MATHEMATIQUES GENERALES II, MATH 1009

Premier bloc en chimie

Deuxième bloc en géologie

Année académique 2025-2026

Françoise BASTIN

Introduction au cours
« Mathématiques générales II (MATH1009) »

Aperçu général du cours

L'enseignement auquel ces notes se réfèrent fait suite au cours intitulé *Mathématiques générales I*, figurant au programme du premier quadrimestre du premier bloc des bacheliers en chimie, géologie, informatique.

Les matières suivantes sont traitées.

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

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).

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/ ;** 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 tout particulier à Madame Christine Amory et à Madame Jacqueline Crasborn, pour leur présence depuis tellement d'années !, pour leurs relectures, avis et suggestions pertinents, travail de fond et de forme, les graphiques, les listes d'exercices et pour leur enthousiasme et dynamisme permanents, facteurs si importants au sein d'une équipe.

Françoise Bastin, 9 décembre 2025(v1 26/12/15)

Table des matières

Chapitre 1

Eléments de calcul matriciel

1.1 Introduction

Les matrices et le calcul matriciel offrent une interprétation de problèmes en termes calculatoires. Par exemple, il est courant de rencontrer des systèmes d'équations linéaires (c'est-à-dire où les inconnues n'apparaissent qu'au premier degré). Imaginons par exemple que l'on doive résoudre le système (S) suivant

$$\left\{\begin{array}{lcl} 2x_1 - 3x_2 + x_4 & = & 9 \\ -x_1 + x_2 + x_3 - x_4 & = & 2 \\ 3x_1 + 3x_2 - 2x_3 + x_4 & = & -1. \end{array}\right.$$

Des résultats et techniques basés notamment sur le rang des matrices donnent des méthodes pour résoudre ce type de système par des manipulations des tableaux des données

$$A = \begin{pmatrix} 2 & -3 & 0 & 1 \\ -1 & 1 & 1 & -1 \\ 3 & 3 & -2 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 9 \\ 2 \\ -1 \end{pmatrix}$$

et du tableau des inconnues

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

En écriture matricielle, le système (S) s'écrit

$$AX = B.$$

1.2 Matrices : définitions générales et notations

Définition 1.2.1. *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 est, par définition, le nombre d'éléments des lignes ; c'est donc aussi le nombre de colonnes de la matrice. La longueur des colonnes de la matrice est, par définition, le nombre d'éléments des colonnes ; c'est donc aussi le 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.

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

$$(A)_{k,j}\,.$$

Par exemple si

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

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 :

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

Lorsque la matrice ne possède qu'une ligne ou qu'une colonne, on l'appelle simplement *vecteur* (ligne ou colonne bien sûr). Dans ce cas, les notations des éléments sont simplifiées. Par exemple, si X est une matrice d'une seule colonne et de n lignes, on écrit tout simplement

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

et pour $j \in \{1,\dots,n\}$, $(X)_{j,1}$ est noté simplement X_j.

Passons au cas important suivant, celui des matrices dites « carrées ».

Définition 1.2.2. *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 identité[1] *de dimension n. Elle est notée I ou $\mathbb{1}$.*

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

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

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_j, $j = 1,\dots,5$, sont des nombres complexes.

1.3 Matrices associées

Définition 1.3.1. *Etant donné une matrice A de type $p \times q$, on lui associe trois autres matrices :*

- *la matrice conjuguée de A, notée $\overline{A}$; il s'agit de la matrice de même type que celui de A dont les éléments sont les conjugués de ceux de A :*

$$\left(\overline{A}\right)_{k,j} = \overline{(A)_{k,j}}, \quad k = 1,\dots,p, \ j = 1,\dots,q.$$

- *la matrice transposée de A, notée $\widetilde{A}$; il s'agit de la matrice de type $q \times p$ telle que*

$$\left(\widetilde{A}\right)_{k,j} = (A)_{j,k}, \quad k = 1,\dots,q, \ j = 1,\dots,p.$$

  Ses lignes sont donc formées des éléments des colonnes de A et ses colonnes sont formées des éléments des lignes de A.

- *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*

$$A^* = \overline{\widetilde{A}} = \widetilde{\overline{A}} \qquad \text{ou encore} \qquad (A^*)_{k,j} = \overline{(A)_{j,k}}, \ k = 1,\dots,q, \ j = 1,\dots,p.$$

1. Parfois on utilise aussi le terme « matrice unité »

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

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

on a

$$\overline{A} = \begin{pmatrix} 0 & -2 & -i \\ i & \sqrt{5} & 6 \\ 3/4 & -i/2+1 & 0 \\ -1 & 3 & 4 \end{pmatrix}, \quad \widetilde{A} = \begin{pmatrix} 0 & -i & 3/4 & -1 \\ -2 & \sqrt{5} & i/2+1 & 3 \\ i & 6 & 0 & 4 \end{pmatrix}$$

et

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

Pour la matrice identité, on a

$$\overline{\mathbb{1}} = \widetilde{\mathbb{1}} = \mathbb{1}^* = \mathbb{1}.$$

En effet, les éléments de cette matrice sont tous réels (donc $\overline{\mathbb{1}} = \mathbb{1}$) et sa ligne numéro k est formée des mêmes éléments, dans le même ordre, que sa colonne numéro k (donc $\widetilde{\mathbb{1}} = \mathbb{1}$).

1.4 Opérations entre matrices

1.4.1 Addition de deux matrices du même type

Définition 1.4.1. *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 :*

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

Par exemple, pour

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

on a

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

1.4.2 Multiplication d'une matrice par un nombre complexe

Définition 1.4.2. *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 :*

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

Par exemple, pour

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

on a

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

1.4.3 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(s) 1.4.3. *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

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

1.4.4 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 [2] dans une base et plus précisément de la représentation matricielle de la composée de deux opérateurs linéaires.

Cette définition permet aussi une modélisation de situations concrètes d'évolution de processus. Après traitement par des méthodes ad hoc, on obtient des réponses aux questions posées initialement (voir la section suivante pour des premiers exemples, puis plus loin pour la présentation des méthodes adéquates).

Définition 1.4.4. *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*

2. Pour rappel, un opérateur linéaire entre espaces vectoriels est une application qui est telle que l'image d'une combinaison linéaire est la combinaison linéaire des images.

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

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

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

Présentons quelques exemples de produits de deux matrices.

Soient

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

On a

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

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

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

1.4.5 Propriétés du produit matriciel

Deux propriétés importantes du produit matriciel sont les suivantes.

Propriété(s) 1.4.5. *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) Les deux membres de l'égalité sont bien définis et sont des matrices de même format. De fait, vu les hypothèses sur les formats, la matrice AB est bien définie et est de format $n \times q$; la matrice C étant de format $q \times r$, le produit $(AB)C$ est bien défini et est une matrice de format $n \times r$. Regardons alors le membre de droite, à savoir $A(BC)$. Ici encore, vu les hypothèses sur les formats, la matrice BC est bien définie et est de format $p \times r$; la matrice A étant de format $n \times p$, le produit $A(BC)$ est bien défini et est une matrice de format $n \times r$.

Cela étant, montrons que les éléments des matrices $(AB)C$ et $A(BC)$ sont les mêmes. L'élément sur la ligne l ($l = 1, \ldots, n$) et la colonne k ($k = 1, \ldots, r$) de la matrice $(AB)C$ est

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

Comme l'addition de nombres est une opération commutative, et comme le produit entre nombres est également une opération commutative, l'expression précédente peut être écrite en permutant les deux sommes et les produits. On a alors

$$\begin{aligned}
\Big((AB)C\Big)_{l,k} &= \sum_{s=1}^{q}\left(\sum_{j=1}^{p}(A)_{l,j}(B)_{j,s}\right)(C)_{s,k}\\
&= \sum_{j=1}^{p}\left(\sum_{s=1}^{q}(B)_{j,s}(C)_{s,k}\right)(A)_{l,j}\\
&= \sum_{j=1}^{p}(BC)_{j,k}(A)_{l,j}\\
&= \sum_{j=1}^{p}(A)_{l,j}(BC)_{j,k} \;=\; \Big(A(BC)\Big)_{l,k}
\end{aligned}$$

et on conclut.

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

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

car l'élément de la première ligne et première colonne du produit du membre de gauche est 0 et celui du produit du membre de droite est 1. Il est bon de remarquer que tout autre exemple est aussi une justification (donc pas besoin d'étudier par coeur celui-ci !!)

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 naturelles d'une matrice carrée

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

Voici quelques autres propriétés du produit matriciel.

1) *Si on travaille entre matrices carrées de même dimension, on a*

$$A0 = 0A = 0, \quad A\mathbb{1} = \mathbb{1}A = A.$$

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

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

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

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

5) 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

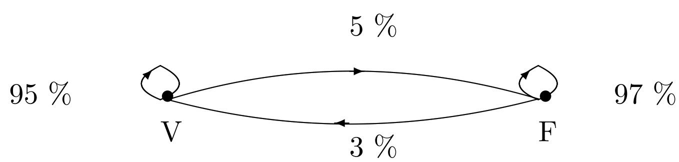
$$\begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}^2 = 0\,, \quad \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = 0.$$

Preuve. Tout s'obtient directement par calcul en utilisant les définitions.□

1.5 Les matrices stochastiques

1.5.1 Migration de la population

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.

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

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

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

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

la répartition de la population après deux années est

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

La matrice T est appelée *matrice de transition.* Plus généralement, après k années, on a

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

On verra plus loin une méthode d'analyse de l'évolution des puissances de la matrice de transition T.

1.6 Déterminants des matrices carrées

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 $\det A$ ou $\det(A)$ pour éviter toute confusion si la matrice dont on prend le déterminant s'écrit de façon plus complexe.

1.6.1 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

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

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

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

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

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

L'expression est la même pour le déterminant d'une matrice de dimension deux, bien sûr seulement avec une somme de deux termes.

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

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

1.6.2 Propriétés

Une première propriété fort utile dans le calcul des déterminants est la suivante. Elle est admise ici.

Propriété(s) 1.6.1 (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

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

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 :

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

Voici deux autres propriétés. Nous les admettrons également.

Propriété(s) 1.6.2. 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 de ce qui précède, les propriétés suivantes se démontrent tout de suite, comme expliqué au cours.

Propriété(s) 1.6.3. 1. *Le déterminant d'une matrice qui a deux rangées parallèles égales est nul.*

2. *Si, à une rangée d'une matrice, on ajoute une combinaison linéaire d'autres rangées parallèles, on définit une nouvelle matrice qui a le même déterminant que celui de la matrice de départ.*

Donnons quelques exemples illustratifs. 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 :

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

Il faut donc bien remarquer que quand on multiplie le déterminant d'une matrice A par un nombre, on obtient le déterminant d'une matrice obtenue à partir de A en multipliant uniquement une rangée de A par le nombre (*).

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 2

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

Ce résultat est à comparer et à ne pas confondre avec ce qui précède (*). Il est incorrect de dire que $r\det(A) = \det(rA)$! !, sauf bien sûr pour $r = 0$ ou $r = 1$.

Ces propriétés permettent aussi de simplifier grandement le calcul de déterminants ; elles sont spécialement utiles 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, sera d'une grande utilité dans l'étude de l'inversion des matrices carrées.

