Un ensemble est parallèle à l'axe Y lorsque toute droite verticale intersecte sa frontière au plus deux fois, exception faite des droites verticales qui composent elles-mêmes éventuellement la frontière,
- un ensemble est parallèle à l'axe X lorsque toute droite horizontale intersecte sa frontière au plus deux fois, exception faite des droites horizontales qui composent elles-mêmes éventuellement la frontière.



Ensembles parallèles à l'axe X et à l'axe Y



Ensemble non parallèle à l'axe X , non parallèle à l'axe Y

Le résultat suivant donne une manière de calculer les intégrales sur des ensembles de ce type.

Proposition 2.4.4 Soit f une fonction continue sur un ensemble A parallèle à l'axe X ou à l'axe Y . Alors f est intégrable sur A et on a

$$\int \int_A f(x, y) \, dx dy = \int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x, y) dy \right) dx$$

si A est parallèle à l'axe Y et

$$\int \int_A f(x, y) \, dx dy = \int_c^d \left(\int_{g_1(y)}^{g_2(y)} f(x, y) dx \right) dy$$

si A est parallèle à l'axe X .

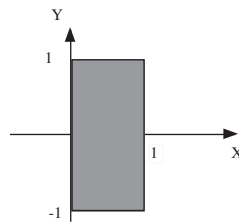
Remarquons que si A est à la fois parallèle à l'axe Y et à l'axe X , on a

$$\begin{aligned} \int \int_A f(x, y) \, dx dy &= \int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x, y) dy \right) dx \\ &= \int_c^d \left(\int_{g_1(y)}^{g_2(y)} f(x, y) dx \right) dy. \end{aligned}$$

On dit que l'on peut *permuter l'ordre d'intégration sans changer la valeur de l'intégrale*.

Exemples

- 1) Par exemple, intégrons la fonction $(x, y) \mapsto f(x, y) = x$ sur l'ensemble A suivant.



On a $A = [0, 1] \times [-1, 1]$. Cet ensemble est un rectangle fermé et borné; la fonction f étant continue sur A , elle est intégrable sur A .

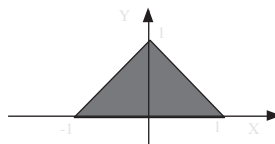
On a

$$\begin{aligned} \int \int_A f(x, y) \, dx dy &= \int_{-1}^1 \left(\int_0^1 f(x, y) \, dx \right) dy \\ &= \int_{-1}^1 \left(\int_0^1 x \, dx \right) dy \\ &= \int_{-1}^1 \left[\frac{x^2}{2} \right]_0^1 dy \\ &= \int_{-1}^1 \frac{1}{2} dy = \frac{1}{2} \int_{-1}^1 1 \, dy = \frac{1}{2} \cdot 2 = 1. \end{aligned}$$

Si on calcule cette intégrale dans l'autre sens, on a

$$\begin{aligned} \int \int_A f(x, y) \, dx dy &= \int_0^1 \left(\int_{-1}^1 f(x, y) \, dy \right) dx \\ &= \int_0^1 \left(\int_{-1}^1 x \, dy \right) dx \\ &= \int_0^1 2x \, dx = 2 \left[\frac{x^2}{2} \right]_0^1 = 1. \end{aligned}$$

- 2) Par exemple, intégrons la fonction $(x, y) \mapsto f(x, y) = y$ sur l'ensemble suivant.



Cet ensemble est fermé et borné, est à la fois parallèle à l'axe X et Y . Posons

$$f_1(x) = 0 \quad (x \in [-1, 1]), \quad f_2(x) = \begin{cases} x+1 & \text{si } x \in [-1, 0] \\ 1-x & \text{si } x \in [0, 1] \end{cases} \quad \text{et} \quad g_1(y) = y-1 \quad (y \in [0, 1]), \quad g_2(y) = 1-y \quad (y \in [0, 1]).$$

Les fonctions f_1, f_2 sont continues sur $[-1, 1]$; les fonctions g_1, g_2 sont continues sur $[0, 1]$. On a

$$A = \{(x, y) \in \mathbb{R}^2 : -1 \leq x \leq 1, f_1(x) \leq y \leq f_2(x)\} = \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq 1, g_1(y) \leq x \leq g_2(y)\}.$$

Dès lors, comme f est continu sur A , on a

$$\begin{aligned} \int \int_A f(x, y) \, dx dy &= \int_{-1}^1 \left(\int_{f_1(x)}^{f_2(x)} y dy \right) dx \\ &= \int_{-1}^0 \left(\int_0^{x+1} y dy \right) dx + \int_0^1 \left(\int_0^{1-x} y dy \right) dx \\ &= \int_{-1}^0 \frac{(x+1)^2}{2} dx + \int_0^1 \frac{(1-x)^2}{2} dx \\ &= \frac{1}{2} \left[\frac{(x+1)^3}{3} \right]_{-1}^0 + \frac{1}{2} \left[\frac{(x-1)^3}{3} \right]_0^1 \\ &= \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{3}. \end{aligned}$$

Calculons cette intégrale dans l'autre ordre. On a

$$\begin{aligned} \int \int_A f(x, y) \, dx dy &= \int_0^1 \left(\int_{g_1(y)}^{g_2(y)} y dx \right) dy \\ &= \int_0^1 \left(\int_{y-1}^{1-y} y dx \right) dy \\ &= \int_0^1 y(1-y-(y-1)) dy \\ &= 2 \int_0^1 y(1-y) dy \\ &= 2 \left[\frac{y^2}{2} - \frac{y^3}{3} \right]_0^1 = 2 \left(\frac{1}{2} - \frac{1}{3} \right) = \frac{1}{3}. \end{aligned}$$

3) Cas des variables séparées : si $A = [a, b] \times [c, d]$ et si $f(x, y) = f_1(x) f_2(y)$ avec f_1 continu sur $[a, b]$ et f_2 continu sur $[c, d]$, la fonction f est intégrable sur le rectangle A et on a

$$\int \int_A f_1(x) f_2(y) \, dx dy = \left(\int_a^b f_1(x) \, dx \right) \left(\int_c^d f_2(y) \, dy \right).$$

Intégration sur une union d'ensembles

Lorsqu'on doit intégrer une fonction continue sur A , avec $A = A_1 \cup A_2$, où A_1 et A_2 sont parallèles aux axes et ne se rencontrent que suivant leur frontière, on utilise la propriété suivante de l'intégrale :

$$\int \int_A f(x, y) \, dx dy = \int \int_{A_1} f(x, y) \, dx dy + \int \int_{A_2} f(x, y) \, dx dy.$$

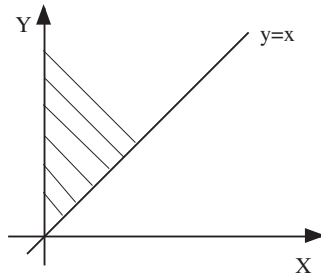
2.4.2 Intégration sur des ensembles non bornés fermés

Introduction

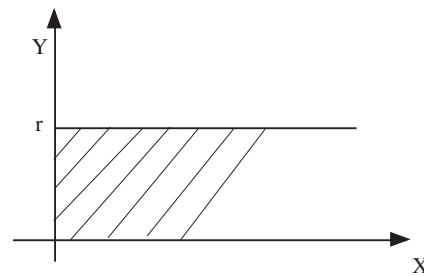
Ici, la définition de l'intégrabilité d'une fonction continue est plus délicate. Nous allons l'introduire de telle sorte que le parallélisme avec ce qui se passe à une variable soit quelque peu conservé.

Ici encore, considérons des fonctions continues sur des ensembles parallèles aux axes. Seulement, ces ensembles ne sont plus nécessairement bornés fermés. Cela revient à dire que, dans la définition introduite précédemment, l'intervalle $[a, b]$ (ou $[c, d]$) est remplacé par un intervalle quelconque de $\mathbb{R}$; de même, les inégalités faisant intervenir les fonctions continues f_1, f_2 (ou g_1, g_2) peuvent être remplacées par des inégalités strictes. De plus, l'une des fonctions f_1, f_2, g_1, g_2 peut ne pas apparaître.

Voici deux exemples. Le premier est un ensemble à la fois parallèle à X et à Y . Si on le considère comme parallèle à Y , l'intervalle de variation de x est $[0, +\infty[$ (au lieu de $[a, b]$), il n'y a pas de fonction f_2 et la fonction f_1 est $f_1(x) = x$. Si on le considère comme parallèle à X , l'intervalle de variation de y est $[0, +\infty[$ (au lieu de $[c, d]$), la fonction g_1 est $g_1(y) = 0$ et la fonction g_2 est $g_2(y) = y$.



$$\begin{aligned} A &= \{(x, y) \in \mathbb{R}^2 : 0 \leq x, x \leq y\} \\ &= \{(x, y) \in \mathbb{R}^2 : 0 \leq x, f_1(x) \leq y\} \\ &= \{(x, y) \in \mathbb{R}^2 : 0 \leq y, 0 \leq x \leq y\} \\ &= \{(x, y) \in \mathbb{R}^2 : 0 \leq y, g_1(y) \leq x \leq g_2(y)\} \end{aligned}$$



$$\begin{aligned} A &= \{(x, y) \in \mathbb{R}^2 : 0 \leq x, 0 \leq y \leq r\} \\ &= [0, +\infty[\times [0, r] \end{aligned}$$

Dans ce qui suit, nous utiliserons des ensembles parallèles aux axes. Afin de ne pas alourdir les notations, nous utiliserons systématiquement la description à l'aide des fonctions f_1, f_2, g_1, g_2 . Bien entendu, dans le cas où la description de A ne les fait pas apparaître, il convient d'adapter les bornes d'intégration dans les notations qui suivent (cela revient à utiliser $+\infty$ ou $-\infty$).

On voudrait conserver la propriété de l'intégrale qui dit que l'on peut permuter l'ordre d'intégration sans altérer la limite ; cette propriété se révèle en effet d'une grande utilité en pratique.

Donnons un exemple montrant que, dans le contexte de l'intégrale d'une fonction continue sur un ensemble qui n'est pas borné fermé et sans hypothèse supplémentaire, on ne peut pas toujours permuter les intégrales sans changer de valeur.

Soit la fonction

$$f(x, y) = \frac{x^2 - y^2}{(x^2 + y^2)^2}, \quad A =]0, 1] \times]0, 1].$$

Cette fonction est continue sur A .

Remarquons d'abord que

$$f(x, y) = D_x \frac{-x}{x^2 + y^2} = D_y \frac{y}{x^2 + y^2}, \quad (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}.$$

Cela étant, d'une part, pour $y \in]0, 1]$ fixé, la fonction $x \mapsto \frac{x^2 - y^2}{(x^2 + y^2)^2}$ est continue sur $[0, 1]$ donc y est intégrable. On a

$$\int_0^1 f(x, y) dx = \int_0^1 D_x \frac{-x}{x^2 + y^2} dx = \left[\frac{-x}{x^2 + y^2} \right]_{x=0}^{x=1} = \frac{-1}{1 + y^2};$$

cette fonction est intégrable sur $[0, 1]$ et

$$\int_0^1 \frac{-1}{1 + y^2} dy = -\arctg(1) = -\frac{\pi}{4}.$$

Dès lors

$$\int_0^1 \left(\int_0^1 f(x, y) dx \right) dy = -\frac{\pi}{4}.$$

D'autre part, pour $x \in]0, 1]$ fixé, la fonction $y \mapsto \frac{x^2 - y^2}{(x^2 + y^2)^2}$ est continue sur $[0, 1]$ donc y est intégrable. On a

$$\int_0^1 f(x, y) dy = \int_0^1 D_y \frac{y}{x^2 + y^2} dy = \left[\frac{y}{x^2 + y^2} \right]_{y=0}^{y=1} = \frac{1}{1 + x^2};$$

cette fonction est intégrable sur $[0, 1]$ et

$$\int_0^1 \frac{1}{1+x^2} dx = \operatorname{arctg}(1) = \frac{\pi}{4}.$$

Dès lors

$$\int_0^1 \left(\int_0^1 f(x, y) dy \right) dx = \frac{\pi}{4}.$$

Cependant, on a le résultat suivant (admis sans démonstration).

Propriété 2.4.5 *Soit f une fonction continue sur A , parallèle à l'axe Y et à l'axe X (ou union d'ensembles de ce type).*

Si

- *pour tout $x \in]a, b[$, la fonction $|f(x, y)|$, $y \in]f_1(x), f_2(x)[$, est intégrable sur $]f_1(x), f_2(x)[$*

- *et si la fonction $\int_{f_1(x)}^{f_2(x)} |f(x, y)| dy$, $x \in]a, b[$, est intégrable sur $]a, b[$*

alors

- *pour presque tout² $y \in]c, d[$, la fonction $|f(x, y)|$, $x \in]g_1(y), g_2(y)[$, est intégrable sur l'intervalle $]g_1(y), g_2(y)[$*

- *la fonction $\int_{g_1(y)}^{g_2(y)} |f(x, y)| dx$, $y \in]c, d[$, est intégrable sur $]c, d[$*

et on a

$$\int_a^b \left(\int_{f_1(x)}^{f_2(x)} |f(x, y)| dy \right) dx = \int_c^d \left(\int_{g_1(y)}^{g_2(y)} |f(x, y)| dx \right) dy$$

et

$$\int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x, y) dy \right) dx = \int_c^d \left(\int_{g_1(y)}^{g_2(y)} f(x, y) dx \right) dy$$

Définition

Nous adopterons la définition suivante.

Définition 2.4.6 *Soit f une fonction continue sur A , parallèle à l'axe Y . On dit que f est intégrable sur A lorsque*

- *pour tout $x \in]a, b[$, la fonction $|f(x, y)|$, $y \in]f_1(x), f_2(x)[$, est intégrable sur $]f_1(x), f_2(x)[$*

- *et lorsque la fonction $\int_{f_1(x)}^{f_2(x)} |f(x, y)| dy$, $x \in]a, b[$, est intégrable sur $]a, b[$.*

L'intégrale de f sur A est alors

$$\int \int_A f(x, y) dx dy = \int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x, y) dy \right) dx.$$

La définition est analogue, avec les modifications naturelles, lorsque A est parallèle à l'axe X .

On remarquera que si la fonction est à valeurs positives, le fait de pouvoir intégrer dans un certain ordre entraîne automatiquement l'intégrabilité de la fonction et l'égalité avec l'intégrale calculée dans l'autre ordre.

On remarquera aussi que puisque le module du module d'un nombre est le module du nombre, la définition donne le résultat suivant : si f est continu sur A alors elle est intégrable sur A si et seulement si $|f|$ est intégrable sur A .

Ainsi, étant donné la définition adoptée et la propriété précédente, on obtient celle-ci.

2. Ce qui signifie rigoureusement "à l'exception d'un ensemble négligeable"

Propriété 2.4.7 Soit f une fonction continue sur A , à la fois parallèle à l'axe X et à l'axe Y . Si f est intégrable sur A alors

$$\int \int_A f(x, y) \, dx \, dy = \int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x, y) \, dy \right) \, dx = \int_c^d \left(\int_{g_1(y)}^{g_2(y)} f(x, y) \, dx \right) \, dy.$$

Exemples fondamentaux

1) Voici un cas où il est très facile de vérifier que f est intégrable sur $A =]a, b[\times]c, d[$ et où le calcul de l'intégrale peut s'effectuer assez aisément : si h est une fonction continue et intégrable sur $]a, b[$ et si H est une fonction continue et intégrable sur $]c, d[$ alors la fonction $(x, y) \mapsto f(x, y) = h(x) H(y)$ est intégrable sur $A =]a, b[\times]c, d[$ et on a

$$\int \int_A f(x, y) \, dx \, dy = \left(\int_a^b h(x) \, dx \right) \left(\int_c^d H(y) \, dy \right).$$

2) On a la propriété suivante (nous l'admettrons bien que sa preuve consiste seulement à utiliser la définition et l'hypothèse).

Propriété 2.4.8 Soient f, F deux fonctions continues sur A . Si F est intégrable sur A et si

$$|f| \leq F \quad \text{sur} \quad A$$

alors f est intégrable sur A et

$$\left| \int \int_A f(x, y) \, dx \, dy \right| \leq \int \int_A F(x, y) \, dx \, dy.$$

Exemples

Déterminer si les intégrales suivantes existent. Si oui, les calculer. Représenter également les ensembles d'intégration.