Propriété(s) 1.6.4 (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.*

Résultat admis.□

Illustrons cette propriété sur deux exemples. Soit la matrice

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

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

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

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

$$\begin{aligned}
& a_{31} \times \text{cofacteur de } a_{21} + a_{32} \times \text{cofacteur de } a_{22} + a_{33} \times \text{cofacteur de } a_{23} \\
&= 3 \times \left(\frac{-3}{2}\right) + 1 \times \left(6 + \frac{i}{2}\right) + \frac{i}{2} \times (-1+3i) \\
&= -\frac{9}{2} + 6 + \frac{i}{2} - \frac{i}{2} - \frac{3}{2} = 0.
\end{aligned}$$

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

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

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

$$\begin{aligned}
& a_{12} \times \text{cofacteur de } a_{11} + a_{22} \times \text{cofacteur de } a_{21} + a_{32} \times \text{cofacteur de } a_{31} \\
&= i \times \left(\frac{i}{4} - 1\right) + \frac{1}{2} \times \left(-\frac{3}{2}\right) + 1 \times (i+1) \\
&= -\frac{1}{4} - i - \frac{3}{4} + i + 1 = 0.
\end{aligned}$$

Enonçons maintenant en une seule fois les deux lois des mineurs. Soit une matrice carrée A de dimension n. La matrice des cofacteurs de A, notée $\mathcal{A}$, est la matrice carrée de même dimension que A dont l'élément $(\mathcal{A})_{k,l}$ est le cofacteur de l'élément $(A)_{k,l}$ de A, quels que soient $k, l = 1, \dots n$. Les deux lois des mineurs s'écrivent alors (c'est évident par définition du produit matriciel)

$$\boxed{\widetilde{\mathcal{A}}\,A \;=\; \det(A)\,\mathbb{1} \;=\; A\,\widetilde{\mathcal{A}}.}$$

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

Propriété(s) 1.6.5. *Si*

$$A = diag(a_1, a_2, \ldots, a_n)$$

alors

$$\det(A) = a_1 a_2 \ldots a_n.$$

En particulier,

$$\det(\mathbb{1}) = 1.$$

Ce résultat s'écrit, en dimension 2 :

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

en dimension 3 :

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

en dimension 4

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

La propriété précédente peut être généralisée au cas de matrices triangulaires inférieures (resp. supérieures), c'est-à-dire au cas de 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(s) 1.6.6. *Si A, B sont deux matrices carrées de même dimension, on a*

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

Résultat admis. □

Attention, il importe de vérifier les hypothèses. De fait, 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.

Terminons par énoncer et démontrer partiellement un résultat extrêmement utile concernant l'annulation d'un déterminant.

Propriété(s) 1.6.7. *Soit A une matrice carrée de dimension n. Alors le déterminant de A est nul si et seulement si l'une des colonnes (resp. des lignes) est une combinaison linéaire des autres.*

Preuve. D'une part si l'une des colonnes (resp. des lignes) est une combinaison linéaire des autres, alors il est clair que le déterminant est nul, par application de la propriété de linéarité et de celle qui affirme qu'un déterminant ayant deux rangées parallèles égales est nul.

La réciproque est admise, à savoir que si le déterminant d'une matrice carrée est nul, alors une colonne (resp. une ligne) est nécessairement une combinaison linéaire des autres.□

1.7 Inversion des matrices carrées

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 $A\mathbb{1} = \mathbb{1}A = A$ 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 carrée A, existe-t-il une matrice carrée A' de même dimension que A, telle que

$$AA' = \mathbb{1} \text{ et } A'A = \mathbb{1}?$$

Remarquons ici que le produit matriciel n'étant pas commutatif, le fait d'avoir $AA' = \mathbb{1}$ n'implique pas, à priori, que $A'A = \mathbb{1}$ comme c'est le cas pour les nombres complexes.

On adopte alors de façon naturelle la définition suivante.

Définition 1.7.1. *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*

$$AA' = \mathbb{1} = A'A.$$

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 carrée 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(s) 1.7.2. *Si A admet un inverse alors son déterminant est non nul.*

Preuve. Soit une matrice A' telle que $AA' = \mathbb{1}$. En prenant le déterminant, on obtient

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

donc $\det(A) \neq 0$. □

En fait, les lois des mineurs permettent de montrer que cette condition nécessaire à l'existence d'un inverse (à savoir $\det(A) \neq 0$) est aussi suffisante. Reprenons en effet la matrice des cofacteurs des éléments d'une matrice A, notée $\mathcal{A}$ (cf section consacrée aux déterminants). Comme expliqué précédemment, les deux lois des mineurs peuvent s'écrire

$$A\,\widetilde{\mathcal{A}} \;=\; (\det A)\;\mathbb{1} \quad \text{et} \quad \widetilde{\mathcal{A}}\,A \;=\; (\det A)\;\mathbb{1}.$$

Il s'ensuit le résultat ci-dessous.

Propriété(s) 1.7.3. *Si A est une matrice carrée dont le déterminant n'est pas nul, alors la matrice*

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

vérifie

$$AA' = \mathbb{1} \quad \textit{et} \quad A'A = \mathbb{1}.$$

Et qu'en est-il de l'unicité ? Montrons que si la matrice A admet une matrice inverse (et même moins, cf résultat ci-dessous), alors cette matrice inverse est unique et est

$$\frac{1}{\det A}\widetilde{\mathcal{A}}.$$

Propriété(s) 1.7.4. *Soit A une matrice carrée.*

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

$$BA = \mathbb{1} \qquad (\textit{resp.}\;\; AB = \mathbb{1}),$$

alors

$$\det A \neq 0$$

et

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

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

$$1 = \det \mathbb{1} = \det(BA) = \det(B)\;\det(A) \qquad (\text{resp.}\;\; 1 = \det \mathbb{1} = \det(AB) = \det(A)\;\det(B)),$$

donc le déterminant de A n'est pas nul.

Notons

$$C = \frac{1}{\det A}\widetilde{\mathcal{A}}.$$

On a

$$B = B\mathbb{1} = B(AC) = (BA)C = \mathbb{1}C = C$$
$$(\text{resp.} \quad B = \mathbb{1}B = (CA)B = C(AB) = C\mathbb{1} = C).$$

□

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

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

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(s) 1.7.5. *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

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

En particulier, $\mathbb{1}^{-1} = \mathbb{1}$.

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

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

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

Preuve. Tout se démontre directement en repassant à la définition. Prenons quelques cas.

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

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

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

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

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

1.8 Systèmes d'équations linéaires

1.8.1 Cas des systèmes carrés

Lorsque la matrice A du système d'équations

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

est carrée et inversible, le système admet une solution unique donnée par

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

Si la dimension de la matrice est n, on trouve rapidement la valeur des inconnues : quel que soit $j = 1, \ldots, n$, on a

$$x_j \; = \; \frac{\det(A^{(j)})}{\det(A)}$$

où $A^{(j)}$ est la matrice dont la colonne numéro j est le vecteur des termes indépendants B et dont les autres colonnes sont celles de A. Il est aisé de se convaincre de cela car il suffit de se rappeler la forme de l'inverse, à savoir la matrice $\mathcal{A}$ des cofacteurs de A transposée divisée par le déterminant de A. On a en effet

$$\begin{aligned} x_j &= \left(A^{-1}B\right)_j \\ &= \frac{1}{\det(A)} \left(\widetilde{\mathcal{A}}B\right)_j \\ &= \frac{1}{\det(A)} \sum_{k=1}^{n} \left(\widetilde{\mathcal{A}}\right)_{j,k} b_k \\ &= \frac{1}{\det(A)} \sum_{k=1}^{n} (\mathcal{A})_{k,j}\, b_k. \end{aligned}$$

En se rappelant que la matrice $A^{(j)}$ est la matrice dont la colonne numéro j est le vecteur des termes indépendants B et dont les autres colonnes sont celles de A, son déterminant, calculé à partir de la colonne j (application de la première loi des mineurs), est

$$\sum_{k=1}^{n} (\mathcal{A})_{k,j}\, b_k.$$

On trouve donc finalement

$$x_j \; = \; \frac{\det(A^{(j)})}{\det(A)}.$$

Dans le cas où la matrice n'est pas inversible, le système est soit incompatible, soit équivalent à un système constitué de moins d'équations. La théorie générale sort du cadre de ce cours. Seuls des exemples seront utilisés, principalement dans le cadre de la recherche de vecteurs propres (cf une section de la suite de ce chapitre).

1.8.2 Cas des systèmes non carrés

Des méthodes standards de résolution doivent normalement déjà être connues et interviennent dans divers exercices. La théorie générale n'est pas difficile mais, faute de temps, il n'est pas possible d'en donner les argumentations théoriques dans le cadre de ce cours.

1.9 Vecteurs propres, valeurs propres, diagonalisation

Ces notions (vecteurs et valeurs propres, diagonalisation) ne s'étudient que dans le cadre de matrices carrées.

1.9.1 Manipulations de matrices-vecteurs

Sont présentées ici quelques manipulations de calcul matriciel qui seront bien utiles dans la suite pour une meilleure et plus rapide compréhension des développements.

Soit une matrice carrée A de dimension n dont les colonnes sont notées $C_1, \ldots, C_n$ et soit un vecteur colonne X de n éléments. La matrice AX est de format $n \times 1$, c'est donc un vecteur colonne avec n éléments. Etant donné la définition du produit matriciel (« ligne par colonne »), le $j$ème élément de AX est donc obtenu en faisant la somme des produits des éléments de la ligne j de A par les éléments de X. Comme les éléments de la ligne j de A sont les éléments j de ses colonnes, pour obtenir AX, on fait donc la combinaison linéaire des colonnes de A avec les éléments de X comme coefficients. Autrement dit, on a

$$AX = \sum_{k=1}^{n} x_k \, C_k.$$

Autre formulation : si

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

et

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

l'élément j du vecteur AX est

$$\sum_{k=1}^{n} a_{jk} \, x_k;$$

le vecteur AX est donc

$$AX = \sum_{k=1}^{n} \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix} x_k = \sum_{k=1}^{n} x_k \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix},$$

c'est-à-dire

$$AX = \sum_{k=1}^{n} x_k \, C_k.$$

1.9.2 Définitions et premières propriétés

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

$$AX = \lambda X.$$

Définition 1.9.1. *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.

Il est important de noter qu'il est indispensable de spécifier ci-dessus « vecteur non nul » car il est clair que la relation $AX = \lambda X$ est vérifiée pour $X = 0$ et n'importe quel complexe λ.

Si A est une matrice carrée, un simple calcul de déterminant (appliquer la définition !) montre que la fonction

$$\lambda \mapsto \det(A - \lambda \mathbb{1})$$

est un polynôme en la variable λ de degré n si A est de dimension n, dont le coefficient de λ^n est $(-1)^n$ et dont le terme indépendant est $\det(A)$.

Cela étant, voici une propriété extrêmement utile pour la recherche des valeurs propres. Comme il s'agit en fait d'un résultat faisant intervenir une notion d'équivalence (ici « si et seulement si »), ce résultat donne une alternative pour une autre définition de la notion de valeur propre.

Théorème 1.9.2. *Les valeurs propres de A sont les zéros du polynôme*

$$\lambda \mapsto P(\lambda) = \det(A - \lambda \mathbb{1}).$$

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

Preuve. 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

$$(A - \lambda_0 \mathbb{1})X_0 = 0.$$

Si le déterminant de $A - \lambda_0 \mathbb{1}$ était non nul, on aurait, en appliquant l'inverse de la matrice $A - \lambda_0 \mathbb{1}$ aux deux membres de l'égalité

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

ce qui est contradictoire. On a donc obtenu que le complexe λ_0 est effectivement un zéro du polynôme $\lambda \mapsto P(\lambda) = \det(A - \lambda \mathbb{1})$.

Démontrons à présent que si le complexe λ_0 vérifie $\det(A - \lambda_0 \mathbb{1}) = 0$ alors c'est une valeur propre de A. Notons $C_1, \ldots C_n$ les colonnes de A. Etant donné l'annulation du déterminant de la matrice $A - \lambda_0 \mathbb{1}$, l'une des colonnes de cette dernière est combinaison linéaire des autres : il existe $k \in \{1, \ldots, n\}$ et des complexes c_j $(j \in \{1, \ldots, n\} \setminus \{k\})$ tels que

$$C_k - \lambda_0 E_k \;=\; \sum_{j=1,\ j\neq k}^{n} c_j \,(C_j - \lambda_0 E_j)$$

où E_l est le vecteur colonne dont tous les éléments sont nuls sauf le numéro l, qui vaut 1. Cette relation peut encore s'écrire

$$C_k - \sum_{j=1,\ j\neq k}^{n} c_j C_j \;=\; \lambda_0 E_k - \lambda_0 \sum_{j=1,\ j\neq k}^{n} c_j E_j \;=\; \lambda_0 \begin{pmatrix} -c_1 \\ \vdots \\ -c_{k-1} \\ 1 \\ -c_{k+1} \\ \vdots \\ -c_n \end{pmatrix}.$$

En définissant le vecteur non nul X par

$$X = \begin{pmatrix} -c_1 \\ \vdots \\ -c_{k-1} \\ 1 \\ -c_{k+1} \\ \vdots \\ -c_n \end{pmatrix},$$

on obtient finalement

$$AX = \lambda_0 X$$

avec $X \neq 0$ et on conclut. □

Rappelons le théorème fondamental de l'algèbre : un polynôme à coefficients et variable complexes de degré n possède exactement n zéros, comptés avec leur multiplicité. Une matrice carrée de dimension n possède donc toujours exactement n valeurs propres, comptées avec leur multiplicité.

Notons que si A est une matrice diagonale, ou même une matrice triangulaire inférieure ou supérieure, le résultat précédent peut être précisé, comme le montre la propriété ci-dessous.

Propriété(s) 1.9.3. *Si les éléments de la matrice carrée A sous (ou au-dessus de) la diagonale sont tous nuls, alors les valeurs propres de cette matrice sont les éléments diagonaux.*

Preuve. C'est immédiat en utilisant le fait que les valeurs propres d'une matrice carrée sont les zéros du polynôme caractéristique de cette matrice. En effet, le calcul du déterminant de $A - \lambda\mathbb{1}$ est direct en utilisant la définition, vu la présence de nombreux zéros : si les $a_k,\ k = 1, \ldots n$ désignent les éléments diagonaux de A de dimension n, on a

$$\det(A - \lambda\mathbb{1}) = (a_1 - \lambda)\ldots(a_n - \lambda).$$

□

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

Propriété(s) 1.9.4. *Si les vecteurs $X_1, X_2, \ldots, X_J$ sont des vecteurs propres relatifs à la même valeur propre λ_0 de la matrice A, alors, pour tous complexes $c_1, \ldots, c_J$, le vecteur*

$$X = c_1X_1 + \ldots + c_JX_J = \sum_{j=1}^{J} c_jX_j,$$

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

Preuve. Vu les propriétés du produit matriciel, on a

$$AX = A(c_1X_1 + \ldots + c_JX_J) = c_1\ AX_1\ +\ \ldots +\ c_J\ AX_J;$$

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

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

Il s'ensuit que

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

□

Propriété(s) 1.9.5. *Si les n valeurs propres d'une matrice de dimension n sont toutes différentes, alors, quels que soient les vecteurs propres $X_1, \ldots, X_n$ relatifs à ces valeurs propres distinctes, la matrice dont les colonnes sont ces vecteurs est une matrice inversible (c'est-à-dire une matrice dont le déterminant n'est pas nul).*

Preuve. Procédons par récurrence sur $n \geq 2$. On utilise la propriété qui affirme que le déterminant d'une matrice est nul si et seulement si l'une des colonnes de la matrice est combinaison linéaire des autres.

Soient λ_1, λ_2 les valeurs propres distinctes de A et X_1, X_2 des vecteurs propres associés. Si $\det(X_1, X_2) = 0$, alors par exemple $X_1 = rX_2$, avec $r \in \mathbb{C}_0$. On obtient alors

$$AX_1 \;=\; \lambda_1 X_1 \;=\; A(rX_2) \;=\; r\lambda_2 X_2 \;=\; \lambda_2 X_1,$$

ce qui implique $(\lambda_1 - \lambda_2)X_1 = 0$, ce qui est absurde puisque $\lambda_1 \neq \lambda_2$ et $X_1 \neq 0$.

Supposons alors que la propriété soit vraie en dimension n et démontrons-la en dimension $n + 1$. Notons encore λ_k et X_k $(k = 1, \ldots, n + 1)$ respectivement les valeurs propres distinctes de A et des vecteurs propres associés. Si le déterminant formé par ces colonnes est nul, alors il existe $k \in \{1, \ldots, n + 1\}$ et des complexes r_j $(j = 1, \ldots, n + 1,\ j \neq k)$ tels que

$$X_k \;=\; \sum_{j=1, j\neq k}^{n+1} r_j X_j;$$

comme X_k n'est pas nul, l'un des r_j au moins est non nul. On obtient alors, en appliquant la matrice $A - \lambda_k \mathbb{1}$ aux deux membres de l'égalité précédente,

$$(A - \lambda_k \mathbb{1})X_k \;=\; 0 \;=\; \sum_{j=1, j\neq k}^{n+1} r_j (A - \lambda_k \mathbb{1})X_j \;=\; \sum_{j=1, j\neq k}^{n+1} r_j (\lambda_j - \lambda_k)X_j.$$

Cela étant, comme $\lambda_j - \lambda_k \neq 0$ quel que soit $j \neq k$ et comme l'un des r_j n'est pas nul, on en déduit que l'un des n vecteurs X_j $(j \in \{1, \ldots, n + 1\} \setminus \{k\})$ est combinaison linéaire des $n - 1$ autres. Le déterminant formé par ces vecteurs est donc nul, ce qui contredit l'hypothèse de récurrence. □

1.9.3 Exemples

1) Les valeurs propres de la matrice A

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

sont 1 et -1.

En effet, le polynôme caractéristique de A est

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

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

$$(A - 1\mathbb{1})X = (A-\mathbb{1})X = 0;$$

comme

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

on a

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

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

$$(A-(-1)\mathbb{1})X = (A+\mathbb{1})X = 0;$$

comme

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

on a

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

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 A

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

sont i et $-i$.
En effet, le polynôme caractéristique de A est

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

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

$$(A - i\mathbb{1})X = 0;$$

comme

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

on a

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

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

$$(A - (-i)\mathbb{1})X = (A + i\mathbb{1})X = 0;$$

comme

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

on a

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

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 A

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

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

En effet, le polynôme caractéristique de A est

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

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

$$(A - 0\mathbb{1})X = AX = 0;$$

comme

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

on a

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

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 A

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

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

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

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

$$(A - 0\mathbb{1})X = AX = 0;$$

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

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

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

$$(A - 1\mathbb{1})X = (A - \mathbb{1})X = 0;$$

comme

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

on a

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

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 A

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

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

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

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

$$(A - 1\mathbb{1})X = (A - \mathbb{1})X = 0;$$

comme

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

on a

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

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

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

6) Les valeurs propres de la matrice A

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

sont $-1, 2, 3$.

En effet, le polynôme caractéristique de A est

$$\det(A - \lambda \mathbb{1}) = \begin{vmatrix} 7-\lambda & -4 & 8 \\ -2 & 3-\lambda & -2 \\ -5 & 4 & -6-\lambda \end{vmatrix}.$$

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

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

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

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

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

$$(A - (-1)\mathbb{1})X = (A + \mathbb{1})X = 0;$$

comme

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

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

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

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

$$(A - 2\mathbb{1})X = 0;$$

comme

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

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

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

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

$$(A - 3\mathbb{1})X = 0;$$

comme

$$(A - 3\mathbb{1})X = \begin{pmatrix} 4 & -4 & 8 \\ -2 & 0 & -2 \\ -5 & 4 & -9 \end{pmatrix} X = \begin{pmatrix} 4x - 4y + 8z \\ -2x - 2z \\ -5x + 4y - 9z \end{pmatrix}$$

on a (on fait disparaître le terme en y)

$$(A-3\mathbb{1})X=0 \quad\Leftrightarrow\quad \left\{\begin{array}{l}4x-4y+8z=0\\-2x-2z=0\\-x-z=0\end{array}\right. \quad\Leftrightarrow\quad \left\{\begin{array}{l}x-y+2z=0\\x=-z\end{array}\right. \quad\Leftrightarrow\quad \left\{\begin{array}{l}y=z\\x=-z\end{array}\right. \quad\Leftrightarrow\quad X=z\left(\begin{array}{c}-1\\1\\1\end{array}\right).$$

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

$$X=z\left(\begin{array}{c}-1\\1\\1\end{array}\right),\qquad z\in\mathbb{C}_0.$$

7) Les valeurs propres de la matrice

$$L \;=\; \left(\begin{array}{cc}1&8\\1/4&0\end{array}\right)$$

sont 2 et -1.

En effet, le polynôme caractéristique de L est

$$\det(L-\lambda\mathbb{1})=\left|\begin{array}{cc}1-\lambda&8\\1/4&-\lambda\end{array}\right|=(1-\lambda)(-\lambda)-2=\lambda^2-\lambda-2=(\lambda-2)(\lambda+1).$$

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

$$X=\left(\begin{array}{c}x\\y\end{array}\right)$$

non nuls tels que

$$(L-2\mathbb{1})X=0;$$

comme

$$(L-2\mathbb{1})X=\left(\begin{array}{cc}-1&8\\1/4&-2\end{array}\right)X=\left(\begin{array}{c}-x+8y\\x/4-2y\end{array}\right),$$

on a

$$(L-2\mathbb{1})X=0 \quad\Leftrightarrow\quad x=8y \quad\Leftrightarrow\quad X=x\left(\begin{array}{c}8\\1\end{array}\right).$$

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

$$X=c\left(\begin{array}{c}8\\1\end{array}\right),\qquad c\in\mathbb{C}_0.$$

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

$$X=\left(\begin{array}{c}x\\y\end{array}\right)$$

non nuls tels que

$$(L-(-1)\mathbb{1})X=(L+\mathbb{1})X=0;$$

comme

$$(L+\mathbb{1})X=\left(\begin{array}{cc}2&8\\1/4&1\end{array}\right)X=\left(\begin{array}{c}2x+8y\\x/4+y\end{array}\right),$$

on a

$$(L+\mathbb{1})X=0 \quad\Leftrightarrow\quad x=-4y \quad\Leftrightarrow\quad X=y\left(\begin{array}{c}-4\\1\end{array}\right).$$

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

$$X=c\left(\begin{array}{c}-4\\1\end{array}\right),\qquad c\in\mathbb{C}_0.$$

8) Les valeurs propres de la matrice stochastique vue précédemment, à savoir la matrice

$$T = \begin{pmatrix} 0,95 & 0,03 \\ 0,05 & 0,97 \end{pmatrix}$$

sont 1 et $0,92$.