1) $\int \int_A e^{-xy} \, dx \, dy$ avec $A = [0, +\infty[\times [1, 2]$.

La fonction donnée par $f(x, y) = e^{-xy}$ est positive et continue sur A . On a, pour tout $y \in [1, 2]$

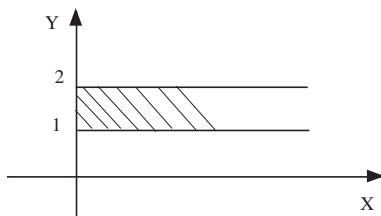
$$\int_0^{+\infty} e^{-xy} \, dx = \left[-\frac{1}{y} e^{-xy} \right]_0^{+\infty} = \frac{1}{y}$$

puis

$$\int_1^2 \frac{1}{y} \, dy = [\ln y]_1^2 = \ln 2.$$

La fonction $(x, y) \mapsto f(x, y) = e^{-xy}$ est donc intégrable sur A et

$$\int \int_A e^{-xy} \, dx \, dy = \ln 2.$$



2) $\int \int_A e^{-y^2} dx dy$ avec $A = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq y\}$.

La fonction donnée par $f(x, y) = e^{-y^2}$ est positive et continue sur A . On a $A = \{(x, y) \in \mathbb{R}^2 : 0 \leq x, x \leq y\}$ donc A apparaît comme parallèle à l'axe Y . Pour $x \geq 0$, on est alors tenté de calculer $\int_x^{+\infty} e^{-y^2} dy$. Ce calcul n'est pas possible de façon élémentaire. Ecrivons alors plutôt A comme étant parallèle à l'axe X :

$$A = \{(x, y) \in \mathbb{R}^2 : 0 \leq y, 0 \leq x \leq y\}.$$

Pour $y \geq 0$, on a alors

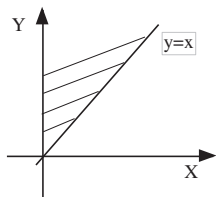
$$\int_0^y e^{-y^2} dx = ye^{-y^2}$$

puis

$$\int_0^{+\infty} ye^{-y^2} dy = \int_0^{+\infty} \left(-\frac{1}{2}\right) D_y e^{-y^2} dy = -\frac{1}{2} [e^{-y^2}]_0^{+\infty} = \frac{1}{2}.$$

La fonction $(x, y) \mapsto e^{-y^2}$ est donc intégrable sur A et

$$\int \int_A e^{-y^2} dx dy = \frac{1}{2}.$$



3) $\int \int_A \frac{y}{\sqrt{x^2 + y^2}} dx dy$ avec $A = \{(x, y) \in \mathbb{R}^2 : 0 < x \leq 1, 0 < y \leq x^2\}$.

La fonction $(x, y) \mapsto f(x, y) = y/\sqrt{x^2 + y^2}$ est positive et continue sur A . Pour $x > 0$, on a

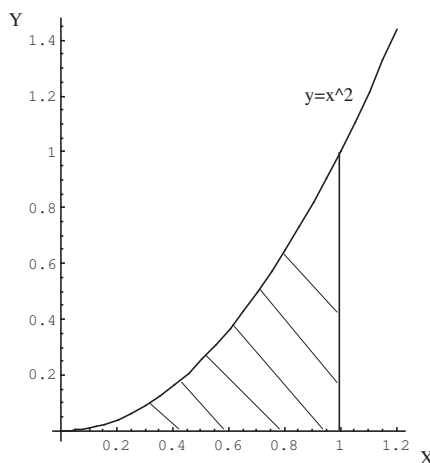
$$\int_0^{x^2} f(x, y) dy = \int_0^{x^2} D_y \sqrt{x^2 + y^2} dy = \left[\sqrt{x^2 + y^2} \right]_{y=0}^{y=x^2} = x\sqrt{1+x^2} - x.$$

On a ensuite

$$\int_0^1 \left(x\sqrt{1+x^2} - x \right) dx = \int_0^1 \left(\frac{1}{3} D_x (1+x^2)^{3/2} - D_x \frac{x^2}{2} \right) dx = \left[\frac{1}{3} (1+x^2)^{3/2} - \frac{x^2}{2} \right]_0^1 = \frac{2^{3/2}}{3} - \frac{1}{2} - \frac{1}{3} = \frac{1}{6} (4\sqrt{2} - 5).$$

La fonction f est donc intégrable sur A et on a

$$\int \int_A f(x, y) dx dy = \frac{1}{6} (4\sqrt{2} - 5).$$



$$4) \int \int_A e^{-x} \frac{\cos(xy)}{\sqrt{1-y^2}} dx dy \quad \text{avec } A = \{(x, y) \in \mathbb{R}^2 : 0 \leq x, 0 \leq y < 1\}.$$

La fonction donnée par $f(x, y) = e^{-x} \frac{\cos(xy)}{\sqrt{1-y^2}}$ est continue sur A . Elle n'est pas positive sur A . Comme

$$|f(x, y)| = \left| e^{-x} \frac{\cos(xy)}{\sqrt{1-y^2}} \right| \leq e^{-x} \frac{1}{\sqrt{1-y^2}}$$

pour montrer l'intégrabilité de f , il suffit de montrer l'intégrabilité de $(x, y) \mapsto e^{-x} \frac{1}{\sqrt{1-y^2}}$. Cette fonction est continue et intégrable sur A . De plus, elle est le produit de $h(x) = e^{-x}$, intégrable sur $[0, +\infty[$ (car fonction continue, positive sur l'intervalle et telle que $\lim_{t \rightarrow +\infty} \int_0^t e^{-x} dx = \lim_{t \rightarrow +\infty} (1 - e^{-t}) = 1$) et de $H(y) = 1/\sqrt{1-y^2}$ intégrable sur $[0, 1[$ (car fonction continue, positive sur cet intervalle et telle que $\int_0^t 1/\sqrt{1-y^2} dy = \arcsin t$ admet une limite finie si $t \rightarrow 1$).

Cela étant, calculons l'intégrale de f sur A (par des méthodes réelles). On a successivement, pour $y \in [0, 1[$:

$$\begin{aligned} \int_0^{+\infty} e^{-x} \cos(xy) dx &= - \int_0^{+\infty} D_x e^{-x} \cos(xy) dx \\ &= - \left[e^{-x} \cos(xy) \right]_0^{+\infty} - y \int_0^{+\infty} e^{-x} \sin(xy) dx \\ &= 1 + y \int_0^{+\infty} D_x e^{-x} \sin(xy) dx \\ &= 1 + y \left[e^{-x} \sin(xy) \right]_0^{+\infty} - y^2 \int_0^{+\infty} e^{-x} \cos(xy) dx \\ &= 1 - y^2 \int_0^{+\infty} e^{-x} \cos(xy) dx \end{aligned}$$

donc

$$\int_0^{+\infty} e^{-x} \cos(xy) dx = \frac{1}{1+y^2}.$$

Ensuite, calculons

$$\int_0^1 \frac{1}{(1+y^2)\sqrt{1-y^2}} dy.$$

Effectuons le changement de variables $y = g(t) = \sin t$, $g :]0, \pi/2[\rightarrow]0, 1[$. On a

$$\int_0^1 \frac{1}{(1+y^2)\sqrt{1-y^2}} dy = \int_0^{\pi/2} \frac{1}{(1+\sin^2 t) \cos t} \cos t dt = \int_0^{\pi/2} \frac{1}{1+\sin^2 t} dt.$$

On a

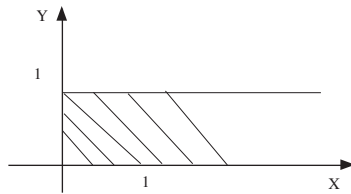
$$\frac{1}{1+\sin^2 t} = \frac{1}{\cos^2 t} \frac{1}{\frac{1}{\cos^2 t} + \tan^2 t} = \frac{1}{\cos^2 t} \frac{1}{1+2\tan^2 t} = \frac{1}{\sqrt{2}} D_t \arctg(\sqrt{2} \tan t)$$

donc

$$\int_0^{\pi/2} \frac{1}{1+\sin^2 t} dt = \frac{1}{\sqrt{2}} \left[\arctg(\sqrt{2} \tan t) \right]_0^{\pi/2} = \frac{1}{\sqrt{2}} \frac{\pi}{2} = \frac{\sqrt{2}}{4} \pi.$$

Finalement,

$$\int \int_A f(x, y) dx dy = \frac{\sqrt{2}}{4} \pi.$$



2.4.3 Intégration par changement de variables

Cas général

Dans le cadre des fonctions de plusieurs variables, il existe également un résultat d'intégration par changement de variables. Il s'énonce comme suit et sera admis sans démonstration. Signalons également que les ensembles ne sont plus nécessairement parallèles aux axes.

Théorème 2.4.9 Soient deux fonctions g_1, g_2 qui sont continûment dérivables sur $A \subset \mathbb{R}^2$ et établissent une bijection entre A et A' telle que la bijection inverse soit aussi formée par deux fonctions continûment dérivables sur A' .

Alors la fonction continue $f(x', y')$ est intégrable sur A' si et seulement si la fonction $f(g_1(x, y), g_2(x, y)) |D_x g_1 D_y g_2 - D_x g_2 D_y g_1|$ est intégrable sur A . Dans ce cas

$$\int \int_{A'} f(x', y') dx' dy' = \int \int_A f(g_1(x, y), g_2(x, y)) |D_x g_1 D_y g_2 - D_x g_2 D_y g_1| dx dy.$$

La fonction

$$(x, y) \mapsto D_x g_1 D_y g_2 - D_x g_2 D_y g_1$$

s'appelle

le jacobien

du changement de variables.

Remarquons que le résultat concernant l'intégration par changement de variables dans le cas des fonctions d'une variable est un cas particulier de ce théorème.

Changement de variables polaires

Un cas particulier très utile de changement de variables est celui qui fait intervenir les coordonnées polaires.

Rappelons que tout $(x', y') \in \mathbb{R}^2$, $(x', y') \neq (0, 0)$ peut s'écrire de manière unique sous la forme³

$$x' = r \cos \theta, \quad y' = r \sin \theta$$

où $r > 0$ et $\theta \in [0, 2\pi[$. Si on note $x = r, y = \theta$, on peut définir

$$g_1(r, \theta) = r \cos \theta, \quad g_2(r, \theta) = r \sin \theta.$$

Ces fonctions sont continûment dérivables sur $\mathbb{R}^2$. Elles établissent une bijection entre l'ensemble $]0, +\infty[\times]0, 2\pi[$ et $\mathbb{R}^2 \setminus \{(0, 0)\}$ et s'inversent selon

$$r = \sqrt{x^2 + y^2}, \quad \theta = \begin{cases} \arctg(\frac{y}{x}) & \text{si } x > 0, y \geq 0 \\ \arctg(\frac{y}{x}) + \pi & \text{si } x < 0, y \in \mathbb{R} \\ \arctg(\frac{y}{x}) + 2\pi & \text{si } x > 0, y < 0 \\ \frac{\pi}{2} & \text{si } x = 0, y > 0 \\ \frac{3\pi}{2} & \text{si } x = 0, y < 0 \end{cases}$$

On montre que ces deux fonctions sont aussi continûment dérivables sur $]0, +\infty[\times]0, 2\pi[$.

Par exemple, si

$$A' =]0, +\infty[\times]0, +\infty[,$$

on a

$$A =]0, +\infty[\times]0, \pi/2[.$$

De même si

$$A' = \{(x', y') \in \mathbb{R}^2 : x' > 0, y' > 0, x'^2 + y'^2 \leq 1\}$$

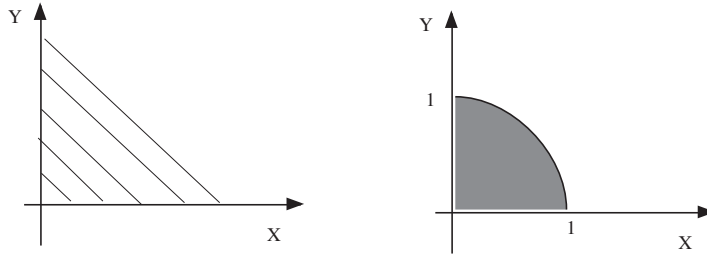
on a

$$A =]0, 1] \times]0, \pi/2[.$$

3. On a également la propriété suivante : soient $x'_0, y'_0 \in \mathbb{R}$. Tout $(x', y') \in \mathbb{R}^2$, $(x, y) \neq (x_0, y_0)$ peut s'écrire de manière unique sous la forme

$$x' - x'_0 = r \cos \theta, \quad y' - y'_0 = r \sin \theta$$

où $r > 0$ et $\theta \in [0, 2\pi[$.



Calculons le jacobien de ce changement de variables. On a

$$D_r g_1(r, \theta) = D_r(r \cos \theta) = \cos \theta, \quad D_\theta g_1(r, \theta) = D_\theta(r \cos \theta) = -r \sin \theta$$

et

$$D_r g_2(r, \theta) = D_r(r \sin \theta) = \sin \theta, \quad D_\theta g_2(r, \theta) = D_\theta(r \sin \theta) = r \cos \theta$$

donc

$$D_r g_1 D_\theta g_2 - D_r g_2 D_\theta g_1 = r \cos^2 \theta + r \sin^2 \theta = r.$$

Il s'ensuit que la formule de changement de variables s'écrit dans ce cas

$$\int \int_{A'} f(x, y) dx dy = \int \int_A f(r \cos \theta, r \sin \theta) r dr d\theta$$

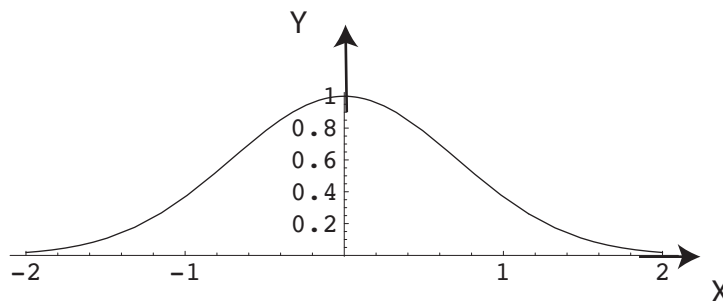
où $A' \subset \mathbb{R}^2$ et $A \subset]0, +\infty[\times]0, 2\pi[$.

2.4.4 Applications

Calcul de $\int_0^{+\infty} e^{-ax^2} dx$ avec $a > 0$

Lorsque $a > 0$, la fonction $f(x) = e^{-ax^2}$ est intégrable sur $\mathbb{R}$: elle y est en effet continue et $\lim_{x \rightarrow +\infty} x^2 e^{-ax^2} = \lim_{x \rightarrow -\infty} x^2 e^{-ax^2} = 0$.

Cette fonction a une importance considérable notamment en probabilités-statistiques. Dans ce cadre, à une constante près, elle s'appelle la densité de probabilité gaussienne (c'est la densité de probabilité d'une variable aléatoire dite normale). Voici la représentation graphique de $f(x) = e^{-x^2}$, $x \in \mathbb{R}$ (La représentation est uniquement effectuée sur $[-2, 2]$ pour que la figure soit un peu esthétique ; les valeurs de cette fonction sont très petites par rapport à 1 déjà pour des valeurs de x telles que $|x| \geq 2$).



Notons I_a l'intégrale demandée. Pour la calculer, considérons la fonction de deux variables

$$F(x, y) = e^{-a(x^2+y^2)} = f(x) f(y).$$

En utilisant le changement de variables polaires, on a

$$\int \int_A F(x, y) dx dy = I_a^2 = \int \int_{A'} e^{-ar^2} r dr d\theta$$

avec

$$A =]0, +\infty[\times]0, +\infty[\quad A' =]0, +\infty[\times]0, \frac{\pi}{2}[.$$

L'intégrale de droite se calcule aisément :

$$\begin{aligned} \int \int_{A'} e^{-ar^2} r \, dr \, d\theta &= \left(\int_0^{\pi/2} 1 \, d\theta \right) \left(\int_0^{+\infty} r e^{-ar^2} \, dr \right) \\ &= \frac{\pi}{2} \int_0^{+\infty} \frac{-1}{2a} D_r e^{-ar^2} \, dr = \frac{-\pi}{4a} \left[e^{-ar^2} \right]_0^{+\infty} \\ &= \frac{\pi}{4a}. \end{aligned}$$

Ainsi, puisque $I_a > 0$, on obtient

$$I_a = \frac{1}{2} \sqrt{\frac{\pi}{a}}.$$

Calcul de $\int_0^{+\infty} \frac{\sin x}{x} dx$

Cette intégrale ressemble aux intégrales *de Fresnel* qui interviennent dans la diffraction⁴. Elle a été introduite spécifiquement notamment par Dirichlet et d'autres pour des calculs intervenant dans la recherche des solutions d'équations liées à l'électromagnétisme, à des phénomènes de pression, ...

Notons aussi que cette intégrale intervient dans la théorie des transformations de Fourier ($\sin x/x$ est, à une constante multiplicative près, la transformée de Fourier de la fonction caractéristique de l'intervalle $[-1, 1]$).

Nous voulons établir que la limite

$$\lim_{t \rightarrow +\infty} \int_0^t \frac{\sin x}{x} dx$$

existe, est finie et calculer sa valeur.

Remarquons d'abord que, pour tout $t > 0$, la fonction $\frac{\sin x}{x}$ est continue sur $]0, t]$ et admet une limite finie en 0. Elle est donc intégrable sur l'intervalle $]0, t]$. Notons

$$I_t = \int_0^t \frac{\sin x}{x} dx.$$

Cela étant, pour tout $x > 0$, on a

$$\frac{1}{x} = \int_0^{+\infty} e^{-sx} ds.$$

On a donc, pour $t > 0$:

$$I_t = \int_0^t \frac{\sin x}{x} dx = \int_0^t \sin x \left(\int_0^{+\infty} e^{-xs} ds \right) dx. \quad (2.1)$$

Pour $t > 0$, on considère alors la fonction

$$|\sin x| e^{-sx}, \quad s \in]0, +\infty[, \quad x \in]0, t].$$

On a

$$\int_0^{+\infty} |\sin x| e^{-sx} ds = \frac{|\sin x|}{x}$$

4. Intégrales de Fresnel : $\frac{1}{\sqrt{2\pi}} \int_0^{x^2} \frac{\cos t}{\sqrt{t}} dt = \sqrt{\frac{2}{\pi}} \int_0^x \cos(t^2) dt$ et $\frac{1}{\sqrt{2\pi}} \int_0^{x^2} \frac{\sin t}{\sqrt{t}} dt = \sqrt{\frac{2}{\pi}} \int_0^x \sin(t^2) dt$.

et $x \mapsto \frac{|\sin x|}{x}$ est une fonction intégrable sur $]0, t]$. Il s'ensuit (cf résultat énoncé auparavant) que, dans (2.1), on peut permuter l'ordre l'intégration sans changer de valeur. On obtient donc

$$I_t = \int_0^t \sin x \left(\int_0^{+\infty} e^{-xs} ds \right) dx = \int_0^{+\infty} \left(\int_0^t \sin x e^{-xs} dx \right) ds.$$

Calculons l'intégrale par rapport à x (par des méthodes réelles). On a

$$\begin{aligned} \int_0^t \sin x e^{-xs} dx &= - \int_0^t D_x \cos x e^{-xs} dx \\ &= - [\cos x e^{-xs}]_{x=0}^{x=t} - s \int_0^t \cos x e^{-xs} dx \\ &= - \cos t e^{-ts} + 1 - s \int_0^t D_x \sin x e^{-xs} dx \\ &= - \cos t e^{-ts} + 1 - s \left([\sin x e^{-xs}]_{x=0}^{x=t} + s \int_0^t \sin x e^{-xs} dx \right) \\ &= - \cos t e^{-ts} + 1 - s \sin t e^{-ts} - s^2 \int_0^t \sin x e^{-xs} dx \end{aligned}$$

donc

$$\int_0^t \sin x e^{-xs} dx = \frac{1}{1+s^2} (1 - \cos t e^{-st} - s \sin t e^{-st}).$$

Dès lors,

$$\begin{aligned} I_t &= \int_0^{+\infty} \frac{1}{1+s^2} (1 - \cos t e^{-st} - s \sin t e^{-st}) ds \\ &= \int_0^{+\infty} \frac{1}{1+s^2} ds - \cos t \int_0^{+\infty} \frac{e^{-st}}{1+s^2} ds - \sin t \int_0^{+\infty} \frac{s e^{-st}}{1+s^2} ds. \end{aligned}$$

On a

$$\int_0^{+\infty} \frac{1}{1+s^2} ds = [\arctg s]_0^{+\infty} = \frac{\pi}{2};$$

de plus, comme

$$\left| \cos t \int_0^{+\infty} \frac{e^{-st}}{1+s^2} ds \right| \leq \int_0^{+\infty} e^{-st} ds = \frac{1}{t}$$

et

$$\left| \sin t \int_0^{+\infty} \frac{s e^{-st}}{1+s^2} ds \right| \leq \int_0^{+\infty} e^{-st} ds = \frac{1}{t}$$

on obtient que

$$\lim_{t \rightarrow +\infty} \cos t \int_0^{+\infty} \frac{e^{-st}}{1+s^2} ds = 0, \quad \lim_{t \rightarrow +\infty} \sin t \int_0^{+\infty} \frac{s e^{-st}}{1+s^2} ds = 0$$

donc finalement

$$\boxed{\lim_{t \rightarrow +\infty} I_t = \int_0^{+\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}.}$$

2.4.5 Intégrales triples

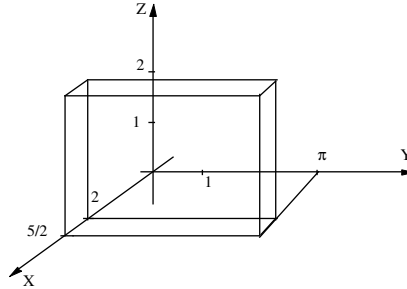
Cas général

De manière analogue à ce qui vient de se faire pour les intégrales doubles, on peut introduire aussi les intégrales triples. Les propriétés sont similaires. L'intégration se ramène à calculer successivement

trois intégrales à une variable (on parle d'intégrale simple quand on traite une intégrale d'une fonction d'une variable) au lieu de deux dans le cas de deux variables.

Soit R le rectangle du plan X, Y délimité par les droites d'équations $x = 2, x = \frac{5}{2}, y = 0, y = \pi$ et soit C le parallélépipède situé entre les plans d'équation $z = 0, z = 2$ sur R . Calculer

$$\int \int \int_C zx \sin(xy) \, dx dy dz.$$



L'ensemble d'intégration C s'écrit

$$C = \{(x, y, z) \in \mathbb{R}^3 : 2 \leq x \leq \frac{5}{2}, 0 \leq y \leq \pi, 0 \leq z \leq 2\} = [2, \frac{5}{2}] \times [0, \pi] \times [0, 2].$$

On a donc

$$\begin{aligned} \int \int \int_C zx \sin(xy) \, dx dy dz &= \int_2^{5/2} dx \int_0^\pi dy \int_0^2 zx \sin(xy) \, dz \\ &= \int_2^{5/2} dx \int_0^\pi dy \left[\frac{z^2}{2} x \sin(xy) \right]_{z=0}^{z=2} \\ &= \int_2^{5/2} dx \int_0^\pi 2x \sin(xy) dy = \int_2^{5/2} dx \int_0^\pi 2D_y(-\cos(xy)) dy \\ &= 2 \int_2^{5/2} dx [-\cos(xy)]_{y=0}^{y=\pi} = 2 \int_2^{5/2} dx (1 - \cos(x\pi)) \\ &= 2 \left[x - \frac{1}{\pi} \sin(\pi x) \right]_2^{5/2} = 2 \left(\frac{5}{2} - \frac{1}{\pi} \sin\left(\frac{5\pi}{2}\right) - 2 + \frac{1}{\pi} \sin(2\pi) \right) \\ &= 1 - \frac{2}{\pi}. \end{aligned}$$

Volume d'un corps

Nous avons introduit précédemment la notion de *volume* d'un corps limité par une surface d'équation $z = f(x, y)$ ($(x, y) \in R$ (f étant à valeurs positives, R étant un rectangle) par

$$V = \int \int_R f(x, y) \, dx dy.$$

Cette notion s'étend à des fonctions qui sont définies sur des ensembles qui ne sont plus des rectangles. De plus, si on remarque que le corps dont on veut calculer le volume se décrit par

$$C = \{(x, y, z) : (x, y) \in R, 0 \leq z \leq f(x, y)\}$$

l'intégrale donnant le volume s'écrit

$$V = \int \int_R f(x, y) \, dx dy = \int \int_R \left(\int_0^{f(x, y)} 1 dz \right) dx dy = \int \int \int_C 1 \, dx dy dz.$$

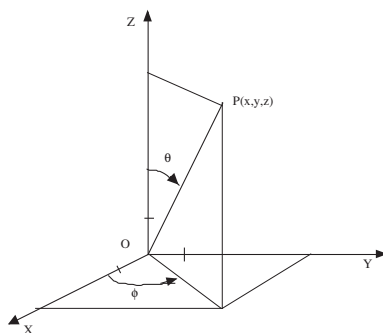
Définition 2.4.10 Soit C un corps borné fermé de l'espace. Le volume de C est

$$V = \int \int \int_C 1 \, dx dy dz.$$

Cette définition recouvre les cas particuliers de calculs de volumes que nous avons introduits dans le cadre de l'intégrale de fonctions d'une variable. Pour le voir, il suffit en fait de décrire analytiquement le corps, c'est-à-dire de trouver C , puis de ramener l'intégrale triple à une intégrale simple en effectuant le calcul de deux des intégrales simples.

Changement de variables polaires dans $\mathbb{R}^3$

Dans l'espace (muni d'un repère orthonormé), on introduit également les coordonnées sphériques (ou coordonnées polaires dans l'espace).



Si (x, y, z) désignent les coordonnées cartésiennes d'un point P de l'espace différent de l'origine, les coordonnées sphériques de ce point sont les réels (r, φ, θ) avec $r > 0$, $\varphi \in [0, 2\pi[$, $\theta \in [0, \pi]$ tels que

$$\begin{cases} x &= r \cos \varphi \sin \theta \\ y &= r \sin \varphi \sin \theta \\ z &= r \cos \theta. \end{cases}$$

Ici encore, on définit un jacobien. La formule de changement de variables dans les intégrales s'écrit alors

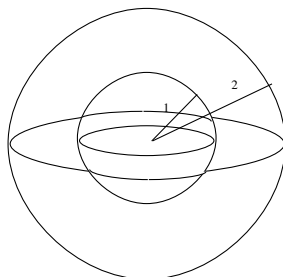
$$\int \int \int_A f(x, y, z) \, dx dy dz = \int \int \int_{A'} f(r \cos \varphi \sin \theta, r \sin \varphi \sin \theta, r \cos \theta) \, r^2 \sin \theta \, dr d\varphi d\theta$$

où l'on suppose que l'un des intégrands existe et où A et A' sont respectivement les expressions d'une partie de l'espace décrite avec les coordonnées cartésiennes d'une part et les coordonnées sphériques d'autre part.

Ainsi par exemple, calculons

$$\int \int \int_C z^2 \, dx dy dz$$

où C est la région de l'espace comprise entre les sphères centrées à l'origine de rayon 1 et 2.



Les descriptions en coordonnées cartésiennes et en coordonnées sphériques de C sont successivement

$$\begin{aligned} A &= \{(x, y, z) \in \mathbb{R}^3 : 1 \leq x^2 + y^2 + z^2 \leq 4\} \\ A' &= \{(r, \varphi, \theta) : 1 \leq r \leq 2, \varphi \in [0, 2\pi[, \theta \in [0, \pi]\}. \end{aligned}$$

On a donc

$$\begin{aligned} \int \int \int_A z^2 \, dx dy dz &= \int \int \int_{A'} (r \cos \theta)^2 r^2 \sin \theta \, dr d\varphi d\theta = \int_0^\pi \cos^2 \theta \sin \theta \, d\theta \int_1^2 r^4 \, dr \int_0^{2\pi} 1 \, d\varphi \\ &= \left[-\frac{1}{3} \cos^3 \theta \right]_0^\pi \left[\frac{r^5}{5} \right]_1^2 2\pi \\ &= \frac{124}{15} \pi. \end{aligned}$$

2.5 Annexe

2.5.1 Exemple où on ne peut pas permuter les dérivées

La fonction

$$f(x, y) = \begin{cases} \frac{x^3 y - x y^3}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

est deux fois dérivable dans $\mathbb{R}^2$ mais

$$D_x D_y f(0, 0) \neq D_y D_x f(0, 0).$$

En effet, vu sa forme, la fonction est deux fois dérivable dans $\mathbb{R}^2 \setminus \{(0, 0)\}$. Examinons ce qui se passe au point $(0, 0)$. On a

$$\lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0$$

donc la fonction est dérivable par rapport à sa première variable en $(0, 0)$ et on a

$$D_x f(x, y) = \begin{cases} \frac{x^4 y + 4x^2 y^3 - y^5}{(x^2 + y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0). \end{cases}$$

De même on a

$$\lim_{h \rightarrow 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0$$

donc la fonction est dérivable par rapport à sa seconde variable en $(0, 0)$ et on a

$$D_y f(x, y) = \begin{cases} \frac{x^5 - 4y^2 x^3 - x y^4}{(x^2 + y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0). \end{cases}$$

Cela étant, montrons que la fonction est deux fois dérivable en $(0, 0)$. On a successivement

$$\lim_{h \rightarrow 0} \frac{D_x f(h, 0) - D_x f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0}{h} = 0$$

donc

$$D_x^2 f(0, 0) = 0;$$

$$\lim_{h \rightarrow 0} \frac{D_y f(0, h) - D_y f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0}{h} = 0$$

donc

$$D_y^2 f(0, 0) = 0;$$

$$\lim_{h \rightarrow 0} \frac{D_y f(h, 0) - D_y f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{h^5/h^4 - 0}{h} = 1$$

donc

$$D_x D_y f(0, 0) = 1;$$

$$\lim_{h \rightarrow 0} \frac{D_x f(0, h) - D_x f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{-h^5/h^4 - 0}{h} = -1$$

donc

$$D_y D_x f(0, 0) = -1.$$

La fonction est donc deux fois dérivable sur $\mathbb{R}^2$ mais

$$D_x D_y f(0, 0) \neq D_y D_x f(0, 0).$$

Chapitre 3

Approximations polynomiales et séries

3.1 Approximation de fonctions par des polynômes, développements limités

Dans cette section, nous allons voir comment on peut estimer les valeurs d'une fonction au voisinage d'un point à l'aide de polynômes.

Remarquons que nous avons déjà à notre disposition les *approximations linéaires*, c'est-à-dire les approximations à l'aide d'un polynôme de degré plus petit ou égal à 1. En effet, si f est dérivable en x_0 , on a

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - (x - x_0)Df(x_0)}{x - x_0} = 0$$

donc le polynôme

$$P(x - x_0) = f(x_0) + (x - x_0)Df(x_0)$$

vérifie

$$\lim_{x \rightarrow x_0} \frac{f(x) - P(x - x_0)}{x - x_0} = 0.$$

On obtient donc qu'au voisinage de x_0 , les valeurs de f sont proches de celles de $P(x - x_0)$, le facteur correctif (c'est-à-dire $R(x) = f(x) - P(x - x_0)$, appelé *reste*) vérifiant

$$\lim_{x \rightarrow x_0} \frac{R(x)}{x - x_0} = 0.$$

Cela signifie que $|R(x)|$ est très petit devant $|x - x_0|$ lorsque x est proche de x_0 .

3.1.1 Définitions

Définition 3.1.1 Soit un intervalle $I =]a, b[$ contenant le point x_0 et soit une fonction f définie sur I . Soit aussi P un polynôme de degré inférieur ou égal à n ($n \in \mathbb{N}$).

On dit que le polynôme $x \mapsto P(x - x_0)$ est une approximation de f à l'ordre n en x_0 lorsque¹

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - P(x - x_0)}{(x - x_0)^n} = 0. \quad (*)$$

Le reste de l'approximation est la fonction

$$R(x) = f(x) - P(x - x_0).$$

1. Si $n \in \mathbb{N}_0$, il est inutile d'imposer $x \neq x_0$; cela est implicitement demandé puisqu'on divise par une puissance non nulle de $x - x_0$.