Simplifions les calculs en posant $a = 0,95$, $b = 0,03$. On a alors

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

donc

$$\det(T - \lambda \mathbb{1}) = \begin{vmatrix} a-\lambda & b \\ 1-a & 1-b-\lambda \end{vmatrix} = \lambda^2 - (a+1-b)\lambda + a(1-b) - b(1-a) = \lambda^2 - (a-b+1)\lambda + a - b.$$

Cela étant, le réalisant (discriminant) du polynôme $\lambda \mapsto \lambda^2 - (a-b+1)\lambda + a - b$ est égal à

$$(a-b+1)^2 - 4(a-b) \;=\; (a-b)^2 + 2(a-b) - 4(a-b) + 1 \;=\; (a-b)^2 - 2(a-b) + 1 \;=\; (a-b-1)^2.$$

Comme $a-b-1$ est négatif, on obtient finalement que les valeurs propres cherchées sont les réels

$$\frac{a-b+1 \;+\; (1+b-a)}{2} \;=\; 1 \quad \text{et} \quad \frac{a-b+1 \;-\; (1+b-a)}{2} \;=\; a-b \;=\; 0,92$$

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

$$(T - \mathbb{1})X = 0;$$

comme

$$(T - \mathbb{1})X = \begin{pmatrix} a-1 & b \\ 1-a & -b \end{pmatrix} X = \begin{pmatrix} (a-1)x + by \\ (1-a)x - by \end{pmatrix},$$

on a

$$(T - \mathbb{1})X = 0 \quad \Leftrightarrow \quad (1-a)x = by \quad \Leftrightarrow \quad X = x \begin{pmatrix} 1 \\ (1-a)/b \end{pmatrix}.$$

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

$$X = c \begin{pmatrix} b \\ 1-a \end{pmatrix} \;=\; c \begin{pmatrix} 0,03 \\ 0,05 \end{pmatrix}, \quad c \in \mathbb{C}_0.$$

Cherchons les vecteurs propres associés à la valeur propre $0,92 = a-b$. On doit trouver les vecteurs

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

non nuls tels que

$$(T - (a-b)\mathbb{1})X = 0;$$

comme

$$(T - (a-b)\mathbb{1})X = \begin{pmatrix} b & b \\ 1-a & 1-a \end{pmatrix} X,$$

on a

$$(T - (a-b)\mathbb{1})X = 0 \quad \Leftrightarrow \quad x = -y \quad \Leftrightarrow \quad X = x \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Il s'ensuit que l'ensemble des vecteurs propres associés à la valeur propre $a-b = 0,92$ sont les vecteurs

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

1.9.4 Retour aux matrices stochastiques

Maintenant que l'on sait ce que sont valeurs propres et vecteurs propres d'une matrice carrée, on revient au traitement du problème posé précédemment, modélisé par des matrices stochastiques.

Rappelons et complétons tout d'abord les définitions.

Définition 1.9.6.
— *Une* matrice stochastique *est une matrice 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 positifs*[3].
— *Si T est une matrice stochastique et X_0 un vecteur de probabilité, la* chaîne de Markov *associée est la suite de vecteurs*

$$X_0,\ X_1 = TX_0,\ X_2 = TX_1 = T^2X_0,\ \ldots$$

Propriété(s) 1.9.7 (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 strictement positif, 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 $j \in \{1,\ldots,n\}$, 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 $C_1,\ldots,C_n$ les colonnes de T^k, la matrice T^{k+1} est

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

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 $C_1,\ldots,C_n$. 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 $\widetilde{T}X = X$, par définition des matrices stochastiques. Le nombre 1 est donc valeur propre de $\widetilde{T}$, donc de T puisque ces matrices possèdent les mêmes valeurs propres.
Soit maintenant une valeur propre quelconque λ de T et soit un vecteur propre X de $\widetilde{T}$ relatif

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

à celle-ci. Si $k \in \{1, \ldots, n\}$ est tel que $|(X)_k| = \sup \{|(X)_j| \ : 1 \leq j \leq n\}$ alors $|(X)_k| \neq 0$ et

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

□

Propriété(s) 1.9.8 (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

$$\begin{aligned}
\det(T - \lambda \mathbb{1}) &= \det \begin{pmatrix} a-\lambda & b \\ 1-a & 1-b-\lambda \end{pmatrix} \\
&= \det \begin{pmatrix} a-\lambda & b \\ 1-a+a-\lambda & 1-b-\lambda+b \end{pmatrix} \\
&= \det \begin{pmatrix} a-\lambda & b \\ 1-\lambda & 1-\lambda \end{pmatrix} \\
&= (1-\lambda) \det \begin{pmatrix} a-\lambda & b \\ 1 & 1 \end{pmatrix} \\
&= (1-\lambda) \ \Big((a-b) - \lambda\Big),
\end{aligned}$$

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 élements de X_0, comme $T^k X$ est un vecteur de probabilité quel que soit k, on a

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

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. □

Cela étant, reprenons l'exemple introductif (mouvement de population ville-faubourg).

La modélisation détermine une matrice stochastique. Les calculs faits précédemment (cf exemples de recherche de valeurs propres et de vecteurs propres) ont conduit à trouver que les vecteurs propres de valeur propre 1 sont les vecteurs

$$X = c \begin{pmatrix} b \\ 1-a \end{pmatrix} = c \begin{pmatrix} 0,03 \\ 0,05 \end{pmatrix}, \quad c \in \mathbb{C}_0.$$

Parmi ces vecteurs, celui qui est de probabilité est celui où $c = 1/(0,03 + 0,05) = 1/0,08$. On obtient donc

$$X = \begin{pmatrix} 0.03/0,08 \\ 0,05/0,08 \end{pmatrix} = \begin{pmatrix} 3/8 \\ 5/8 \end{pmatrix} = \begin{pmatrix} 0,375 \\ 0,625 \end{pmatrix}.$$

Ainsi, à partir d'une population de départ quelconque de 1 000 000 d'habitants, on évolue vers une population de

375 000 habitants en ville et 625 000 habitants dans les faubourgs.

1.9.5 La diagonalisation des matrices carrées

Définition 1.9.9. *Une matrice A est diagonalisable s'il existe une matrice inversible S de même dimension telle que la matrice $S^{-1}AS$ soit une matrice diagonale.*

Définition 1.9.10. *Si la matrice inversible S est telle que $S^{-1}AS$ est 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.

Identifions les éléments diagonaux de la matrice $S^{-1}AS$ lorsque celle-ci est diagonale.

Propriété(s) 1.9.11. *Si S est une matrice inversible telle que $S^{-1}AS$ est une matrice diagonale, alors les éléments diagonaux de cette matrice diagonale sont les valeurs propres de A.*

Preuve. On sait que les valeurs propres d'une matrice B sont les zéros du polynôme $\lambda \mapsto \det(B - \lambda\mathbb{1})$ et que les valeurs propres d'une matrice diagonale sont les éléments diagonaux de celle-ci. Pour prouver le résultat annoncé, il suffit donc de montrer que les matrices A et $S^{-1}AS$ ont le même polynôme caractéristique.

De fait, on a

$$\begin{aligned} \det\left(S^{-1}AS - \lambda\mathbb{1}\right) &= \det\left(S^{-1}AS - \lambda S^{-1}\mathbb{1}S\right) \\ &= \det\left(S^{-1}(A - \lambda\mathbb{1})S\right) \\ &= \det\left(S^{-1}\right)\,\det(A - \lambda\mathbb{1})\,\det(S), \end{aligned}$$

la dernière égalité étant obtenue par le fait que le déterminant d'un produit de matrices carrées de même dimension est égal au produit des déterminants des matrices. Ensuite, encore à cause de cette propriété et du fait que $S^{-1}S = \mathbb{1}$, on a $1 = \det(S^{-1})\,\det(S)$ donc finalement

$$\det\left(S^{-1}AS - \lambda\mathbb{1}\right) = \det\left(S^{-1}\right)\,\det(A - \lambda\mathbb{1})\,\det(S) = \det(A - \lambda\mathbb{1}).$$

□

Venons-en maintenant à une autre propriété de la diagonalisation, laquelle va fournir un critère de diagonalisabilité.

Commençons par établir une propriété de manipulation de certaines égalités matricielles, lesquelles seront bien utiles pour obtenir la caractérisation de la diagonalisabilité d'une matrice carrée.

Propriété(s) 1.9.12. *Soit une matrice carrée A de dimension n. Notons $\lambda_1, \ldots, \lambda_n$ ses valeurs propres et $X_1, \ldots, X_n$ des vecteurs propres associés à celles-ci, respectivement. Si S est la matrice carrée de dimension n dont les colonnes sont respectivement $X_1, \ldots, X_n$ alors on a l'égalité*

$$A\,S = S\,diag(\lambda_1, \ldots, \lambda_n).$$

Preuve. D'une part, par définition du produit matriciel, AS est la matrice dont les colonnes sont $AX_1, \ldots, AX_n$, ou encore $\lambda_1 X_1, \ldots, \lambda_n X_n$ puisque X_k est vecteur propre de valeur propre λ_k quel que soit $k \in \{1, \ldots, n\}$.

D'autre part, par définition du produit matriciel encore, la matrice $S\,diag(\lambda_1, \ldots, \lambda_n)$ a pour vecteurs-colonnes les vecteurs

$$S \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \lambda_k \text{ à la ligne numéro } k,$$

c'est-à-dire les vecteurs $\lambda_k X_k$ $(k = 1, \ldots, n)$.

Les matrices AS et $S\,diag(\lambda_1, \ldots, \lambda_n)$ ayant les mêmes colonnes, elles sont égales. □

Propriété(s) 1.9.13. *Si S est une matrice inversible telle que $S^{-1}AS$ est une matrice diagonale, alors les colonnes de S sont des vecteurs propres de la matrice A.*

Si S est une matrice dont les colonnes sont des vecteurs propres de A et qui est inversible, alors la matrice $S^{-1}AS$ est une matrice diagonale.

Preuve. Si

$$S^{-1}AS \;=\; diag(\lambda_1, \ldots, \lambda_n),$$

on a aussi

$$A\,S \;=\; S\,diag(\lambda_1, \ldots, \lambda_n),$$

ce qui montre que les colonnes de S sont des vecteurs propres de A vu la forme des colonnes de la matrice du membre de droite de l'égalité.

Réciproquement, si S est une matrice dont les colonnes $X_1, \ldots, X_n$ sont des vecteurs non nuls tels que $AX_k = \lambda_k X_k$ pour tout $k = 1, \ldots, n$, alors on a

$$A\,S \;=\; S\,diag(\lambda_1, \ldots, \lambda_n).$$

Dès lors, si S est inversible, on obtient bien que $S^{-1}AS$ est une matrice diagonale. □

De ce qui précède on déduit alors le critère suivant.

Propriété(s) 1.9.14. *Une matrice de dimension n est diagonalisable si et seulement s'il existe des vecteurs propres $X_1, \ldots, X_n$ tels que la matrice formée par ceux-ci soit inversible.*

Il est utile de se rappeler ici la propriété 1.9.5, laquelle implique alors qu'*une matrice dont les valeurs propres sont distinctes est toujours diagonalisable.*

Voici en détail ce qui se passe pour les matrices de dimension 2*.*

Deux cas peuvent se présenter. Dans le cas où les valeurs propres de la matrice sont distinctes, la matrice est diagonalisable. Dans le cas où on a une valeur propre double, la matrice est diagonalisable si et seulement si elle admet deux vecteurs propres qui ne sont pas multiples l'un de l'autre.

Et voici en détail ce qui se passe pour les matrices de dimension 3*.*

Ici, on a davantage de situations différentes.

(1) Dans le cas où les valeurs propres de la matrice sont distinctes, la matrice est diagonalisable.

(2) Dans le cas où une valeur propre est double, la matrice est diagonalisable si et seulement si elle admet deux vecteurs propres relatifs à la valeur propre double qui ne sont pas multiples

l'un de l'autre.
(3) Dans le cas où la matrice admet une valeur propre triple, la matrice est diagonalisable si et seulement si elle admet trois vecteurs propres relatifs à la valeur propre triple qui forment une matrice inversible lorsqu'on les place sur les colonnes de cette dernière.

1.9.6 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 A

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

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 A

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

a comme valeurs propres les complexes i et $-i$. La matrice est donc diagonalisable car elle est de dimension 2 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 A

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

a la valeur propre double 0.

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

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

Ils sont tous multiples du vecteur

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

deux vecteurs propres sont donc toujours multiples l'un de l'autre. Il s'ensuit que la matrice A n'est pas diagonalisable.

4) La matrice A

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

a les valeurs propres 0 et 1 (valeur propre double 0 et 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 qui ne sont pas multiples l'un de l'autre.

Le vecteur

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

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

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 A

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

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 forment donc toujours une matrice dont le déterminant est nul. La matrice de départ n'est donc pas diagonalisable.

6) Les valeurs propres de la matrice A

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

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

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

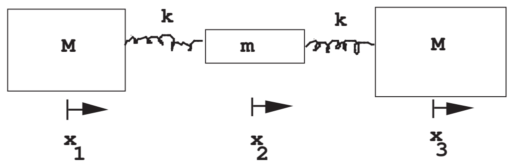
1.10 Application : découplage d'équations différentielles

Nous allons expliquer comment la diagonalisation peut être utile dans la résolution d'équations différentielles couplées (Exemple des modes normaux de vibration).