On écrit aussi

$$f(x) = P(x - x_0) + o((x - x_0)^n)$$

où le reste $o((x - x_0)^n)$ est une fonction qui, après division par $(x - x_0)^n$ tend vers 0 lorsque x tend vers x_0 .

L'interprétation de l'approximation à l'ordre n de f en x_0 est donc celle-ci : "lorsque x est proche de x_0 , $f(x)$ est proche du polynôme $P(x - x_0)$ ", le terme "proche" ayant une définition claire (limite ci-dessus).

Ici, plusieurs remarques concernant la définition peuvent être faites. Elles répondent à des questions naturelles que l'on se pose en voyant la formulation de la définition. Ces remarques se trouvent dans l'annexe et le lecteur curieux ne manquera pas d'y jeter un coup d'oeil.

3.1.2 Propriétés

Voici quelques propriétés des fonctions qui possèdent une approximation.

Propriété 3.1.2 Soit f une fonction définie sur $]a, b[$ et soit $x_0 \in]a, b[$.

1) Si f a une approximation à l'ordre n en x_0 , alors f a une approximation à tout ordre inférieur à n en x_0 .

2) Si f admet une approximation à l'ordre n en x_0 , cette approximation est unique.

3) Si f est continu en x_0 et admet une approximation à l'ordre 1 alors f est dérivable en x_0 et l'approximation à l'ordre 1 est

$$f(x_0) + (x - x_0)Df(x_0).$$

4) Si f est continu en x_0 et admet une approximation à l'ordre 2 en x_0 , f n'est pas nécessairement deux fois dérivable en x_0 .

Preuve. Les démonstrations figurent dans l'annexe. $\square$

3.1.3 Recherche de la forme de l'approximation.

FORME GÉNÉRALE QUAND LA FONCTION EST ASSEZ DÉRIVABLE

Les propriétés précédentes sont nécessaires à l'existence d'une approximation. Mais bien sûr, c'est surtout le calcul, la forme explicite de cette approximation qui est utile !

Nous allons donner un résultat généralisant celui de l'introduction, c'est-à-dire le résultat dans lequel nous avons vu qu'une fonction dérivable en un point possédait une approximation à l'ordre 1 en ce point et que cette approximation était donnée par

$$f(x_0) + (x - x_0)Df(x_0).$$

Théorème 3.1.3 (Approximation taylorienne) Soit un naturel strictement positif n . Si f est une fonction n fois dérivable dans $]a, b[$ alors, pour tout $x_0 \in]a, b[$, on a ²

$$\lim_{x \rightarrow x_0} \frac{f(x) - \sum_{j=0}^n \frac{D^j f(x_0)}{j!} (x - x_0)^j}{(x - x_0)^n} = 0.$$

L'approximation à l'ordre n de f en x_0 est donc

$$P(x - x_0) = f(x_0) + (x - x_0)Df(x_0) + \frac{(x - x_0)^2}{2!} D^2 f(x_0) + \dots + \frac{(x - x_0)^n}{n!} D^n f(x_0)$$

ou encore

2. Si on fixe x_0 , il suffit en fait que la fonction soit $n - 1$ fois dérivable sur $]a, b[$ et que $D^{n-1}f$ soit dérivable en $x_0 \in]a, b[$

$$P(x - x_0) = \sum_{j=0}^n \frac{D^j f(x_0)}{j!} (x - x_0)^j$$

Preuve. Si $n = 1$, on a directement

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - (x - x_0)Df(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} \right) - Df(x_0) = 0.$$

Si $n = 2$, on procède comme suit. Posons

$$g(x) = \sum_{j=0}^2 \frac{D^j f(x_0)}{j!} (x - x_0)^j = f(x_0) + (x - x_0)Df(x_0) + \frac{(x - x_0)^2}{2} D^2 f(x_0).$$

On a

$$\lim_{x \rightarrow x_0} (f(x) - g(x)) = 0 \quad \text{et} \quad \lim_{x \rightarrow x_0} (x - x_0)^2 = 0.$$

Comme

$$\begin{aligned} \lim_{x \rightarrow x_0} \frac{Df(x) - Dg(x)}{D(x - x_0)^2} &= \lim_{x \rightarrow x_0} \frac{Df(x) - Df(x_0) - (x - x_0)D^2 f(x_0)}{2(x - x_0)} \\ &= \frac{1}{2} \lim_{x \rightarrow x_0} \frac{Df(x) - Df(x_0)}{x - x_0} - \frac{1}{2} D^2 f(x_0) \\ &= 0 \end{aligned}$$

le théorème de l'Hospital donne

$$\lim_{x \rightarrow x_0} \frac{f(x) - g(x)}{(x - x_0)^2} = 0$$

c'est-à-dire la thèse.

En toute généralité, pour $n > 1$, on procède comme précédemment mais en appliquant $n - 1$ fois le théorème de l'Hospital. $\square$

CAS DES APPROXIMATIONS D'ORDRE 1 (APPROXIMATIONS DITES LINÉAIRES)

Pour une fonction f vérifiant les hypothèses du résultat précédent avec $n = 1$, rappelons que l'on dit que

$$P(x - x_0) = f(x_0) + (x - x_0)Df(x_0)$$

est

l'approximation linéaire

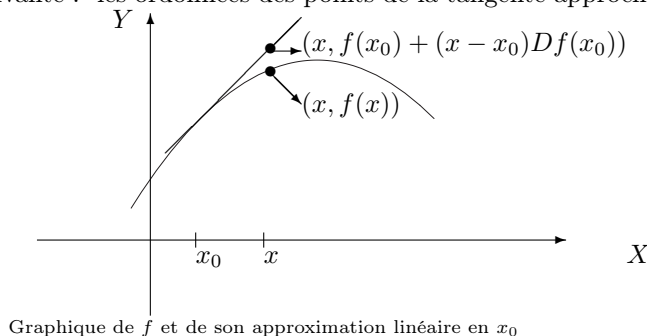
de f en x_0 . Rappelons aussi que l'on a défini la

tangente au graphique de f en x_0

comme étant la droite d'équation cartésienne

$$y = f(x_0) + (x - x_0)Df(x_0).$$

On a donc l'interprétation suivante : "les ordonnées des points de la tangente approchent les valeurs de f à l'ordre 1 en x_0 ".



CAS DES APPROXIMATIONS D'ORDRE 2 (APPROXIMATIONS DITES QUADRATIQUES)

Pour une fonction f vérifiant les hypothèses du résultat précédent avec $n = 2$, on dit que

$$P(x - x_0) = f(x_0) + (x - x_0)Df(x_0) + \frac{(x - x_0)^2}{2}D^2f(x_0)$$

est

l'approximation quadratique

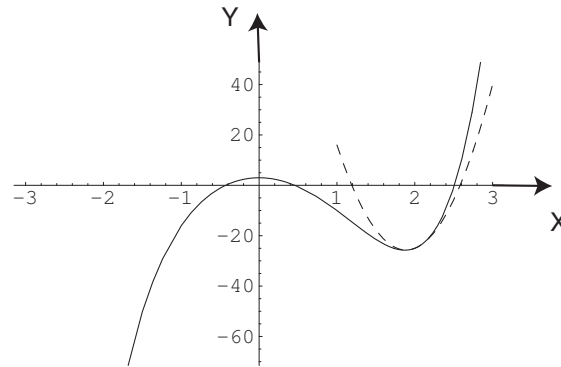
de f en x_0 . Le graphique de

$$g(x) = f(x_0) + (x - x_0)Df(x_0) + \frac{(x - x_0)^2}{2}D^2f(x_0),$$

lorsque $D^2f(x_0) \neq 0$, est une

parabole d'axe parallèle à Y passant par le point de coordonnées $(x_0, f(x_0))$.

On a donc l'interprétation suivante : “les ordonnées des points de la parabole d'équation cartésienne $y = f(x_0) + (x - x_0)Df(x_0) + \frac{(x - x_0)^2}{2}D^2f(x_0)$ approchent les valeurs de f à l'ordre 2 en x_0 ”.



Graphique de f et de son approximation (en pointillés) à l'ordre 2 en $x_0 = 2$

AUTRES ORDRES D'APPROXIMATION

Pour une fonction f vérifiant les hypothèses du résultat précédent avec $p > 2$, on obtient aussi “un polynôme dont les valeurs approchent les valeurs de f à l'ordre p en x_0 ”.

3.1.4 Retour aux polynômes

Le cas où la fonction f est elle-même un polynôme conduit à un résultat précis, que nous énonçons et démontrons ci-dessous.

Proposition 3.1.4 Soit $P : x \mapsto P(x) = a_0 + a_1x + \dots + a_nx^n$ un polynôme de degré³ n . Pour tout naturel N plus grand ou égal à n , l'approximation à l'ordre N en x_0 est le polynôme P lui-même et on a

$$P(x) = P(x_0) + (x - x_0)DP(x_0) + \dots + \frac{(x - x_0)^n}{n!}D^nP(x_0), \quad x \in \mathbb{R}.$$

Preuve. La fonction P est indéfiniment continûment dérivable sur $\mathbb{R}$. Etant donné les réels x et x_0 , le développement limité de Taylor à l'ordre $N + 1$ fournit donc un réel u compris entre x et x_0 tel que

$$P(x) = P(x_0) + (x - x_0)DP(x_0) + \dots + \frac{(x - x_0)^N}{N!}D^NP(x_0) + \frac{(x - x_0)^{N+1}}{(N+1)!}D^{N+1}P(u) \quad (*)$$

3. on a donc $a_n \neq 0$.

Par ailleurs, comme $D^k P$ est la fonction nulle lorsque k est strictement supérieur à n , on a toujours

$$(D^{N+1}P)(u) = 0$$

et lorsque $N > n$, les derniers termes⁴ de cette expression (*) sont nuls également.

On obtient donc

$$P(x) = P(x_0) + (x - x_0)DP(x_0) + \dots + \frac{(x - x_0)^n}{n!}D^n P(x_0), \quad x \in \mathbb{R}.$$

□

Ce résultat permet de donner un critère pratique pour la recherche de la multiplicité des zéros d'un polynôme. Il s'agit du corollaire suivant.

Corollaire 3.1.5 *Soit P un polynôme de degré $n \in \mathbb{N}_0$. Le réel x_0 est un zéro de P de multiplicité $\alpha \in \mathbb{N}_0$ si et seulement si*

$$D^k P(x_0) = 0, \quad k = 0, \dots, \alpha - 1, \quad D^\alpha P(x_0) \neq 0.$$

Preuve. Par définition, le réel x_0 est un zéro de multiplicité α de P s'il existe un polynôme Q tel que

$$P(x) = (x - x_0)^\alpha Q(x), \quad \text{et} \quad Q(x_0) \neq 0.$$

Supposons que x_0 soit un zéro de multiplicité α pour P et calculons les dérivées successives de P . Par la formule de Leibniz, on a

$$D^k P(x) = \sum_{j=0}^k C_k^j D^j (x - x_0)^\alpha D^{k-j} Q(x).$$

En prenant $k \in \{0, \dots, \alpha - 1\}$ et en considérant cette relation en x_0 , on obtient

$$D^k P(x_0) = 0$$

car, si $j \leq \alpha - 1$, la dérivée d'ordre j de $(x - x_0)^\alpha$ comporte au moins un facteur $(x - x_0)$, lequel s'annule pour $x = x_0$. Par contre, on a

$$D^\alpha P(x_0) = \alpha! Q(x_0) \neq 0.$$

Supposons à présent que les dérivées de P en x_0 soient nulles jusqu'à l'ordre $\alpha - 1$ et que la dérivée d'ordre α en x_0 diffère de 0. Vu le développement de Taylor pour P (cf 3.1.4), on a

$$\begin{aligned} P(x) &= \frac{(x - x_0)^\alpha}{\alpha!} D^\alpha P(x_0) + \dots + \frac{(x - x_0)^n}{n!} D^n P(x_0) \\ &= (x - x_0)^\alpha Q(x) \end{aligned}$$

avec

$$Q(x_0) = \frac{D^\alpha P(x_0)}{\alpha!} \neq 0.$$

□

3.1.5 Exemples des fonctions sin et cos

A titre d'exemples d'approximations, recherchons les approximations à l'ordre n des fonctions sin et cos en 0 pour $n = 0, 1, 2, 3$.

Ces fonctions sont indéfiniment continûment dérivables sur $\mathbb{R}$. De plus, on a

$$D \sin x = \cos x, \quad D^2 \sin x = -\sin x, \quad D^3 \sin x = -\cos x$$

4. le nombre de ces termes dépend bien sûr de N

et

$$D \cos x = -\sin x, \quad D^2 \cos x = -\cos x, \quad D^3 \cos x = \sin x$$

donc

$$D \sin(0) = 1, \quad D^2 \sin(0) = 0, \quad D^3 \sin(0) = -1$$

et

$$D \cos(0) = 0, \quad D^2 \cos(0) = -1, \quad D^3 \cos(0) = 0.$$

Il s'ensuit que

les approximations à l'ordre 0, 1, 2, 3 de sin

sont successivement

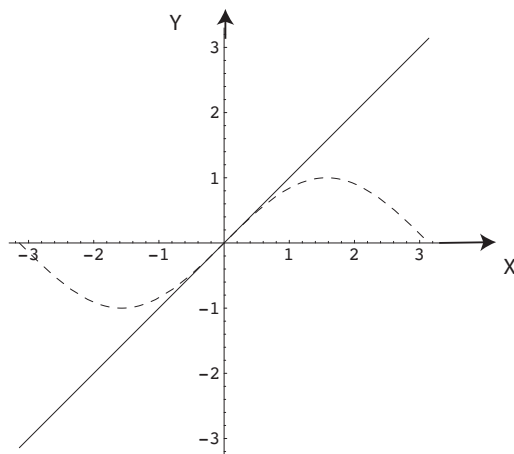
$$P_0(x) = 0, \quad P_1(x) = x, \quad P_2(x) = x, \quad P_3(x) = x - \frac{x^3}{6}$$

et que

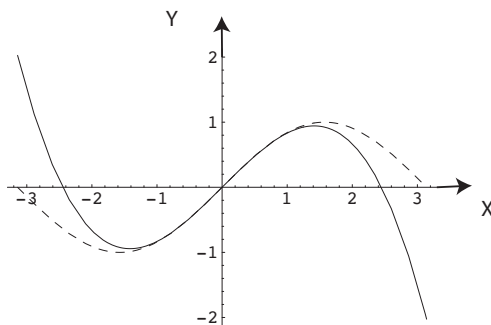
les approximations à l'ordre 0, 1, 2, 3 de cos

sont successivement

$$P_0(x) = 1, \quad P_1(x) = 1, \quad P_2(x) = 1 - \frac{x^2}{2}, \quad P_3(x) = 1 - \frac{x^2}{2}.$$



Graphiques de $\sin x$, ($x \in [-\pi, \pi]$) (en pointillés) et de $P(x) = x$, ($x \in [-\pi, \pi]$).



Graphiques de $\sin x$, ($x \in [-\pi, \pi]$) (en pointillés) et de $P(x) = x - x^3/6$, ($x \in [-\pi, \pi]$).

Plus généralement, on a

$$D^{2n} \sin x = (-1)^n \sin x, \quad D^{2n+1} \sin x = (-1)^n \cos x$$

pour tout $x \in \mathbb{R}$ et tout naturel n positif ou nul. L'approximation à l'ordre $2n + 1$ de $\sin$ en 0 est donc un polynôme de degré $2n + 1$, à savoir

$$P_{2n+1}(x) = \sum_{k=0}^n \frac{x^{2k+1}}{(2k+1)!} D^{2k+1} \sin(0) = \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!}$$

et l'approximation à l'ordre $2n$ de $\sin$ en 0 est un polynôme de degré $2n - 1$, à savoir

$$P_{2n}(x) = \sum_{k=0}^{n-1} \frac{x^{2k+1}}{(2k+1)!} D^{2k+1} \sin(0) = \sum_{k=0}^{n-1} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = P_{2n-1}(x).$$

On procède de même avec la fonction $\cos$. On a

$$D^{2n} \cos x = (-1)^n \cos x, \quad D^{2n+1} \cos x = (-1)^{n+1} \sin x$$

pour tout $x \in \mathbb{R}$ et tout naturel n positif ou nul. L'approximation à l'ordre $2n$ de $\cos$ en 0 est donc un polynôme de degré $2n$, à savoir

$$P_{2n}(x) = \sum_{k=0}^n \frac{x^{2k}}{(2k)!} D^{2k} \cos(0) = \sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!}$$

et l'approximation à l'ordre $2n + 1$ de $\cos$ en 0 est un polynôme de degré $2n$, à savoir

$$P_{2n+1}(x) = \sum_{k=0}^n \frac{x^{2k}}{(2k)!} D^{2k} \cos(0) = \sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!} = P_{2n}(x).$$

3.1.6 Estimation du reste

Le résultat qui suit est une généralisation du théorème des accroissements finis ; il correspond en effet à ce théorème lorsque $n = 1$.