1.10.1 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 x_1, x_2, x_3 par rapport à leur position d'équilibre. La physique indique que ce sytème est régi par les équations différentielles suivantes (t=variable temporelle, m, M, k constantes strictement positives)

$$\left\{\begin{array}{lcl} D^2x_1 &=& -\frac{k}{M}(x_1-x_2) \\ D^2x_2 &=& -\frac{k}{m}(x_2-x_1)-\frac{k}{m}(x_2-x_3) \\ D^2x_3 &=& -\frac{k}{M}(x_3-x_2). \end{array}\right.$$



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.

1.10.2 Résolution

Le système précédent peut s'écrire sous la forme

$$D^2\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -\frac{k}{M} & \frac{k}{M} & 0 \\ \frac{k}{m} & -\frac{2k}{m} & \frac{k}{m} \\ 0 & \frac{k}{M} & -\frac{k}{M} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{1.1}$$

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).

En remplaçant d'abord la première colonne par la somme des trois colonnes, puis en faisant une mise en évidence et enfin en remplaçant les deuxième et troisième lignes par la différence entre ces lignes et la première, on obtient

$$\begin{aligned} \begin{vmatrix} -\frac{k}{M}-\lambda & \frac{k}{M} & 0 \\ \frac{k}{m} & -\frac{2k}{m}-\lambda & \frac{k}{m} \\ 0 & \frac{k}{M} & -\frac{k}{M}-\lambda \end{vmatrix} &= \begin{vmatrix} -\lambda & \frac{k}{M} & 0 \\ -\lambda & -\frac{2k}{m}-\lambda & \frac{k}{m} \\ -\lambda & \frac{k}{M} & -\frac{k}{M}-\lambda \end{vmatrix} \\ &= -\lambda\begin{vmatrix} 1 & \frac{k}{M} & 0 \\ 1 & -\frac{2k}{m}-\lambda & \frac{k}{m} \\ 1 & \frac{k}{M} & -\frac{k}{M}-\lambda \end{vmatrix} \\ &= -\lambda\begin{vmatrix} 1 & \frac{k}{M} & 0 \\ 0 & -\frac{2k}{m}-\frac{k}{M}-\lambda & \frac{k}{m} \\ 0 & 0 & -\frac{k}{M}-\lambda \end{vmatrix} \\ &= -\lambda\left(\frac{k}{M}+\lambda\right)\left(\lambda+\frac{2k}{m}+\frac{k}{M}\right). \end{aligned}$$

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 $-k/M$ 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 $-k/M - 2k/m$ vérifient

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

Si

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

cette équation est équivalente au système

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

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

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

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 2).

Résolution de $D^2y_1 = 0$.

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

$$y_1(t) = r_1t + r_2, \quad t \in \mathbb{R}$$

où $r_1, r_2 \in \mathbb{R}$ sont des constantes réelles arbitraires.

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

On a

$$D^2y_2 = -\frac{k}{M}y_2 \ \Leftrightarrow \ D^2y_2 + \frac{k}{M}y_2 = 0.$$

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

$$y_2(t) = c_1e^{i\sqrt{k/M}t} + c_2e^{-i\sqrt{k/M}\,t}, \quad t \in \mathbb{R}$$

où $c_1, c_2 \in \mathbb{C}$ sont des constantes complexes arbitraires ; l'ensemble des solutions réelles est l'ensemble des fonctions qui s'écrivent

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

où $r_1, r_2 \in \mathbb{R}$ sont des constantes réelles arbitraires.

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

On a

$$D^2y_3 = -\left(\frac{k}{M}+\frac{2k}{m}\right)y_3 \ \Leftrightarrow \ D^2y_3 + \left(\frac{k}{M}+\frac{2k}{m}\right)y_3 = 0.$$

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

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

où $r_1, r_2 \in \mathbb{R}$ sont des constantes arbitraires.

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**.

1.10.3 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 :

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

qui correspondent respectivement aux mouvements suivants

(a) $x_1 = x_2 = x_3$:
les masses bougent sans vibration entre elles

(b) $x_1 = -x_3,\ x_2 = 0$:
immobilité de la masse centrale et mouvement opposé des masses extrêmes

(c) $x_1 = x_3,\ x_2 = -2M/m\ x_3$:
même mouvement pour les masses extrêmes et mouvement opposé pour la masse centrale.

Chapitre 2

Fonctions de plusieurs variables

2.1 Introduction, définitions, représentations

Jusqu'à présent, nous avons considéré des fonctions d'une variable réelle. 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.

2.1.1 Définitions, représentations

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*

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

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

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

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

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

— Dans un repère orthonormé, la distance entre un point P de coordonnées (x, y, z) et l'origine des axes s'exprime par

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

— 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) :

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

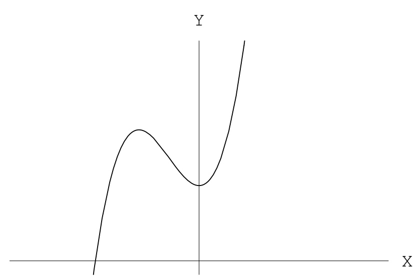
(c est une constante strictement positive).

Représenter une fonction d'une variable réelle à valeurs réelles, c'est représenter son graphe, à savoir l'ensemble

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

Pour cela, on utilise (le plus souvent) 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.*

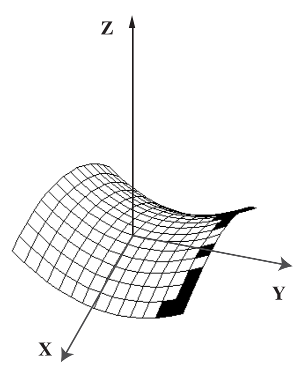


De façon analogue, représenter une fonction de deux variables réelles, à valeurs réelles, c'est représenter l'ensemble

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

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 ≫ $(x, y, z, f(x, y, z))$! Pour avoir une idée du comportement des valeurs de la fonction, on en considère alors des *surfaces de niveau*, lesquelles sont définies dans ce qui suit.

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.

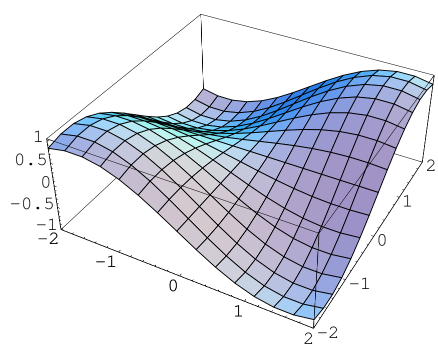
Si f est une fonction de deux variables et si r appartient à l'image de f, l'ensemble

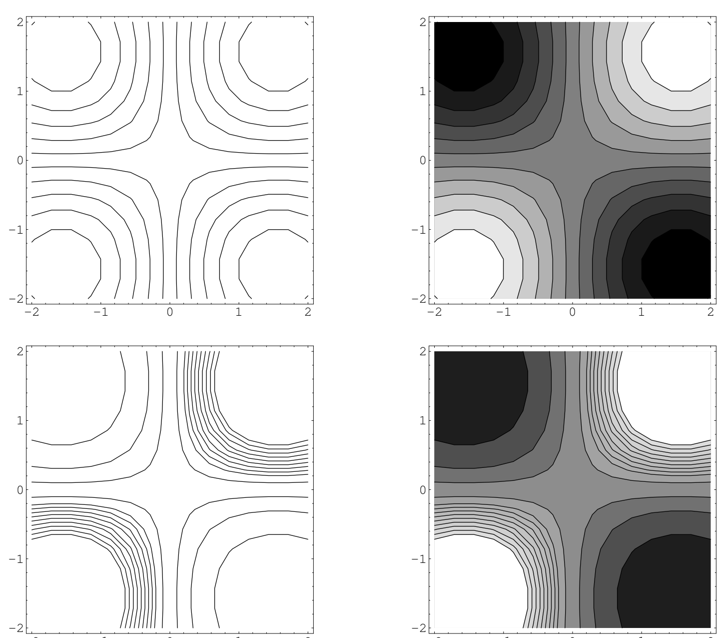
$$\{(x, y) \in dom(f) \ : \ f(x, y) = r\}$$

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.





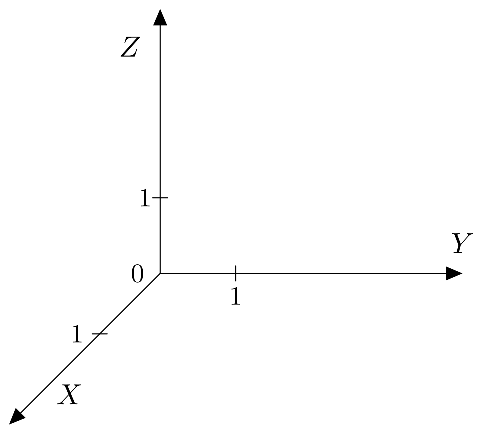
Remarquons que la représentation graphique d'une fonction f d'une variable est une courbe de niveau. En effet, si on pose $g(x, y) = y - f(x)$, on a

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

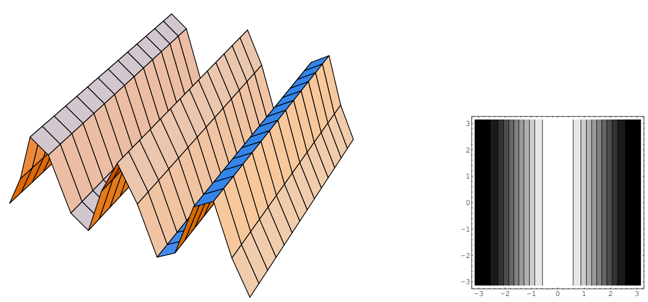
donc la représentation de f est une courbe de niveau de $g(x, y) = y - f(x)$.

Exemples

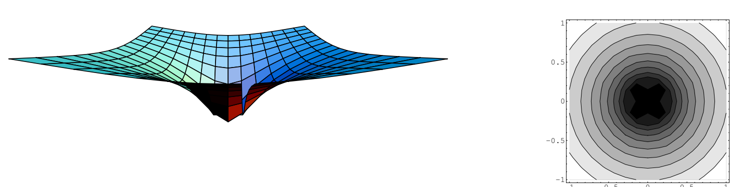
Dans ce qui suit, les représentations à trois dimensions sont toujours faites dans un sytème du type représenté ci-dessous. Nous avons omis les axes afin de ne pas alourdir les figures.



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



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



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

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

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

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

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

Surfaces quadriques

Dans l'espace muni d'un repère, de façon analogue à la définition des coniques du plan via une équation cartésienne, on appelle *surface quadrique* l'ensemble des points dont les coordonnées cartésiennes x, y, z vérifient une équation du type

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + a_{12}xy + a_{13}xz + a_{23}yz + b_1x + b_2y + b_3z + c = 0$$

où les a_{ij}, b_i, c sont des réels tels que les a_{ij} ne soient pas tous nuls.

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.

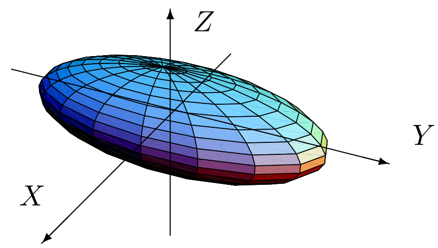
Une aide à la représentation d'une quadrique consiste à considérer les traces de cette surface dans des plans orthogonaux aux axes (c'est-à-dire les intersections de la surface avec des plans orthogonaux aux axes) [1].

— $\boxed{\text{Ellipsoïde}}$: 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.

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.



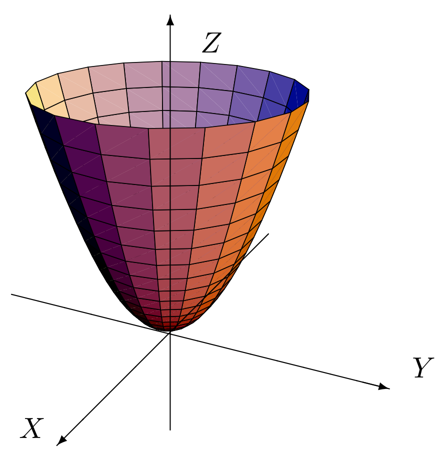
La trace de cette surface sur un plan orthogonal à l'un des axes de coordonnées est vide ou est une ellipse. Par exemple, si $z = r$ avec $r \in]-c, c[$ alors la trace correspondante a pour équation $x^2/a^2 + y^2/b^2 = 1 - r^2/c^2$; ceci est bien l'équation d'une ellipse.

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).

— $\boxed{\text{Paraboloïde elliptique}}$: surface d'équation

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

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

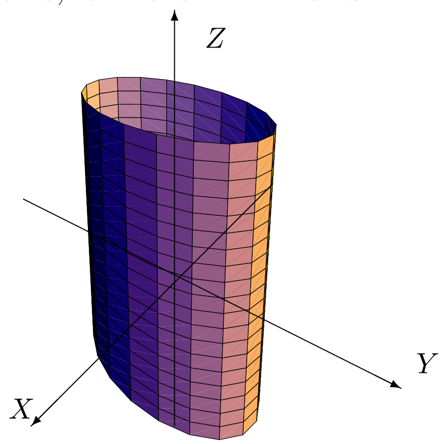


La trace sur un plan orthogonal à l'axe Z est vide ou est une ellipse. La trace sur un plan orthogonal à l'axe X (ou Y) est une parabole.

— $\boxed{\text{Cylindre elliptique}}$: surface d'équation

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

où a, b sont des réels strictement positifs.

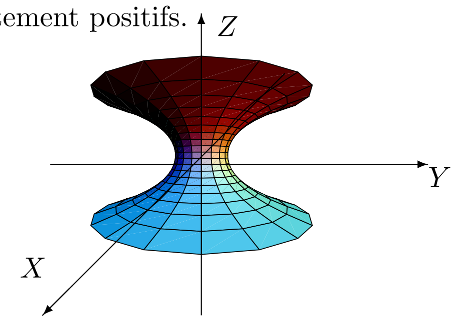


La trace sur un plan orthogonal à l'axe Z est une ellipse. La trace sur un plan orthogonal à l'axe X (ou Y) est vide ou formée de deux droites parallèles.

— $\boxed{\text{Hyperboloïde à une nappe}}$: 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.

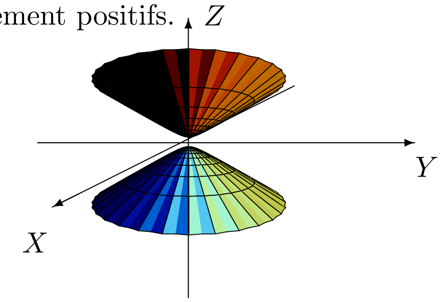


La trace sur un plan orthogonal à l'axe Z est une ellipse ; la trace sur un plan orthogonal à l'axe X (ou à l'axe Y) est une hyperbole.

— $\boxed{\text{Hyperboloïde à deux nappes}}$: 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.

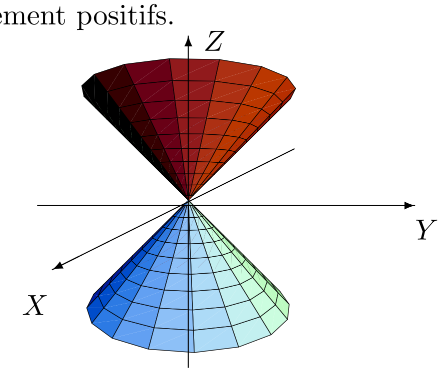


La trace sur un plan orthogonal à l'axe Z est vide ou une ellipse ; la trace sur un plan orthogonal à l'axe X (ou à l'axe Y) est une hyperbole.

— $\boxed{\text{Cône}}$: 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.

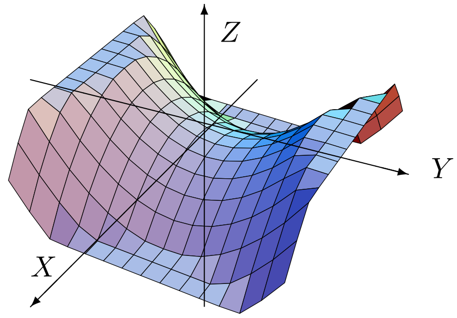


La trace sur un plan orthogonal à l'axe Z ne passant pas par l'origine est une ellipse ; la trace sur le plan d'équation $z = 0$ est l'origine. La trace sur un plan orthogonal à l'axe X (ou Y) est formée par deux droites sécantes si le plan passe par l'origine et est une hyperbole dans les autres cas.

— $\boxed{\text{Paraboloïde hyperbolique}}$: surface d'équation cartésienne

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

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

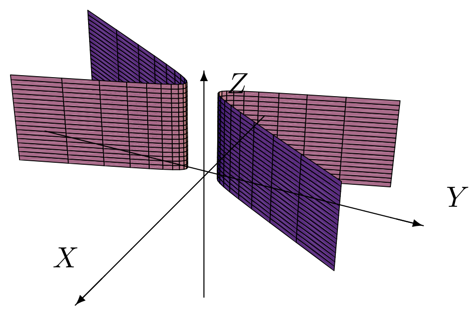


La trace sur un plan orthogonal à l'axe Z ne passant pas par l'origine est une hyperbole ; la trace sur le plan d'équation $z = 0$ est formée par l'union de deux droites sécantes. La trace sur un plan orthogonal à l'axe X (ou Y) est une parabole.

— $\boxed{\text{Cylindre hyperbolique}}$: surface d'équation cartésienne

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

où a, b sont des réels strictement positifs.

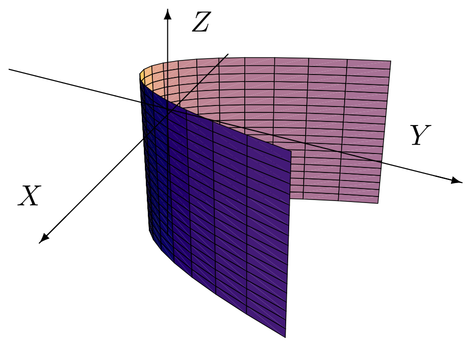


La trace sur un plan orthogonal à l'axe Z est une hyperbole. La trace sur un plan orthogonal à l'axe X (ou Y) est vide ou formée de deux droites parallèles à l'axe Z.

— 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 signes 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.

2.1.2 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).

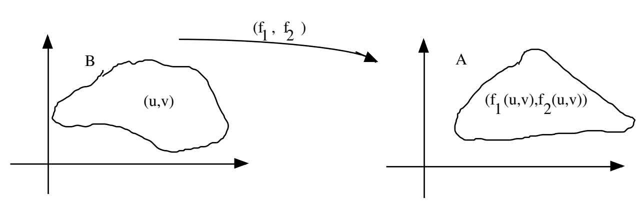
Quant à la composition de fonctions (terme plus général que « fonction de fonction »), elle se définit aussi de manière analogue. Accordons une attention particulière à ce point car il est important lorsque l'on parle de « dérivée partielle, dérivée totale, règle de dérivation en chaîne (« chain rule ») ». 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 $B \subset \mathbb{R}^2$ alors la fonction composée définie explicitement par

$$F(u,v) = f\left(f_1(u,v), f_2(u,v)\right)$$

est définie dans

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



De même si f est de domaine $A \subset \mathbb{R}^3$, si f_1, f_2, f_3 sont à valeurs réelles et définies dans $B \subset \mathbb{R}$ alors la fonction composée définie explicitement par $F(t) = f\left(f_1(t), f_2(t), f_3(t)\right)$ est définie dans

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

2.2 Limites, continuité, dérivation

2.2.1 Limites, continuité

La notion de limite s'introduit de manière analogue au cas d'une variable.

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).

La continuité s'introduit aussi 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 2.2.1. *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*

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) \quad \text{existe.}$$

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.

L'ensemble des fonctions continues sur A est noté

$$C_0(A).$$

2.2.2 Dérivation

DÉFINITIONS

Comme dans le cas des fonctions d'une variable, on travaille dans des ensembles particuliers, appelés ouverts (se rappeler la raison de ceci [2] !) : une partie Ω de $\mathbb{R}^2$ est appelée un ouvert si tout point de Ω est le centre d'un rectangle qui reste inclus dans Ω.

Définition 2.2.2. *Soit f une fonction définie sur un ouvert Ω de $\mathbb{R}^2$ et soit $(x_0, y_0) \in \Omega$.*

a) La fonction f est dérivable par rapport à sa première variable x en (x_0, y_0) si la limite

$$\lim_{h\to 0} \frac{f(x_0+h, y_0) - f(x_0, y_0)}{h}$$

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

$$D_x f(x_0, y_0) \quad \text{ou encore} \quad \frac{\partial f}{\partial x}(x_0, y_0) \qquad \textit{ou encore } D_1 f(x_0, y_0).$$

Remarquons que les expressions

$$\lim_{h\to 0} \frac{f(x_0+h, y_0) - f(x_0, y_0)}{h}$$

2. La définition de la dérivabilité en un point x_0 demande de considérer la valeur de la fonction en $x_0 + h$ (h réel positif ou négatif, proche de 0) ; il est donc indispensable que l'on puisse évaluer la fonction en ces points, ce qui n'est possible que si x_0 est dans un intervalle ouvert

et

$$\lim_{x\to x_0} \frac{f(x,y_0)-f(x_0,y_0)}{x-x_0}$$

signifient exactement la même chose (elles ne diffèrent que par les notations utilisées).

b) La fonction f est dérivable par rapport à sa deuxième variable y en (x_0,y_0) si la limite

$$\lim_{h\to 0} \frac{f(x_0,y_0+h)-f(x_0,y_0)}{h}$$

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

$$D_y f(x_0,y_0) \quad \text{ou encore} \quad \frac{\partial f}{\partial y}(x_0,y_0) \quad \text{ou encore } D_2 f(x_0,y_0).$$

Comme ci-dessus, notons que les expressions

$$\lim_{h\to 0} \frac{f(x_0,y_0+h)-f(x_0,y_0)}{h}$$

et

$$\lim_{y\to y_0} \frac{f(x_0,y)-f(x_0,y_0)}{y-y_0}$$

signifient exactement la même chose (elles ne diffèrent que par les notations utilisées).

c) la fonction f est dérivable sur (ou dans) Ω si elle est dérivable par rapport à x et par rapport à y en tout point de Ω.

Définition 2.2.3. *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 Ω est un ouvert de $\mathbb{R}^2$, la notation

$$C_1(\Omega)$$

désigne l'ensemble des fonctions dérivables dans Ω dont les dérivées partielles sont continues dans Ω. Un élément de cet ensemble s'appelle une fonction continûment dérivable dans (sur) Ω.

Bien sûr ces définitions se généralisent de façon tout à fait naturelle pour les fonctions de plus de deux variables !