Théorème 3.1.6 (Développement limité de Taylor) *Soient n un naturel strictement positif et f une fonction réelle n fois dérivable sur $]a, b[$. Pour tous $x, x_0 \in]a, b[$, $x \neq x_0$ il existe un point u_0 compris strictement entre x_0 et x tel que*

$$f(x) = \sum_{j=0}^{n-1} \frac{(x-x_0)^j}{j!} D^j f(x_0) + \frac{(x-x_0)^n}{n!} D^n f(u_0).$$

Si $x = x_0$ le résultat est encore vrai avec u_0 quelconque dans l'intervalle (en particulier égal à x_0).

Preuve. Pour $n = 1$, il s'agit du théorème des accroissements finis.

Pour démontrer ce résultat lorsque $n > 1$, on procède de la manière suivante. De manière analogue à ce qui a été fait pour la preuve du TAF, définissons la constante k par

$$k = \left(f(x) - f(x_0) - (x-x_0)Df(x_0) - \frac{(x-x_0)^2}{2!} D^2 f(x_0) - \dots - \frac{(x-x_0)^{n-1}}{(n-1)!} D^{n-1} f(x_0) \right) \frac{n!}{(x-x_0)^n}$$

et la fonction F dans l'intervalle fermé d'extrémités x_0, x par

$$F(u) = f(x) - f(u) - (x-u)Df(u) - \frac{(x-u)^2}{2!} D^2 f(u) - \dots - \frac{(x-u)^{n-1}}{(n-1)!} D^{n-1} f(u) - \frac{(x-u)^n}{n!} k.$$

Vu les hypothèses sur f , cette fonction est continue sur l'intervalle fermé d'extrémités x_0, x , dérivable sur l'intervalle ouvert correspondant. De plus, on a

$$F(x) = 0, \quad F(x_0) = 0.$$

5. La preuve du TAF, c'est-à-dire le cas $n = 1$, peut aussi être effectuée en suivant la même preuve

Le théorème de Rolle donne donc un point u_0 strictement compris entre x_0 et x tel que

$$D_u F(u_0) = 0.$$

Comme on a

$$D_u \left(\frac{(x-u)^j}{j!} D_u^j f(u) \right) = -\frac{(x-u)^{j-1}}{(j-1)!} D_u^j f(u) + \frac{(x-u)^j}{j!} D_u^{j+1} f(u)$$

pour $j = 1, \dots, n-1$, on obtient

$$D_u F(u) = -\frac{(x-u)^{n-1}}{(n-1)!} D_u^n f(u) + \frac{(x-u)^{n-1}}{(n-1)!} k = \frac{(x-u)^{n-1}}{(n-1)!} (-D_u^n f(u) + k)$$

donc

$$k = D^n f(u_0);$$

dès lors la thèse est démontrée. $\square$

Sous les hypothèses du résultat précédent (développement limité de Taylor), on obtient donc, si $P_n(x - x_0)$ désigne l'approximation à l'ordre n de f en x_0 :

$$f(x) = P_n(x - x_0) + \frac{(x - x_0)^n}{n!} (D^n f(u_0) - D^n f(x_0)).$$

Le reste $R_n(x) = f(x) - P(x - x_0)$ de l'approximation à l'ordre n s'exprime donc sous la forme

$$R_n(x) = \frac{(x - x_0)^n}{n!} (D^n f(u_0) - D^n f(x_0))$$

et peut être estimé lorsqu'on connaît une estimation pour la dérivée d'ordre n de f .

Lorsque la fonction est $n + 1$ fois dérivable dans l'intervalle, le développement de Taylor, utilisé avec $n + 1$, donne

$$f(x) = f(x_0) + (x - x_0) Df(x_0) + \frac{(x - x_0)^2}{2!} D^2 f(x_0) + \dots + \frac{(x - x_0)^n}{n!} D^n f(x_0) + \frac{(x - x_0)^{n+1}}{(n+1)!} D^{n+1} f(u_1)$$

donc le reste de l'approximation à l'ordre n est

$$R_n(x) = \frac{(x - x_0)^{n+1}}{(n+1)!} D^{n+1} f(u_1)$$

ce qui donne aussi lieu à des estimations.

Par exemple, pour la fonction $\sin$,
- le reste de l'approximation à l'ordre $2n$ en 0 s'écrit

$$R_{2n}(x) = \frac{x^{2n+1}}{(2n+1)!} D^{2n+1} \sin(u_0) = \frac{x^{2n+1}}{(2n+1)!} (-1)^n \cos(u_0)$$

donc

$$|R_{2n}(x)| \leq \frac{|x|^{2n+1}}{(2n+1)!};$$

- le reste de l'approximation à l'ordre $2n + 1$ en 0 s'écrit

$$R_{2n+1}(x) = \frac{x^{2n+2}}{(2n+2)!} D^{2n+2} \sin(u_0) = \frac{x^{2n+2}}{(2n+2)!} (-1)^{n+1} \sin(u_0);$$

comme la fonction $\sin$ est petite au voisinage de 0, on s'attend à pouvoir estimer le reste bien mieux qu'en $|x|^{2n+2}$; comme l'approximation de $\sin$ à l'ordre $2n+2$ en 0 est la même que l'approximation à l'ordre $2n+1$, en utilisant le développement de Taylor pour $2n+3$, on obtient

$$R_{2n+1}(x) = R_{2n+2}(x) = \frac{x^{2n+3}}{(2n+3)!} D^{2n+3} \sin(u_0) = \frac{x^{2n+3}}{(2n+3)!} (-1)^{n+1} \cos(u_0)$$

et ainsi

$$|R_{2n+1}(x)| \leq \frac{|x|^{2n+3}}{(2n+3)!}.$$

Revenons à un résultat cité précédemment (dans l'étude des extrema) et dont la preuve était annoncée dans cette partie.

Propriété 3.1.7 *Si f est deux fois continûment dérivable dans $]a, b[$, si $x_0 \in]a, b[$ et tel que $Df(x_0) = 0$, $D^2f(x_0) > 0$ (resp. $Df(x_0) = 0$, $D^2f(x_0) < 0$) alors x_0 est un minimum (resp. maximum) local strict de f dans $]a, b[$.*

La réciproque de cette propriété est fausse.

Preuve. Supposons $D^2f(x_0) > 0$. D'une part, la continuité de D^2f en x_0 implique l'existence de $r > 0$ tel que $D^2f(x) > 0$ pour tout $x \in]x_0 - r, x_0 + r[$. D'autre part, en utilisant le développement de Taylor, on obtient aussi que, pour tout x dans $]x_0 - r, x_0 + r[$, il existe u_0 compris entre x_0 et x tel que

$$f(x) = f(x_0) + (x - x_0)Df(x_0) + \frac{(x - x_0)^2}{2} D^2f(u_0)$$

donc tel que

$$f(x) - f(x_0) = \frac{(x - x_0)^2}{2} D^2f(u_0)$$

vu l'hypothèse sur f . On obtient donc

$$f(x) > f(x_0) \quad \forall x \in]x_0 - r, x_0 + r[, \quad x \neq x_0$$

ce qui signifie que x_0 est un minimum local strict de f dans $]a, b[$.

Une preuve tout à fait analogue conduit au résultat concernant le maximum.

Pour se convaincre que la réciproque est fausse, il suffit de prendre l'exemple de la fonction $f(x) = x^4$.

□

3.2 Développement illimités-Séries

3.2.1 Rappels sur les suites numériques

3.2.2 Introduction aux développements illimités

Tout ce qui précède concerne les *développements limités* de f , c'est-à-dire qu'on suppose la fonction n fois dérivable sur un intervalle et l'on "approche" f par un polynôme de degré inférieur ou égal à n .

Que se passe-t-il si $f \in C_\infty(]a, b[)$ et que l'on considère des approximations de degré de plus en plus élevé? Vu le développement de Taylor, f admet une approximation à tous les ordres. Etant donné $x_0 \in]a, b[$, pour tout $x \in]a, b[$ et pour tout $N \in \mathbb{N}_0$, il existe donc un point $u(x_0, x, N)$ compris entre x et x_0 et tel que

$$f(x) = \sum_{n=0}^N \frac{(x - x_0)^n}{n!} D^n f(x_0) + R_N(x_0, x)$$

où le reste R_N à l'ordre N s'écrit

$$R_N(x_0, x) = \frac{(x - x_0)^{N+1}}{(N+1)!} D^{N+1} f(u(x_0, x, N)).$$

Si

$$\lim_{N \rightarrow +\infty} R_N(x_0, x) = 0 \quad (*)$$

on a

$$f(x) = \lim_{N \rightarrow +\infty} \left(\sum_{n=0}^N \frac{(x - x_0)^n}{n!} D^n f(x_0) + R_N(x_0, x) \right) = \lim_{N \rightarrow +\infty} \sum_{n=0}^N \frac{(x - x_0)^n}{n!} D^n f(x_0)$$

ce qui se note habituellement

$$f(x) = \sum_{n=0}^{+\infty} \frac{(x - x_0)^n}{n!} D^n f(x_0).$$

Les fonctions qui sont indéfiniment continûment dérivables et admettent un tel développement au voisinage de x_0 sont dites *analytiques* en x_0 ; les fonctions qui sont indéfiniment continûment dérivables et admettent un tel développement au voisinage de tout $x_0 \in]a, b[$ sont dites analytiques dans $]a, b[$. Si la plupart des fonctions élémentaires sont de ce type, il est faux de dire que toutes les fonctions sont analytiques⁶.

3.2.3 Séries (définition)

Dans ce qui précède, on a déjà considéré des “sommes infinies”, c'est-à-dire des expressions du type $1 + x + x^2 + x^3 + \dots$, qui sont plutôt notées $\sum_{m=0}^{+\infty} x^m$ ou $\sum_{m=m_0}^{+\infty} x^m$ ($m_0 \in \mathbb{N}$ fixé). Il faut être prudent lors de la manipulation de telles expressions car en fait la notation sous-entend un passage à la limite et on sait qu'une telle opération peut amener des surprises.

Une série n'est en fait rien d'autre qu'une suite particulière.

Définition 3.2.1 *Etant donné une suite de réels (ou de complexes) x_m ($m \in \mathbb{N}_0$), la série de terme général x_m est la suite des sommes partielles*

$$\sum_{m=1}^M x_m \quad (M \in \mathbb{N}_0).$$

Une série est dite convergente si la suite des sommes partielles converge vers une limite finie. Cette limite est alors notée

$$\sum_{m=1}^{+\infty} x_m$$

et on dit que c'est la somme de la série.

Une série est dite divergente si la suite des sommes partielles ne converge pas vers une limite finie.

Etant donné une suite de nombres x_m ($m \in \mathbb{N}_0$), la série de terme général x_m est donc la suite

$$S_M = \sum_{m=1}^M x_m \quad (M \in \mathbb{N}_0);$$

6. La fonction

$$f(x) = \begin{cases} \exp(\frac{-1}{x}) = e^{\frac{-1}{x}} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

appartient à $C_\infty(\mathbb{R})$ et ses dérivées sont toutes nulles en 0. On ne peut donc pas avoir un tel développement au voisinage de 0.

il s'agit vraiment d'une suite très particulière car il y a des liens étroits entre ses différents éléments. On a en effet

$$S_{M+1} = S_M + x_{M+1}$$

quel que soit le naturel M .

Par définition, si la série de terme général x_m converge, on a donc

$$\lim_{M \rightarrow +\infty} \sum_{m=1}^M x_m = \lim_{M \rightarrow +\infty} S_M = \sum_{m=1}^{+\infty} x_m.$$

Attention : bien souvent, on parle aussi de la série $\sum_{m=1}^{+\infty} x_m$. Cela signifie que l'on doit étudier la suite des sommes partielles $\sum_{m=1}^M x_m$ ($M \in \mathbb{N}_0$) et que, si cette suite converge vers une limite finie, cette limite est notée $\sum_{m=1}^{+\infty} x_m$. Il faut donc bien prendre garde à l'utilisation de cette notation : soit on parle de la limite, soit on parle de la suite des sommes partielles. Tout cela doit être clair dans le contexte.

Pour tout $m_0 \in \mathbb{N}_0$, $m_0 \neq 1$, on peut aussi bien sûr considérer la suite des sommes partielles S_{m_0+m} ($m \in \mathbb{N}$). La série que l'on étudie dans ce cas est la série

$$\sum_{m=m_0}^{+\infty} x_m.$$

De plus, la série $\sum_{m=m_0}^{+\infty} x_m$ converge si et seulement si la série $\sum_{m=1}^{+\infty} x_m$ converge, auquel cas on a l'égalité suivante entre les limites

$$x_1 + \dots + x_{m_0-1} + \sum_{m=m_0}^{+\infty} x_m = \sum_{m=1}^{+\infty} x_m.$$

L'étude de la convergence des séries repose sur plusieurs résultats dont la démonstration sort du cadre de ce cours. Nous ne citerons que ceux qui seront le plus utilisés ici.

3.2.4 Exemples

1) Considérons par exemple la série de terme général $x_m = (-1)^m$ en prenant m à partir de 0. Il s'agit de la suite

$$S_M = \sum_{m=0}^M (-1)^m \quad (M \in \mathbb{N})$$

c'est-à-dire de la suite de réels $S_0, S_1, S_2, \dots$ suivante

$$\begin{aligned} S_0 &= (-1)^0 = 1 \\ S_1 &= S_0 + (-1)^1 = 1 - 1 = 0 \\ S_2 &= S_1 + (-1)^2 = 1 \\ S_3 &= S_2 + (-1)^3 = 1 - 1 = 0 \\ S_4 &= S_3 + (-1)^4 = 0 + 1 = 1 \\ &\vdots \end{aligned}$$

Cette suite ne converge pas vers une limite finie, ni vers l'infini.

2) Considérons la série de terme général $x_m = \frac{1}{2^m}$ en prenant m à partir de 1. Il s'agit de la suite

$$S_M = \sum_{m=1}^M \frac{1}{2^m} \quad (M \in \mathbb{N}_0)$$

c'est-à-dire de la suite de réels

$$S_1 = \frac{1}{2^1} = \frac{1}{2}$$

$$S_2 = S_1 + \frac{1}{2^2} = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

$$S_3 = S_2 + \frac{1}{2^3} = \frac{3}{4} + \frac{1}{8} = \frac{7}{8}$$

$$S_4 = S_3 + \frac{1}{2^4} = \frac{7}{8} + \frac{1}{16} = \frac{15}{16}$$

$\vdots$

Etudier la convergence de cette série peut se faire comme suit.

On a

$$S_M = \frac{1}{2} \sum_{m=0}^{M-1} \frac{1}{2^m} = \frac{1}{2} \frac{1 - \frac{1}{2^M}}{1 - \frac{1}{2}} = 1 - \frac{1}{2^M}$$

pour tout M . Comme la suite $\frac{1}{2^M}$ ($M \in \mathbb{N}_0$) converge vers 0 on en déduit que

$$\lim_{M \rightarrow +\infty} S_M = \sum_{m=1}^{+\infty} \frac{1}{2^m} = 1.$$

3) Considérons la série de terme général $x_m = \frac{1}{m}$ en prenant m à partir de 1. Il s'agit de la suite

$$S_M = \sum_{m=1}^M \frac{1}{m} \quad (M \in \mathbb{N}_0)$$

c'est-à-dire de la suite de réels

$$S_1 = \frac{1}{1} = 1$$

$$S_2 = S_1 + \frac{1}{2} = 1 + \frac{1}{2} = \frac{3}{2}$$

$$S_3 = S_2 + \frac{1}{3} = \frac{3}{2} + \frac{1}{3} = \frac{11}{6}$$

$\vdots$

Cette série ne converge pas : on a

$$\lim_{M \rightarrow +\infty} S_M = +\infty.$$

Démontrons ce résultat ⁷.