Notons aussi que pour bien lever toute ambiguïté, on utilise fréquemment des parenthèses supplémentaires : par exemple la valeur de la dérivée de f par rapport à sa première variable au point (s,t) sera notée

$$(D_1 f)(s,t).$$

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 2.2.4. *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 d'une variable réelle, à valeurs réelles, dérivables sur un ouvert J de $\mathbb{R}$.
Alors la fonction définie par

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

est dérivable sur $I = \{t \in J : (f_1(t), f_2(t)) \in U\}$ et on a

$$DF(t) \;=\; D_1f(X,Y)\,Df_1(t) \;+\; D_2f(X,Y)\,Df_2(t)$$

avec $t \in I$ et $X = f_1(t)$, $Y = f_2(t)$.

b) 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, à valeurs réelles, dérivables sur un ouvert V de $\mathbb{R}^2$.
Alors la fonction définie par

$$F(x,y) = f(f_1(x,y), f_2(x,y))$$

est dérivable sur $\Omega = \{(x,y) \in V : (f_1(x,y), f_2(x,y)) \in U\}$ et on a

$$\begin{aligned} D_1F(x,y) &= D_1f(X,Y)\,D_1f_1(x,y) + D_2f(X,Y)\,D_1f_2(x,y) \\ D_2F(x,y) &= D_1f(X,Y)\,D_2f_1(x,y) + D_2f(X,Y)\,D_2f_2(x,y) \end{aligned}$$

avec $(x,y) \in \Omega$ et $X = f_1(x,y)$, $Y = f_2(x,y)$.

c) Soient
- f une fonction d'une variable réelle dérivable sur un ouvert I de $\mathbb{R}$,
- g une fonction de deux variables réelles dérivable sur l'ouvert V de $\mathbb{R}^2$.
Alors la fonction définie par

$$F(x,y) = f(g(x,y))$$

est dérivable sur $\Omega = \{(x,y) \in V : g(x,y) \in I\}$ et on a

$$\begin{aligned} D_1F(x,y) &= Df(X)\,D_1g(x,y) \\ D_2F(x,y) &= Df(X)\,D_2g(x,y) \end{aligned}$$

avec $(x,y) \in \Omega$ et $X = g(x,y)$.

d) On a bien sûr des résultats analogues pour des fonctions du type $f(f_1, \ldots, f_n)$ avec f fonction de n variables et $f_1, \ldots, f_n$ fonctions de p variables, à valeurs réelles.

Remarque. Pour avoir la dérivabilité de la fonction composée ET l'expression explicite des dérivées présentées dans le résultat précédent, il importe de bien vérifier les hypothèses.
• Il faut 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 (des exemples l'attestent).
• 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, des exemples l'attestent.

Voici un exemple de dérivation de fonction composée.

Soit la fonction F définie explicitement par $F(t) = f\left(\ln(t-1), \sqrt{t^2-1}\right)$ avec f continûment dérivable dans $U =]0,+\infty[\times]4,+\infty[$. On demande où elle est dérivable et l'expression de sa dérivée.

La fonction $t \mapsto \ln(t-1)$ est dérivable sur $]1,+\infty[$ et la fonction $t \mapsto \sqrt{t^2-1}$ est dérivable sur $\{t \in \mathbb{R} : t^2 > 1\} =]-\infty,-1[\cup]1,+\infty[$. L'intersection des domaines de dérivabilité de ces deux fonctions est donc $]1,+\infty[$. Vu la proposition précédente, la fonction F est donc dérivable sur

$$\left\{t \in]1,+\infty[\ :\ \ln(t-1) \in]0,+\infty[\ \text{ et } \ \sqrt{t^2-1} \in]4,+\infty[\right\}.$$

On a $\ln(t-1) \in]0,+\infty[$ si et seulement si $t-1>1$ et on a $\sqrt{t^2-1} \in]4,+\infty[$ si et seulement si $t^2-1>16$. Les deux conditions sur t sont donc

$$\begin{cases} t > 2 \\ |t| > \sqrt{17}, \end{cases}$$

ce qui donne $t > \sqrt{17}$. La fonction F est donc dérivable sur $]\sqrt{17},+\infty[$ et pour tout t dans cet intervalle, on a

$$DF(t) = \frac{D_1 f(X,Y)}{t-1} + \frac{t D_2 f(X,Y)}{\sqrt{t^2-1}}$$

avec $X = \ln(t-1)$ et $Y = \sqrt{t^2-1}$.

2.2.3 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 (des exemples permettent de le montrer). Néanmoins on démontre que *si une fonction est continûment dérivable sur un ouvert Ω de $\mathbb{R}^2$ alors cette fonction y est continue*; ce résultat est aussi valable quand on travaille avec un ouvert de $\mathbb{R}^n$ avec $n > 2$.

2.2.4 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)

$$D_1 f, \quad D_2 f$$

sont encore dérivables. Si c'est le cas, on introduit alors les fonctions dérivées d'ordre deux

$$D_1 D_1 f = D_1^2 f,\ D_2 D_1 f,\ D_2 D_2 f = D_2^2 f,\ D_1 D_2 f.$$

Dans beaucoup de cas, on a $D_1 D_2 f = D_2 D_1 f$ mais pas toujours! Néanmoins on démontre que cette égalité est correcte si la fonction est deux fois continûment dérivable.

On introduit de manière analogue les dérivées d'ordre p pour tout naturel non nul p.

On a aussi la propriété suivante : *Si, dans les énoncés de la Proposition (2.2.4), les fonctions f, f_1, f_2 sont q fois continûment dérivables, alors la fonction composée F est aussi q fois continûment dérivable.*

2.2.5 Des opérateurs de dérivation fort utiles

En physique (notamment), les opérateurs de dérivation appelés

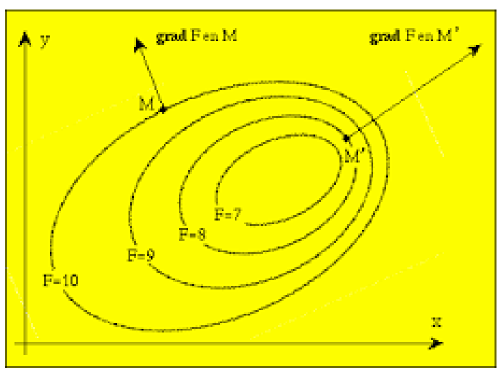
gradient, divergence, rotationnel

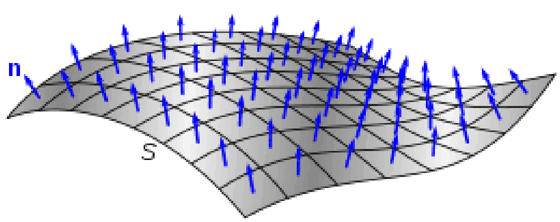
apparaissent à de multiples endroits, comme outils de modélisation de divers phénomènes (flux au travers de parois, équations de Maxwell en électromagnétisme, équation de la chaleur, des ondes, ...). Il s'agit d'*opérateurs faisant intervenir des dérivées partielles* et les équations impliquant celles-ci 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.

Donnons simplement ici la définition de l'opérateur gradient : si f est une fonction dérivable dans un ouvert Ω de $\mathbb{R}^n$ alors le gradient de f est la fonction à valeurs vectorielles

$$grad\, f \ : \ x \in \mathbb{R}^n \mapsto \ (D_1 f(x), \ldots, D_n f(x)).$$

Dans le cas de deux variables ($n = 2$), cas régulier (c'est-à-dire qu'on a la dérivabilité requise), le gradient en un point est un vecteur orthogonal à la tangente à la courbe d'équation $f(x, y) = 0$ en ce point. Dans le cas de trois variables ($n = 3$), cas régulier (c'est-à-dire qu'on a la dérivabilité requise) le gradient en un point est un vecteur orthogonal au plan tangent à la surface d'équation $f(x, y, z) = 0$ en ce point. Lorsque $n = 2$ et que la courbe est le graphique d'une fonction g (c'est-à-dire $f(x, y) = y - g(x)$), cela se justifie aisément : les composantes d'un vecteur directeur de la tangente au point du graphique d'abscisse x est $(1, Dg(x))$; le gradient de f dans ce cas est $(D_1 f(x, y), D_2 f(x, y)) = (-Dg(x), 1)$; et on a bien l'orthogonalité des vecteurs de composantes $(-Dg(x), 1)$ et $(1, Dg(x))$.





2.2.6 La dérivée directionnelle

En analyse mathématique, la notion de dérivée directionnelle permet de quantifier la variation locale d'une fonction dépendant de plusieurs variables, en un point donné et le long d'une direction donnée dans l'espace de ces variables. Dans la version la plus simple, la dérivée directionnelle généralise la notion de dérivées partielles : on les retrouve en prenant comme directions de dérivation les axes de coordonnées.

Voici la définition et une propriété qui relie cette dérivée aux dérivées partielles.

Définition 2.2.5. *Cas de deux variables. Soit f une fonction définie dans un ouvert de $\mathbb{R}^2$, soit un point (x_0, y_0) de cet ouvert et soit un vecteur non nul de composantes (a, b). On dit que f admet une dérivée dans la direction (a, b) au point (x_0, y_0) si la limite suivante existe et est finie*

$$\lim_{t\to 0} \frac{f(x_0 + ta, y_0 + tb) - f(x_0, y_0)}{t}.$$

Si on pose $u = (a, b)$, on utilise fréquemment la notation $D_u f(x_0, y_0)$ pour désigner la valeur de cette limite et on l'appelle la dérivée directionnelle en (x_0, y_0) dans la direction u.

Cas de trois variables. Soit f une fonction définie dans un ouvert de $\mathbb{R}^3$, soit un point (x_0, y_0, z_0) de cet ouvert et soit un vecteur non nul de composantes (a, b, c). On dit que f admet une dérivée dans la direction (a, b, c) au point (x_0, y_0, z_0) si la limite suivante existe et est finie

$$\lim_{t\to 0} \frac{f(x_0 + ta, y_0 + tb, z_0 + tc) - f(x_0, y_0, z_0)}{t}.$$

Si on pose $u = (a, b, c)$, on utilise fréquemment la notation $D_u f(x_0, y_0, z_0)$ pour désigner la valeur de cette limite et on l'appelle la dérivée directionnelle en (x_0, y_0, z_0) dans la direction u.

On obtient alors tout de suite un lien entre cette notion et celle des dérivées partielles (qui en sont bien sûr un cas particulier comme on le voit en prenant les vecteurs $(1, 0)$, $(0, 1)$ dans le cas de deux variables et $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ dans le cas de trois variables). Clairement, c'est la dérivation des fonctions composées qui va intervenir ici.

Proposition 2.2.6. *Cas de deux variables. Si Ω est un ouvert de $\mathbb{R}^2$ et si $f \in C_1(\Omega)$ alors la dérivée directionnelle existe en tout point (x, y) de l'ouvert en toute direction (a, b) et on a*

$$D_u f(x, y) \;=\; D_1 f(x, y)\, a \;+\; D_2 f(x, y)\, b.$$

Cas de trois variables. Si Ω est un ouvert de $\mathbb{R}^3$ et si $f \in C_1(\Omega)$ alors la dérivée directionnelle existe en tout point (x, y, z) de l'ouvert en toute direction (a, b, c) et on a

$$D_u f(x, y, z) \;=\; D_1 f(x, y, z)\, a \;+\; D_2 f(x, y, z)\, b \;+\; D_3 f(x, y, z)\, c.$$

Preuve. C'est une application du théorème de dérivation des fonctions composées. En effet avec

$$F(t) \;=\; f(x + ta, y + tb),$$

on a

$$\begin{aligned} D_u f(x, y) &= \lim_{t\to 0} \frac{f(x + ta, y + tb) - f(x, y)}{t} \\ &= \lim_{t\to 0} \frac{F(t) - F(0)}{t} \\ &= DF(0) \\ &= D_1 f(x, y)\, a \;+\; D_2 f(x, y)\, b. \end{aligned}$$

Avec

$$F(t) \;=\; f(x + ta, y + tb, z + tc),$$

on a

$$\begin{aligned} D_u f(x, y, z) &= \lim_{t\to 0} \frac{f(x + ta, y + tb, z + tc) - f(x, y, z)}{t} \\ &= \lim_{t\to 0} \frac{F(t) - F(0)}{t} \\ &= DF(0) \\ &= D_1 f(x, y, z)\, a \;+\; D_2 f(x, y, z)\, b \;+\; D_3 f(x, y, z)\, c. \end{aligned}$$

□

Une question qui se pose de façon naturelle dans les modélisations, quand f représente une grandeur, une quantité ..., est de savoir dans quelle direction cette grandeur, quantité croît ou décroît le plus vite en un point donné. Autrement dit, on se demande quelle est la direction u pour laquelle $D_u f(x,y)$ (resp. $D_u f(x,y,z)$) est le plus grand ou le plus petit (on examine en fait la « vitesse » de variation de f dans une direction).