Démontrons d'abord que l'on a

$$x \geq \ln(x+1) \quad \forall x > 0.$$

En effet, la fonction $f(x) = x - \ln(x+1)$ est dérivable sur $] -1, +\infty[$. On a $Df(x) = 1 - \frac{1}{x+1} = \frac{x}{x+1} > 0$ pour tout $x \in]0, +\infty[$, donc la fonction f est donc strictement croissante sur $]0, +\infty[$. Il s'ensuit que

$$f(x) = x - \ln(x+1) > \lim_{y \rightarrow 0^+} f(y) = 0, \quad \forall x > 0.$$

Utilisons ce résultat pour toutes les valeurs $x = 1/m$ ($m \in \mathbb{N}_0$) : on a donc

$$\frac{1}{m} \geq \ln\left(\frac{1}{m} + 1\right) = \ln(m+1) - \ln m, \quad \forall m \in \mathbb{N}_0.$$

7. Une autre démonstration consiste à montrer que $S_{2^j} \geq 1 + \frac{j}{2}$ pour tout $j \in \mathbb{N}_0$; voir par exemple *Calculus*, R. Ellis and D. Gulick, 1994, p. 572

Il s'ensuit que

$$S_M = \sum_{m=1}^M \frac{1}{m} \geq \sum_{m=1}^M (\ln(m+1) - \ln m) \geq \ln(M+1), \quad \forall M \in \mathbb{N}_0.$$

Dès lors

$$\lim_{M \rightarrow +\infty} S_M = +\infty.$$

4) Considérons la série de terme général $x_m = \frac{(-1)^m}{m}$ en prenant m à partir de 1. Il s'agit de la suite

$$S_M = \sum_{m=1}^M \frac{(-1)^m}{m} \quad (M \in \mathbb{N}_0)$$

c'est-à-dire de la suite de réels

$$\begin{aligned} S_1 &= \frac{(-1)^1}{1} = -1 \\ S_2 &= S_1 + \frac{(-1)^2}{2} = -1 + \frac{1}{2} = -\frac{1}{2} \\ S_3 &= S_2 + \frac{(-1)^3}{3} = -\frac{1}{2} - \frac{1}{3} = -\frac{5}{6} \\ &\vdots \end{aligned}$$

Contrairement à la série précédente, les éléments successifs de la suite S_M ($M \in \mathbb{N}_0$) sont obtenus en ajoutant, puis retranchant un nombre positif de plus en plus petit. Ce comportement est typique des séries dites “alternées”. S'il n'est pas aisé de trouver la valeur de la limite, on démontre cependant qu'une telle série est convergente (voir la suite de ce chapitre). On démontre que (voir les annexes du chapitre 4 du cours A du premier quadrimestre et voir aussi le développement en série de puissances de la fonction logarithme dans ce qui suit)

$$\sum_{m=1}^{+\infty} \frac{(-1)^m}{m} = -\ln 2.$$

3.2.5 Cas de référence

Deux exemples de séries sont fondamentaux. L'étude de la convergence de beaucoup d'autres les utilise. C'est la raison pour laquelle on parle de “cas de référence” en parlant de ces deux exemples.

Le premier cas fondamental est celui de la série géométrique. Cette série (on devrait plutôt parler de “séries” au pluriel car on la définit pour tout réel ou complexe) est définie de la manière suivante : pour $q \in \mathbb{R}$, il s'agit de la série de terme général q^m , c'est-à-dire de la suite des sommes partielles

$$S_M = \sum_{m=0}^M q^m, \quad M \in \mathbb{N}.$$

Cette série est notée

$$\sum_{m=0}^{+\infty} q^m.$$

Le second cas fondamental est celui de la série de Riemann. Cette série (on devrait plutôt parler de “séries” au pluriel car on la définit pour tout réel ou complexe) est définie de la manière suivante : pour $\alpha \in \mathbb{R}$, il s'agit de la série de terme général $\frac{1}{m^\alpha}$, c'est-à-dire de la suite des sommes partielles

$$S_M = \sum_{m=1}^M \frac{1}{m^\alpha}, \quad M \in \mathbb{N}_0.$$

Cette série est notée

$$\sum_{m=1}^{+\infty} \frac{1}{m^\alpha}.$$

Le résultat régissant la convergence de ces séries est le suivant.

Théorème 3.2.2 1) La série géométrique de terme général q^m converge si et seulement si $q \in]-1, 1[$. De plus, on a

$$\sum_{m=0}^{+\infty} q^m = \frac{1}{1-q} \quad q \in]-1, 1[$$

et même

$$\sum_{m=M}^{+\infty} q^m = q^M \frac{1}{1-q} \quad q \in]-1, 1[, \quad M \in \mathbb{N}$$

2) La série de Riemann de terme général $\frac{1}{m^\alpha}$ converge si et seulement si $\alpha > 1$.

Preuve. 1) Si $q = 1$, la suite des sommes partielles de la série géométrique est la suite $S_M = M+1$ ($M \in \mathbb{N}$). Cette suite ne converge pas vers une limite finie.

Si $q \neq 1$, la suite des sommes partielles de la série géométrique est la suite formée à partir de la somme des premiers termes d'une progression géométrique de raison q . On a donc

$$S_M = \frac{1 - q^{M+1}}{1 - q} \quad M \in \mathbb{N}.$$

Ainsi, la suite S_M ($M \in \mathbb{N}$) converge vers une limite finie si et seulement si la suite q^{M+1} ($M \in \mathbb{N}$) converge vers une limite finie. L'étude de la convergence de cette suite a été effectuée dans le chapitre 2 (partim A) (lors de l'étude d'exemples fondamentaux de suites). On reprend les résultats établis à ce moment-là : la suite q^{M+1} ($M \in \mathbb{N}$), converge vers une limite finie si et seulement si $q \in]-1, 1]$. Comme nous sommes dans le cas $q \neq 1$, on conclut.

2) Résultat admis. $\square$

3.2.6 Conditions de convergence

Propriété 3.2.3 Si la série $\sum_{m=1}^{+\infty} x_m$ converge, alors la suite formée par le terme général, c'est-à-dire la suite x_m ($m \in \mathbb{N}_0$), converge vers 0. La réciproque est fausse.

Preuve. Notons l la limite de la série. Soit $S_M = \sum_{m=1}^M x_m$ ($M \in \mathbb{N}_0$) la suite des sommes partielles. On a

$$x_1 = S_1, \quad x_M = S_M - S_{M-1}, \quad M \geq 2.$$

Comme les suites S_M ($M \in \mathbb{N}_0, M \geq 2$) et S_{M-1} ($M \in \mathbb{N}_0, M \geq 2$) convergent vers l , la suite $x_M = S_M - S_{M-1}$, $M \geq 2$ converge vers 0.

La réciproque est fausse comme le montre l'exemple de la série $\sum_{m=1}^{+\infty} \frac{1}{m}$ appelée série harmonique $\square$

Trouver la limite d'une série n'est pas chose aisée. Cependant, il existe des critères permettant de savoir si, oui ou non, une série est convergente. Nous renvoyons à l'annexe pour le *critère de Cauchy* qui n'est rien d'autre que le critère de Cauchy pour les suites.

Proposition 3.2.4 1) **Critère de comparaison.** Si la série de terme général $|x_m|$ converge et si le réel strictement positif R est tel que $|y_m| \leq R|x_m| \quad \forall m$, alors la série de terme général y_m est aussi une série convergente.

2) **Critère des séries alternées.** Si r_m ($m \in \mathbb{N}_0$) est une suite de réels qui décroît vers 0, alors la série de terme général $(-1)^m r_m$ est une série convergente. De plus, pour tout $p \in \mathbb{N}_0$, on a

$$\left| \sum_{m=p}^{+\infty} (-1)^m r_m \right| \leq r_p.$$

Preuve. Les preuves figurent dans l'annexe. $\square$

Remarquons que la première propriété énoncée ci-dessus indique également que si la série $\sum_{m=1}^{+\infty} |x_m|$ converge, alors la série $\sum_{m=1}^{+\infty} x_m$ converge également.

Les notions de série absolument convergente et semi-convergente, de même que les critères pratiques de convergence des séries ne seront pas abordés dans un premier temps. Nous renvoyons à l'annexe pour une première information sur ces sujets.

Voici quelques exemples.

- La série $\sum_{m=1}^{+\infty} 2^m$ ne converge pas.
- La série $\sum_{m=1}^{+\infty} 3^{-m}$ converge (est même absolument convergente).
- La série $\sum_{m=1}^{+\infty} (-1)^m / m^2$ converge (est même absolument convergente).
- La série $\sum_{m=1}^{+\infty} \frac{1}{m+1}$ ne converge pas.
- La série $\sum_{m=1}^{+\infty} \frac{(-1)^m}{m+1}$ converge.
- La série $\sum_{m=1}^{+\infty} \frac{1}{m^2+1}$ converge (est même absolument convergente).
- La série $\sum_{m=1}^{+\infty} \frac{\sqrt{m}}{m+1}$ ne converge pas.
- La série $\sum_{m=1}^{+\infty} \frac{(-1)^m}{m^m}$ converge (est absolument convergente).
- La série $\sum_{m=1}^{+\infty} \frac{(-2)^m}{m!}$ converge (est absolument convergente).

3.2.7 Séries de puissances

Dans le cas des développements illimités de Taylor, les séries qui apparaissent sont particulières en ce sens qu'elles s'écrivent sous la forme

$$\sum_{m=0}^{+\infty} a_m (x - x_0)^m$$

où a_m , ($m \in \mathbb{N}$), est une suite de réels (ou de complexes) et où $x, x_0 \in \mathbb{R}$ (ou $\in \mathbb{C}$). Une telle série s'appelle une série de puissances de $x - x_0$. Si on fixe x_0 et la suite a_m ($m \in \mathbb{N}$), il s'agit en fait d'une fonction de x dont on doit déterminer le domaine, c'est-à-dire l'ensemble des x pour lesquels la série converge.

Pour l'étude des séries de puissances, on a le résultat fort utile suivant.

Théorème 3.2.5 *On considère la série de puissances*

$$\sum_{m=0}^{+\infty} a_m x^m.$$

Supposons que la série converge pour tout $x \in]-R, R[$ où R est un réel strictement positif. Alors la fonction

$$x \in]-R, R[\mapsto S(x) = \sum_{m=0}^{+\infty} a_m x^m$$

est indéfiniment continûment dérivable dans l'intervalle $] -R, R[$ et y est dérivable terme à terme c'est-à-dire

$$\begin{aligned} DS(x) &= D(a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots) \\ &= Da_0 + a_1 Dx + a_2 Dx^2 + a_3 Dx^3 + \dots = a_1 + 2a_2 x + 3a_3 x^2 + \dots \end{aligned}$$

et ainsi de suite pour les dérivées d'ordre supérieur.

Le théorème précédent peut être aussi énoncé pour une série de puissances de $x - x_0$, à savoir

$$\sum_{m=0}^{+\infty} a_m (x - x_0)^m$$

au lieu de la série de puissances de x

$$\sum_{m=0}^{+\infty} a_m x^m.$$

3.3 Fonction exponentielle (définie par une série)

Nous allons à présent définir la fonction exponentielle et en démontrer les propriétés fondamentales. Normalement, en suivant un processus déductif, nous aurions dû commencer par là pour définir plusieurs des fonctions élémentaires ($\ln$, $\sin$, $\cos$). Cependant, vu la grande utilité des fonctions élémentaires et de leurs propriétés dans les cours de sciences abordés dès le début de l'année, nous avons préféré présenter la matière de cette manière-ci.

3.3.1 Définition

Définition 3.3.1 *La fonction exponentielle est la fonction définie sur $\mathbb{R}$ par*

$$\exp(x) = \sum_{m=0}^{+\infty} \frac{x^m}{m!}, \quad x \in \mathbb{R}$$

On définit de la même manière la fonction $\exp(z)$, avec z nombre complexe quelconque.

Montrons que cette définition a un sens, c'est-à-dire que la série

$$\sum_{m=0}^{+\infty} \frac{x^m}{m!}$$

est une série convergente⁸ quel que soit le réel x .

Soit $x \in \mathbb{R}$. Il existe $M \in \mathbb{N}_0$ tel que $|x| \leq M$. Il s'ensuit que, pour $m > M$, on a

$$\frac{|x|^m}{m!} = \frac{|x|}{1} \cdots \frac{|x|}{M} \frac{|x|}{M+1} \cdots \frac{|x|}{m} \leq \frac{|x|}{1} \cdots \frac{|x|}{M} \left(\frac{|x|}{M+1} \right)^{m-M} \leq C \left(\frac{|x|}{M+1} \right)^m$$

où C est une constante strictement positive qui ne dépend pas de m (cette constante ne dépend que de x et de M). Comme $\frac{|x|}{M+1} \in [0, 1[$, l'expression $\left(\frac{|x|}{M+1} \right)^m$ est le terme général d'une série géométrique convergente. D'où la conclusion.

3.3.2 Propriétés fondamentales

Nous sommes à présent en mesure de démontrer les propriétés fondamentales de la fonction exponentielle, annoncées au chapitre 2 (partim A).

Théorème 3.3.2 1. *La fonction $\exp$ est indéfiniment continûment dérivable sur $\mathbb{R}$ et⁹ on a*

$$D_x^k \exp(x) = \exp(x), \quad x \in \mathbb{R}, \quad k \in \mathbb{N}; \quad \exp(0) = 1.$$

2. *On a la première estimation*

$$2 \leq \exp(1) \leq 3.$$

3. *Plus précisément, on a l'estimation*

$$e = \exp(1) = 2.71828182 \dots$$

8. On peut directement aussi appliquer le critère du quotient. Nous avons utilisé ici une méthode directe car elle permet de donner une justification même si les critères pratiques n'ont pas été abordés au cours.

9. C'est en fait l'unique fonction dérivable sur $\mathbb{R}$ qui vérifie ces conditions. La preuve de l'unicité est un cas particulier de la preuve régissant la structure des solutions des équations différentielles linéaires à coefficients constants d'ordre 1.

4. Pour tous $x, y \in \mathbb{R}$, on a

$$\exp(x + y) = \exp(x) \cdot \exp(y).$$

5. Pour tout $x \in \mathbb{R}$, on a

$$\exp(x) > 0.$$

6. On a $\lim_{x \rightarrow +\infty} \exp(x) = +\infty$ et $\lim_{x \rightarrow -\infty} \exp(x) = 0$. Plus généralement, pour tout $p \in \mathbb{N}$, on a

$$\lim_{x \rightarrow +\infty} \frac{\exp(x)}{x^p} = +\infty, \text{ et } \lim_{x \rightarrow -\infty} x^p \exp(x) = 0$$

(“à l’infini, la fonction exponentielle domine toute puissance antagoniste de x ”).

Preuve. 1) On vient de voir que cette fonction est effectivement définie sur $\mathbb{R}$. Elle est en outre dérivable sur $\mathbb{R}$ et dérivable terme à terme, par utilisation du théorème 3.2.5. On a donc

$$D_x \exp(x) = \sum_{m=1}^{+\infty} \frac{mx^{m-1}}{m!} = \sum_{m=0}^{+\infty} \frac{x^m}{m!}$$

pour tout $x \in \mathbb{R}$. Les autres dérivées s’obtiennent alors directement. De plus, comme $0^m = 0$ pour tout $m \in \mathbb{N}_0$, on a $\exp(0) = 1$.

2) Pour tout $m \in \mathbb{N}, m \geq 2$, on a $m! = m \cdot (m-1) \dots 2 \geq 2 \dots 2 = 2^{m-1}$. Ainsi,

$$2 \leq \exp(1) \leq 2 + \sum_{m=2}^{+\infty} \frac{1}{m!} \leq 2 + \sum_{m=2}^{+\infty} 2^{-(m-1)} = 3.$$

3) est démontré dans l’annexe.

4) Soit $y \in \mathbb{R}$. Considérons la fonction de x

$$f(x) = \exp(x + y) \exp(-x).$$

Cette fonction est définie et dérivable sur $\mathbb{R}$ car la fonction exponentielle est définie et dérivable sur $\mathbb{R}$. De plus,

$$\begin{aligned} D_x f(x) &= D_x \exp(x + y) \exp(-x) + \exp(x + y) D_x \exp(-x) \\ &= \exp(x + y) \exp(-x) - \exp(x + y) \exp(-x) \\ &= 0. \end{aligned}$$

Il existe donc une constante r telle que

$$f(x) = \exp(x + y) \exp(-x) = r, \quad \forall x \in \mathbb{R}.$$

En prenant cette égalité pour $x = 0$, on obtient $r = \exp(y)$. En conclusion, on a obtenu

$$\exp(x + y) \exp(-x) = \exp(y) \quad \forall x, y \in \mathbb{R}. \quad (*)$$

En prenant alors cette égalité pour $y = 0$, on trouve

$$\exp(x) \exp(-x) = 1 \quad \forall x \in \mathbb{R};$$

il s’ensuit alors que (*) devient

$$\exp(x + y) = \exp(x) \exp(y) \quad \forall x, y \in \mathbb{R}.$$

5) Pour tout $x \in \mathbb{R}$, on a

$$\exp(x) = \exp\left(\frac{x}{2} + \frac{x}{2}\right) = \left(\exp\left(\frac{x}{2}\right)\right)^2$$

et

$$\exp(x) \exp(-x) = 1.$$

Dès lors $\exp(x)$ est positif et non nul.

6) Soit $p \in \mathbb{N}$. Pour $x > 0$, on a

$$\exp(x) = \sum_{m=0}^{+\infty} \frac{x^m}{m!} \geq \frac{x^{p+1}}{(p+1)!}$$

donc

$$\frac{\exp(x)}{x^p} \geq \frac{x}{(p+1)!};$$

il s'ensuit que

$$\lim_{x \rightarrow +\infty} \frac{\exp(x)}{x^p} = +\infty.$$

Si $x \neq 0$, on a

$$x^p \exp(x) = \frac{1}{\frac{\exp(-x)}{x^p}} = (-1)^p \frac{1}{\frac{\exp(-x)}{(-x)^p}};$$

comme

$$\lim_{x \rightarrow -\infty} \frac{\exp(-x)}{(-x)^p} = \lim_{y \rightarrow +\infty} \frac{\exp(y)}{y^p} = +\infty$$

on obtient

$$\lim_{x \rightarrow -\infty} x^p \exp(x) = 0.$$

□

Remarques 1) Le nombre réel $\exp(1)$ est noté e . On démontre que c'est un nombre irrationnel. Comme la fonction exponentielle vérifie $\exp(x+y) = \exp(x)\exp(y)$, ($x, y \in \mathbb{R}$) on a, pour tout naturel m

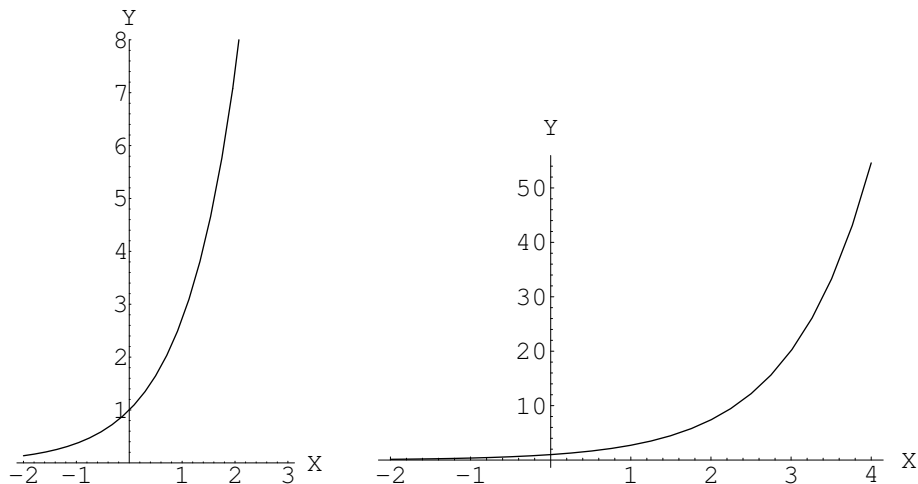
$$\exp(m) = \exp(1 + \dots + 1) = \underbrace{\exp(1) \exp(1) \dots \exp(1)}_{m \text{ facteurs}} = \underbrace{e e \dots e}_{m \text{ facteurs}} = e^m;$$

la fonction exponentielle est donc tout naturellement également notée

$$\exp(x) = e^x, \quad x \in \mathbb{R}.$$

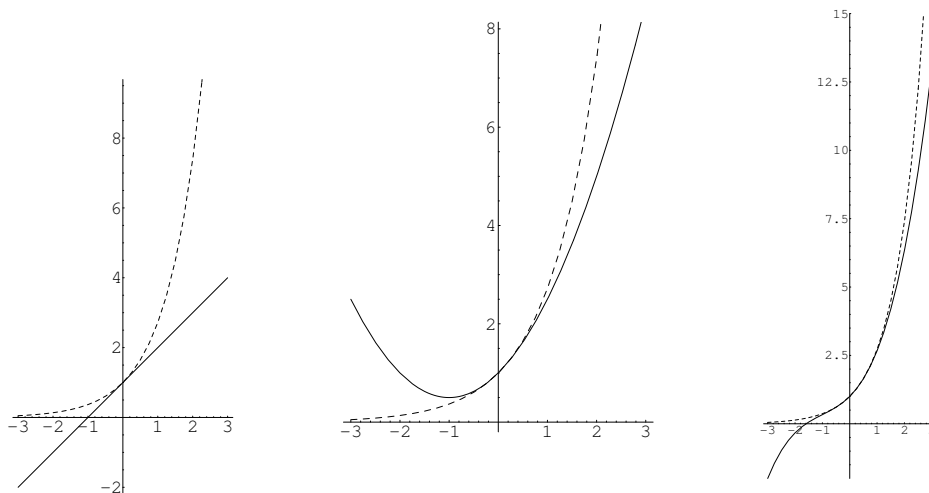
2) A partir des propriétés 1), 5) et 6), on obtient bien sûr que la fonction $\exp$ est une fonction bijective de $\mathbb{R}$ sur $]0, +\infty[$.

Voici la représentation graphique de la fonction exponentielle; la première représentation est effectuée dans un repère orthonormé; on voit bien son comportement quand l'argument augmente (comportement "exponentiel"!). La seconde représentation est effectuée dans un repère orthogonal non normé.



Voici la représentation graphique de la fonction exponentielle (en pointillés) et de ses approximations en 0 respectivement à l'ordre 1, 2, 3, c'est-à-dire des polynômes

$$P_1(x) = 1 + x, \quad P_2(x) = 1 + x + \frac{x^2}{2}, \quad P_3(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}.$$



3.3.3 Exponentielle complexe

DÉFINITIONS ET PROPRIÉTÉS

Terminons cette partie consacrée à la fonction exponentielle par une introduction à l'exponentielle complexe, à la définition des fonctions sin, cos (pas par un dessin !) et quelques compléments concernant les nombres complexes.

Pour tout $z \in \mathbb{C}$, la série

$$\sum_{m=0}^{+\infty} \frac{z^m}{m!}$$

converge absolument. La fonction ainsi définie dans $\mathbb{C}$ est notée

$$\exp(z) \quad \text{ou encore} \quad e^z.$$

La fonction exponentielle de domaine $\mathbb{R}$ en est sa restriction. De nombreuses propriétés sont encore vérifiées par cette fonction complexe MAIS plus question de parler de croissance, ni de positivité, ... car il s'agit d'une fonction définie dans $\mathbb{C}$ et à valeurs dans $\mathbb{C}$.

Propriété 3.3.3 1) On a

$$\exp(z + z') = \exp(z) \exp(z') \quad \forall z, z' \in \mathbb{C}.$$

2) On a

$$\overline{\exp(z)} = \exp(\bar{z}) \quad \forall z \in \mathbb{C}.$$

3) On a

$$D_t \exp(z_0 t) = z_0 \exp(z_0 t) \quad \forall t \in \mathbb{R} \text{ et } z_0 \in \mathbb{C} \text{ fixé.}$$

4) On a

$$\exp(z) = \exp(z') \Leftrightarrow \exists k \in \mathbb{Z} : z = z' + 2ik\pi.$$

En particulier, pour $x, x' \in \mathbb{R}$, on a

$$\exp(ix) = \exp(ix') \Leftrightarrow \exists k \in \mathbb{Z} : x = x' + 2k\pi.$$

Preuve. Résultat admis. $\square$

On utilise encore la notation

$$\exp(z) = e^z, \quad z \in \mathbb{C}.$$

Définissons à présent les fonctions sin et cos sur $\mathbb{R}$.

Définition 3.3.4 On définit

$$\cos x = \Re(e^{ix}), \quad \sin x = \Im(e^{ix}), \quad x \in \mathbb{R}$$

On en déduit que

$$\cos x + i \sin x = \exp(ix) = e^{ix}, \quad \cos x - i \sin x = \overline{e^{ix}} = e^{-ix}$$

pour tout réel x .

De plus, comme $\Re z = \frac{z+\bar{z}}{2}$ et $\Im z = \frac{z-\bar{z}}{2i}$ pour tout complexe z , on déduit aussi de la définition que

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}.$$

On a les propriétés suivantes.

Propriété 3.3.5 1) Pour tout $x \in \mathbb{R}$, on a

$$|e^{ix}| = 1.$$

2) $\cos^2 x + \sin^2 x = 1$ pour tout réel x .

3) On a $(\cos x + i \sin x)^m = \cos(mx) + i \sin(mx)$ pour tout naturel m et tout réel x .

4) Pour tout $z \in \mathbb{C}$ tel que $|z| = 1$, il existe $x \in [0, 2\pi[$ unique tel que

$$z = e^{ix}.$$

Preuve. 1) On a $z\bar{z} = |z|^2$ pour tout complexe z . Il s'ensuit que

$$|e^{ix}|^2 = e^{ix} \overline{e^{ix}} = e^{ix} e^{-ix} = e^{ix-ix} = 1$$

d'où la conclusion car le module d'un complexe est un réel positif ou nul.

2) On a

$$1 = |e^{ix}|^2 = (\Re(e^{ix}))^2 + (\Im(e^{ix}))^2 = \cos^2 x + \sin^2 x.$$

3) On a

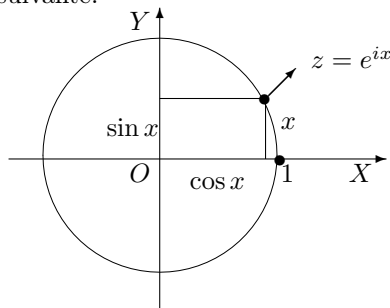
$$(\cos x + i \sin x)^m = (e^{ix})^m = e^{imx} = \cos(mx) + i \sin(mx).$$

4) Résultat admis. $\square$

COMPLEXES ET TRIGONOMETRIE

1) Etant donné un complexe z de module 1, c'est-à-dire un point du plan situé sur le cercle centré à l'origine et de rayon 1, on sait qu'il existe un réel unique $x \in [0, 2\pi[$ tel que $z = e^{ix}$. On montre aussi que la longueur de l'arc de cercle joignant le complexe 1 au complexe z vaut x .

On obtient donc la représentation suivante.



2) La forme trigonométrique d'un nombre complexe consiste simplement à écrire celui-ci en se servant des coordonnées polaires du point du plan qu'il détermine.

Etant donné $z \in \mathbb{C}$, $z \neq 0$, on sait qu'il existe un réel $x \in [0, 2\pi[$, unique, tel que

$$\frac{z}{|z|} = e^{ix}.$$

En posant

$$r = |z|$$

on a

$$z = re^{ix},$$

c'est ce que l'on appelle la forme trigonométrique du complexe z . Les réels r et x constituent également les coordonnées polaires du point P d'abscisse $\Re z$ et d'ordonnée $\Im z$.

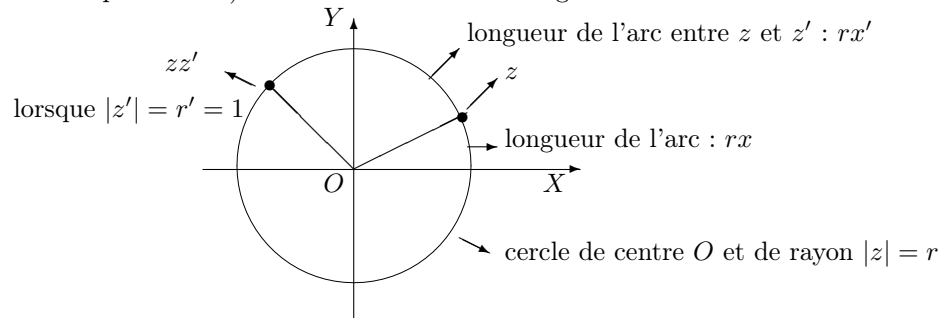
3) Interprétons à présent la *multiplication de deux complexes*. Soient z, z' deux complexes non nuls. On peut écrire

$$z = re^{ix}, \quad z' = r'e^{ix'}$$

donc

$$zz' = rr'e^{i(x+x')}.$$

La multiplication de z par z' consiste donc en une multiplication par le réel r' (qui s'interprète comme la multiplication d'un vecteur par un réel) et en une rotation d'un angle x' .



4) Grâce à la forme trigonométrique des complexes, on peut aussi démontrer que, pour tout complexe non nul z et tout naturel $n \geq 1$, il existe n complexes distincts $z_0, z_1, \dots, z_{n-1}$ tels que

$$z_k^n = z.$$

On dit que les complexes z_k ($k = 0, \dots, n-1$) sont les racines n -ièmes du complexe z .

La preuve est constructive : si

$$z = re^{ix}, \quad r > 0, x \in [0, 2\pi[$$

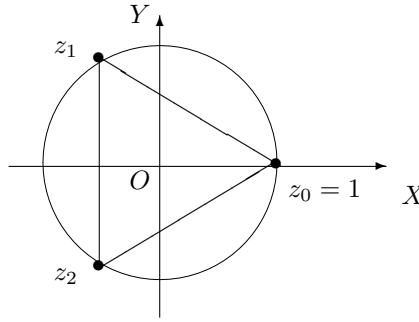
on obtient en effet

$$z_k = \sqrt[n]{r} e^{ix'_k}, \quad x'_k = \frac{x + 2k\pi}{n}, \quad k = 0, \dots, n-1.$$

Les n racines n -ièmes d'un complexe sont donc les sommets d'un polygone régulier à n côtés inscrit dans un cercle centré à l'origine et de rayon $\sqrt[n]{r}$; la mesure des angles entre les vecteurs joignant l'origine à deux racines consécutives est $\frac{2\pi}{n}$ radian(s).

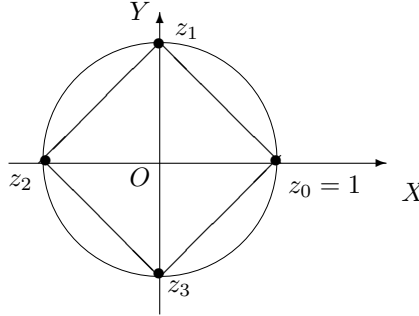
Voici trois exemples.

Les racines cubiques de $1 = e^{i0}$ sont $z_0 = 1$, $z_1 = e^{i\frac{2\pi}{3}}$, $z_2 = e^{i\frac{4\pi}{3}}$. Leur représentation est la suivante.



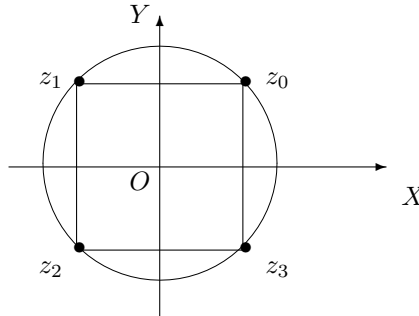
Représentation des racines cubiques de 1

Les racines quatrièmes de $1 = e^{i0}$ sont $z_0 = 1$, $z_1 = e^{i\frac{\pi}{2}} = i$, $z_2 = e^{i\pi} = -1$, $z_3 = e^{i\frac{3\pi}{2}} = -i$. Leur représentation est la suivante.



Représentation des racines quatrièmes de 1

Les racines quatrièmes de $-1 = e^{i\pi}$ sont $z_0 = e^{i\frac{\pi}{4}}$, $z_1 = e^{i\frac{3\pi}{4}}$, $z_2 = e^{i\frac{5\pi}{4}}$, $z_3 = e^{i\frac{7\pi}{4}}$. Leur représentation est la suivante.



Représentation des racines quatrièmes de -1

3.4 Quelques exemples fondamentaux de développements en série de puissances

Pour tout complexe z , on a

$$\begin{aligned} \exp z &= e^z = \sum_{m=0}^{+\infty} \frac{z^m}{m!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \dots \\ \cos z &= \sum_{m=0}^{+\infty} \frac{(-1)^m}{(2m)!} z^{2m} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \dots \\ \sin z &= \sum_{m=0}^{+\infty} \frac{(-1)^m}{(2m+1)!} z^{2m+1} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \dots \end{aligned}$$

et pour tout réel $x > -1$, on a

$$\ln(x+1) = \sum_{m=1}^{+\infty} \frac{(-1)^{m-1}}{m} x^m = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$$

3.5 Annexe

3.5.1 Approximations polynomiales

REMARQUES

1) Si $P(x) = a_N x^N + \dots + a_n x^n + \dots + a_1 x + a_0$ est un polynôme de degré strictement supérieur à n et vérifie

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - P(x - x_0)}{(x - x_0)^n} = 0$$

alors le polynôme $P_1(x) = a_n x^n + \dots + a_1 x + a_0$ vérifie

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - P_1(x - x_0)}{(x - x_0)^n} = 0.$$

En effet, on a

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{P(x - x_0) - P_1(x - x_0)}{(x - x_0)^n} = \sum_{j=n+1}^N \lim_{x \rightarrow x_0, x \neq x_0} a_j (x - x_0)^{j-n} = 0.$$

2) Si f est continu en x_0 alors

$$\lim_{x \rightarrow x_0} (f(x) - f(x_0)) = \lim_{x \rightarrow x_0, x \neq x_0} (f(x) - f(x_0)) = 0.$$

Il s'ensuit que f admet la constante $f(x_0)$ comme approximation à l'ordre 0 en x_0 .

3) Pour une fonction f et un point x_0 du domaine de définition de f , la propriété $\lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - P(x - x_0)}{(x - x_0)^n} = 0$ n'implique pas nécessairement la continuité de f au point x_0 . Il suffit en effet de considérer la fonction nulle pour tout $x \neq 0$ et qui vaut 1 en $x = 0$. Cette fonction admet le polynôme 0 comme approximation à l'ordre n (pour tout $n \in \mathbb{N}$) en $x_0 = 0$ mais n'est pas continue en 0.

4) Si f est continu en x_0 alors, pour tout polynôme P , on a

$$\lim_{x \rightarrow x_0, x \neq x_0} (f(x) - P(x - x_0)) = \lim_{x \rightarrow x_0} (f(x) - P(x - x_0)) (= f(x_0) - P(0)).$$

La restriction faite dans la définition ($x \neq x_0$ pour $n = 0$) peut donc être omise. Dans ce cas, on appelle donc approximation à l'ordre n de f en x_0 un polynôme P de degré inférieur ou égal à n tel que

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - P(x - x_0)}{(x - x_0)^n} = \lim_{x \rightarrow x_0} \frac{f(x) - P(x - x_0)}{(x - x_0)^n} = 0.$$

PROPRIÉTÉS

Propriété 3.5.1 Soit f une fonction définie sur $]a, b[$ et soit $x_0 \in]a, b[$.

1) Si f a une approximation à l'ordre n en x_0 , alors f a une approximation à tout ordre inférieur à n en x_0 .

2) Si f admet une approximation à l'ordre n en x_0 , cette approximation est unique.

3) Si f est continu en x_0 et admet une approximation à l'ordre 1 alors f est dérivable en x_0 et l'approximation à l'ordre 1 est

$$f(x_0) + (x - x_0)Df(x_0).$$

4) Si f est continu en x_0 et admet une approximation à l'ordre 2 en x_0 , f n'est pas nécessairement deux fois dérivable en x_0 .

Preuve. 1) En effet, si P est un polynôme de degré inférieur ou égal à n tel que

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - P(x - x_0)}{(x - x_0)^n} = 0$$

alors, pour tout $k \in \{0, \dots, n-1\}$, on a

$$\begin{aligned} \lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - P(x - x_0)}{(x - x_0)^k} &= \lim_{x \rightarrow x_0, x \neq x_0} \left(\frac{f(x) - P(x - x_0)}{(x - x_0)^n} (x - x_0)^{n-k} \right) \\ &= \lim_{x \rightarrow x_0, x \neq x_0} \left(\frac{f(x) - P(x - x_0)}{(x - x_0)^n} \right) \lim_{x \rightarrow x_0, x \neq x_0} (x - x_0)^{n-k} \\ &= 0 \end{aligned}$$

On conclut alors en utilisant la première remarque ci-dessus.

2) Soit $P(x) = a_0 + a_1x + \dots + a_nx^n$ un polynôme de degré inférieur ou égal à n tel que

$$\lim_{x \rightarrow x_0} \frac{f(x) - P(x - x_0)}{(x - x_0)^n} = 0.$$

Recherche du coefficient a_0 . Vu 1) on a

$$\lim_{x \rightarrow x_0, x \neq x_0} (f(x) - P(x - x_0)) = 0.$$

Il s'ensuit que

$$\lim_{x \rightarrow x_0, x \neq x_0} f(x) = \lim_{x \rightarrow x_0, x \neq x_0} (f(x) - P(x - x_0)) + \lim_{x \rightarrow x_0, x \neq x_0} P(x - x_0) = a_0.$$

Le coefficient a_0 est donc unique.

Recherche du coefficient a_1 . Vu 1) on a

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - P(x - x_0)}{x - x_0} = 0.$$

Il s'ensuit que

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - a_0}{x - x_0} = \lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - P(x - x_0)}{x - x_0} + \lim_{x \rightarrow x_0, x \neq x_0} \frac{P(x - x_0) - a_0}{x - x_0} = a_1.$$

Le coefficient a_1 est donc unique.

On continue de cette manière jusqu'au coefficient a_n .

3) Soit $P(x) = ax + b$ tel que $P(x - x_0)$ soit l'approximation de f à l'ordre 1 en x_0 . Vu ce qui précède, la constante b et $f(x_0)$ sont des approximations à l'ordre 0 en x_0 . On a donc nécessairement

$$b = f(x_0).$$

Dès lors, pour $x \in]a, b[$, $x \neq x_0$, on a

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{f(x) - a(x - x_0) - f(x_0)}{x - x_0} + \frac{a(x - x_0)}{x - x_0} = \frac{f(x) - a(x - x_0) - f(x_0)}{x - x_0} + a$$

donc la limite

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{f(x) - P(x - x_0)}{x - x_0} + a = a$$

existe et est finie.

4) La fonction

$$f(x) = \begin{cases} x^3 \sin(1/x) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

est définie sur $\mathbb{R}$ et est continue sur $\mathbb{R}$. Cette fonction est dérivable sur $\mathbb{R}$ mais n'est pas deux fois dérivable sur $\mathbb{R}$. On a en effet

$$Df(x) = \begin{cases} 3x^2 \sin(1/x) - x \cos(1/x) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

et cette fonction n'est pas dérivable en 0 car la limite

$$\lim_{x \rightarrow 0} \frac{3x^2 \sin(1/x) - x \cos(1/x)}{x}$$

n'existe pas.

Cependant, la fonction f admet une approximation à l'ordre 2 en $x_0 = 0$ car

$$\lim_{x \rightarrow 0} \frac{f(x) - 0}{x^2} = \lim_{x \rightarrow 0} x \sin(1/x) = 0.$$

3.5.2 Critères de convergence pour les séries

Proposition 3.5.2 (Critère de Cauchy) *La série de terme général x_m converge si et seulement si $\forall \varepsilon > 0$, il existe $M \in \mathbb{N}_0$ tel que*

$$\left| \sum_{m=p}^q x_m \right| \leq \varepsilon, \quad \forall q \geq p \geq M.$$

En particulier, si la série de terme général $|x_m|$ converge et si $R > 0$ est tel que $|y_m| \leq R|x_m| \forall m$, alors la série de terme général y_m est aussi une série convergente.

Preuve. Résultat qui se déduit du critère de Cauchy pour les suites car une série est une suite particulière. $\square$

Proposition 3.5.3 *1) Si la série de terme général $|x_m|$ converge et si $R > 0$ est tel que $|y_m| \leq R|x_m| \forall m$, alors la série de terme général y_m est aussi une série convergente.*

2) Critère des séries alternées. Si r_m ($m \in \mathbb{N}_0$) est une suite de réels qui décroît vers 0, alors la série de terme général $(-1)^m r_m$ est une série convergente. De plus, pour tout $p \in \mathbb{N}_0$, on a

$$\left| \sum_{m=p}^{+\infty} (-1)^m r_m \right| \leq r_p.$$

Preuve. 1) Il s'agit d'une application du critère de Cauchy car on a

$$\left| \sum_{m=p}^q y_m \right| \leq \sum_{m=p}^q |y_m| \leq R \sum_{m=p}^q |x_m|.$$

2) Ce résultat peut être démontré en utilisant le critère de Cauchy. En effet, pour tous $p, q \in \mathbb{N}_0$, $p \leq q$, on a

$$\begin{aligned} & \sum_{m=p}^q (-1)^m r_m \\ &= (-1)^p (r_p - r_{p+1} + r_{p+2} - r_{p+3} + \dots + (-1)^{q-p} r_q) \\ &= (-1)^p \begin{cases} (r_p - r_{p+1}) + (r_{p+2} - r_{p+3}) + (r_{p+4} - r_{p+5}) + \dots + (r_{q-1} - r_q) & \text{si } q-p \text{ est impair} \\ (r_p - r_{p+1}) + (r_{p+2} - r_{p+3}) + (r_{p+4} - r_{p+5}) + \dots + r_q & \text{si } q-p \text{ est pair} \end{cases} \end{aligned}$$

donc

$$\begin{aligned} & \left| \sum_{m=p}^q (-1)^m r_m \right| = \\ & \begin{cases} (r_p - r_{p+1}) + (r_{p+2} - r_{p+3}) + (r_{p+4} - r_{p+5}) + \dots + (r_{q-1} - r_q) & \text{si } q-p \text{ est impair} \\ (r_p - r_{p+1}) + (r_{p+2} - r_{p+3}) + (r_{p+4} - r_{p+5}) + \dots + r_q & \text{si } q-p \text{ est pair} \end{cases} \end{aligned}$$

car la suite r_m ($m \in \mathbb{N}_0$) est décroissante. En regroupant encore les termes d'une autre manière, on a

$$\left| \sum_{m=p}^q (-1)^m r_m \right| =$$

$$\begin{cases} r_p + (r_{p+2} - r_{p+1}) + (r_{p+4} - r_{p+3}) + \dots + (r_q - r_{q-1}) & \text{si } q - p \text{ est pair} \\ r_p + (r_{p+2} - r_{p+1}) + (r_{p+4} - r_{p+3}) + \dots - r_q & \text{si } q - p \text{ est impair} \end{cases}$$

donc cette expression est égale à r_p auquel on ajoute des termes tous négatifs. Il s'ensuit que

$$\left| \sum_{m=p}^q (-1)^m r_m \right| \leq r_p.$$

On peut alors conclure car la suite r_m ($m \in \mathbb{N}_0$) converge vers 0. $\square$

3.5.3 Critères pratiques de convergence des séries

Définition 3.5.4 La série de terme général x_m , c'est-à-dire la série $\sum_{m=1}^{+\infty} x_m$ est dite

- absolument convergente si la série

$$\sum_{m=1}^{+\infty} |x_m| \quad \text{converge.}$$

- semi-convergente si

$$\sum_{m=1}^{+\infty} x_m \text{ converge mais } \sum_{m=1}^{+\infty} |x_m| \text{ ne converge pas.}$$

Vu ce qui précède, une série absolument convergente est convergente mais la réciproque est fausse.

On démontre les résultats suivants, appelés *critères pratiques de convergence*, à l'aide des séries de référence (série géométrique et série de Riemann).

Proposition 3.5.5 (Critères pratiques de convergence) Soit $\sum_{m=1}^{+\infty} x_m$ une série.

1) **Critère de la racine.**

- Si $\lim_{m \rightarrow +\infty} \sqrt[m]{|x_m|} = \theta \in [0, 1[$ alors la série $\sum_{m=1}^{+\infty} x_m$ est absolument convergente (donc est convergente).

- Si on a $\lim_{m \rightarrow +\infty} \sqrt[m]{|x_m|} = \theta > 1$ ou $\lim_{m \rightarrow +\infty} \sqrt[m]{|x_m|} = +\infty$ ou $\lim_{m \rightarrow +\infty} \sqrt[m]{|x_m|} = 1^+$ alors la série $\sum_{m=1}^{+\infty} x_m$ diverge (car son terme général ne tend pas vers 0).

2) **Critère du quotient.** Supposons $x_m \neq 0$ pour tout m .

- Si $\lim_{m \rightarrow +\infty} \frac{|x_{m+1}|}{|x_m|} = \theta \in [0, 1[$ alors la série $\sum_{m=1}^{+\infty} x_m$ est absolument convergente (donc est convergente).

- Si on a $\lim_{m \rightarrow +\infty} \frac{|x_{m+1}|}{|x_m|} = \theta > 1$ ou $\lim_{m \rightarrow +\infty} \frac{|x_{m+1}|}{|x_m|} = +\infty$ ou $\lim_{m \rightarrow +\infty} \frac{|x_{m+1}|}{|x_m|} = 1^+$ alors la série diverge (car son terme général ne tend pas vers 0).

3) **Critère de Riemann.**

- S'il existe $\alpha > 1$ tel que $\lim_{m \rightarrow +\infty} m^\alpha |x_m|$ existe et est finie alors la série $\sum_{m=1}^{+\infty} x_m$ est absolument convergente (donc convergente).

- Si $\lim_{m \rightarrow +\infty} m |x_m|$ existe et vaut soit $+\infty$, soit un réel strictement positif, alors la série $\sum_{m=1}^{+\infty} |x_m|$ diverge (mais la série $\sum_{m=1}^{+\infty} x_m$ peut converger).

3.5.4 Une approximation de e

Propriété 3.5.6 Pour tous $x \in \mathbb{R}$ et $M \in \mathbb{N}_0$ vérifiant $M \geq 2(|x| - 1)$, les sommes partielles de la série définissant $\exp(x)$ sont telles que

$$\left| \exp(x) - \sum_{m=0}^M \frac{x^m}{m!} \right| \leq 2 \frac{|x|^{M+1}}{(M+1)!}.$$

En particulier pour $x = 1$ on trouve

$$\sum_{m=0}^M \frac{1}{m!} \leq \exp(1) \leq \sum_{m=0}^M \frac{1}{m!} + \frac{2}{(M+1)!}$$

d'où l'estimation $\exp(1) = 2.718$ ($M = 6$) ; plus précisément $\exp(1) = 2.71828182\dots$

Preuve. Pour tout $M \in \mathbb{N}$ et tout réel x , on a

$$\begin{aligned} \exp(x) - \sum_{m=0}^M \frac{x^m}{m!} &= \sum_{m=M+1}^{+\infty} \frac{x^m}{m!} \\ &= \frac{x^{M+1}}{(M+1)!} \left(1 + \sum_{m=1}^{+\infty} \frac{x^m}{(m+M+1) \dots (M+2)} \right). \end{aligned}$$

Dès lors, si $M+2 \geq 2|x|$,

$$\left| \sum_{m=1}^{+\infty} \frac{x^m}{(m+M+1) \dots (M+2)} \right| = \left| \sum_{m=1}^{+\infty} \frac{x}{m+M+1} \cdots \frac{x}{M+2} \right| \leq \sum_{m=1}^{+\infty} \frac{1}{2^m} = 1.$$

□