La propriété suivante [3] va permettre de répondre à la question.

Propriété(s) 2.2.7. *Pour tous points (a,b), (x,y) de $\mathbb{R}^2$, on a*

$$|ax+by| \ \leq \ \sqrt{a^2+b^2}\ \sqrt{x^2+y^2},$$

l'égalité ayant lieu si et seulement si (a,b) est multiple de (x,y) ou (x,y) est multiple de (a,b).

Pour tous points (a,b,c), (x,y,z) de $\mathbb{R}^3$, on a

$$|ax+by+cz| \ \leq \ \sqrt{a^2+b^2+c^2}\ \sqrt{x^2+y^2+z^2},$$

l'égalité ayant lieu si et seulement si (a,b,c) est multiple de (x,y,z) ou (x,y,z) est multiple de (a,b,c).

Preuve. On l'obtient directement par le calcul.□

Revenons maintenant au cas qui nous occupe. Que ce soit dans le cas de deux ou trois variables, en supposant que le gradient de f en (x,y) (resp. en (x,y,z)) n'est pas nul, les expressions

$$D_u f(x,y) = D_1 f(x,y)\,a \ + \ D_2 f(x,y)\,b$$

$$D_u f(x,y,z) = D_1 f(x,y,z)\,a \ + \ D_2 f(x,y,z)\,b \ + \ D_3 f(x,y,z)\,c$$

sont donc bien extrémales lorsque (a,b) (resp. (a,b,c)) est un multiple du gradient de f en (x,y) (resp. (x,y,z)). Si a^2+b^2 (resp. $a^2+b^2+c^2$) vaut 1, les valeurs extrémales sont la norme (longueur) du gradient de f en (x,y) (resp. en (x,y,z)) et son opposé.

2.3 « Description d'ensembles »

Pour le calcul intégral à plusieurs variables réelles, il est indispensable de pouvoir manipuler les descriptions géométriques (dessins, lieux) et analytiques (à l'aide de coordonnées cartésiennes, parfois en utilisant aussi un, plusieurs paramètres) d'ensembles. Il est tout aussi indispensable de pouvoir décrire ces ensembles en fixant l'une ou l'autre variable d'abord et ensuite de décrire comment varie la seconde ; nous donnons quelques exemples ci-dessous pour illustrer ce propos.

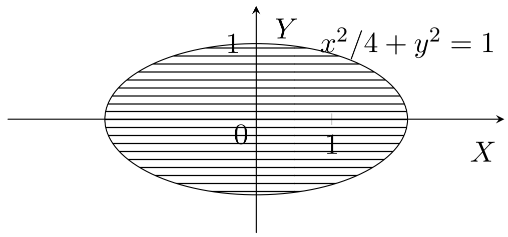
2.3.1 Exemple 1

L'ensemble décrit analytiquement par

$$\left\{(x,y)\ :\ x,y\in\mathbb{R} \text{ et } \frac{x^2}{4}+y^2\leq 1\right\}$$

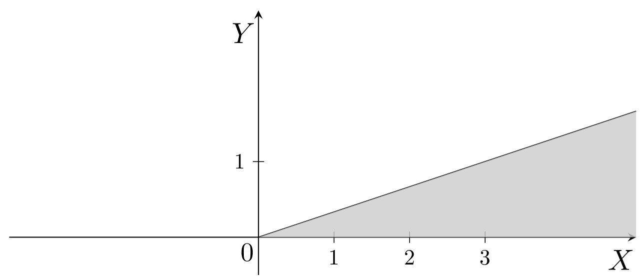
se représente géométriquement par la partie hachurée ci-dessous ; il s'agit des points du plan situés « à l'intérieur » de l'ellipse d'équation cartésienne $x^2/4+y^2=1$ et des points de l'ellipse elle-même.

3. Cette propriété est un cas particulier de ce que l'on appelle l'inégalité de Cauchy-Schwarz, dont l'interprétation est claire avec la définition géométrique du produit scalaire.



2.3.2 Exemple 2

Soit l'ensemble (fermé non borné) grisé suivant



Cet ensemble est limité par la droite d'équation cartésienne $y = 0$ (axe des abscisses) et la droite oblique d'équation cartésienne $x - 3y = 0$. Analytiquement, on a donc

$$\left\{(x, y) \in \mathbb{R}^2 : y \geq 0 \text{ et } x - 3y \geq 0\right\}.$$

Décrivons-le en fixant d'abord x et en donnant ensuite l'ensemble de variation de y. On a

$$\left\{(x, y) \in \mathbb{R}^2 : x \geq 0 \text{ et } y \in [0, x/3]\right\}.$$

Décrivons-le maintenant en fixant d'abord y et en donnant ensuite l'ensemble de variation de x. On a

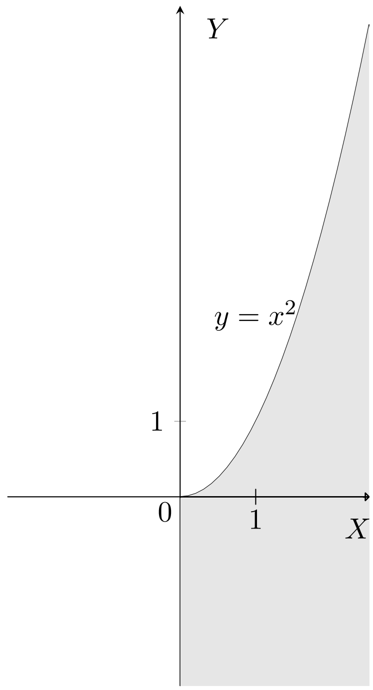
$$\left\{(x, y) \in \mathbb{R}^2 : y \geq 0 \text{ et } x \in [3y, +\infty[\right\}.$$

2.3.3 Exemple 3

Soit l'ensemble

$$A = \left\{(x, y) \in \mathbb{R}^2 \ : \ x > 0 \text{ et } x^2 > y\right\};$$

il s'agit de l'ensemble des points du plan dont l'abscisse est strictement positive et qui sont situés « sous » la parabole d'équation cartésienne $y = x^2$, les points de la parabole n'étant pas compris. Une représentation graphique de cet ensemble est la suivante.



Décrivons cet ensemble en fixant d'abord x et en donnant ensuite l'ensemble de variation de y. On a

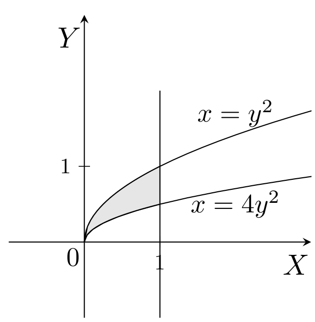
$$A = \left\{(x, y) \in \mathbb{R}^2 \ : \ x > 0 \text{ et } y \in]-\infty, x^2[\right\}.$$

Décrivons-le maintenant en fixant d'abord y et en donnant ensuite l'ensemble de variation de x. On a

$$A = \left\{(x, y) \in \mathbb{R}^2 \ : \ y \geq 0 \text{ et } x \in]\sqrt{y}, +\infty[\right\} \cup \left\{(x, y) \in \mathbb{R}^2 \ : \ y < 0 \text{ et } x \in]0, +\infty[\right\}.$$

2.3.4 Exemple 4

L'ensemble grisé dans la figure ci-dessous



a les descriptions analytiques suivantes

$$\left\{(x, y) \in \mathbb{R}^2 : x \in [0, 1] \text{ et } y \in \left[\sqrt{x}/2, \sqrt{x}\ \right]\right\}$$

et

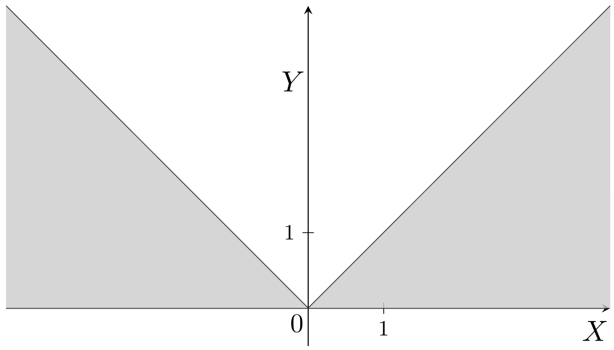
$$\left\{(x, y) \in \mathbb{R}^2 : y \in [0, 1/2] \text{ et } x \in \left[y^2, 4y^2\right]\right\} \cup \left\{(x, y) \in \mathbb{R}^2 : y \in [1/2, 1] \text{ et } x \in \left[y^2, 1\right]\right\}.$$

2.3.5 Exemple 5

L'ensemble décrit analytiquement par

$$A = \{(x,y) \in \mathbb{R}^2 : x^2 - y^2 \geq 0,\ y \geq 0\}$$

se représente comme suit dans un repère orthonormé (en grisé)



On a aussi

$$A = \{(x,y) \in \mathbb{R}^2 : x \in \mathbb{R} \text{ et } 0 \leq y \leq |x|\}$$

et

$$A = \{(x,y) \in \mathbb{R}^2 : 0 \leq y \text{ et } x \in]-\infty, -y]\} \cup \{(x,y) \in \mathbb{R}^2 : 0 \leq y \text{ et } x \in [y, +\infty[\}\,.$$

2.4 Calcul intégral

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 <u>fonctions continues</u>.

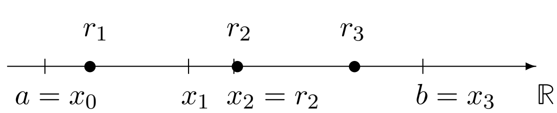
Dans le cadre des fonctions d'une variable réelle, on introduit la notion d'intégrale sur un intervalle borné fermé comme limite de sommes de Riemann obtenues à partir de découpages de l'intervalle. On étend ensuite cette notion au cas des autres intervalles.

Dans $\mathbb{R}^2$ et $\mathbb{R}^3$, 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 et analytiques sont alors beaucoup plus lourdes à manipuler.

2.4.1 Intégration sur des rectangles

Pour rappel, à une variable, on considère un découpage de $[a,b]$ et, dans chacun des sous-intervalles ainsi déterminés, on prend un réel. On construit de la sorte des ≪ sommes de Riemann ≫.



Sous-intervalles : $[a, x_1],\ [x_1, x_2],\ [x_2, b]$

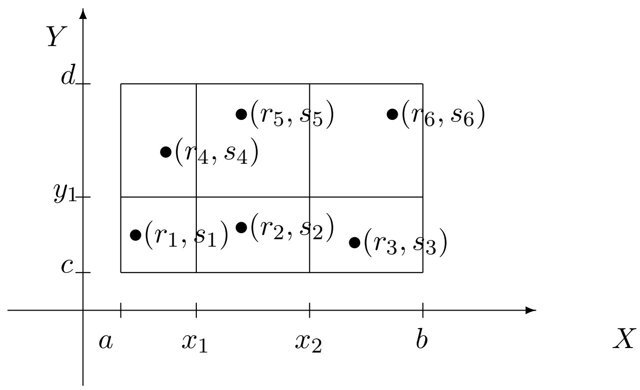
Réels dans les sous-intervalles : $r_1 \in [a, x_1],\ r_2 \in [x_1, x_2],\ r_3 \in [x_2, b]$

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 considère alors des suites de découpages dont la largeur tend vers 0 (cf le rappel consacré au calcul intégral à une variable).

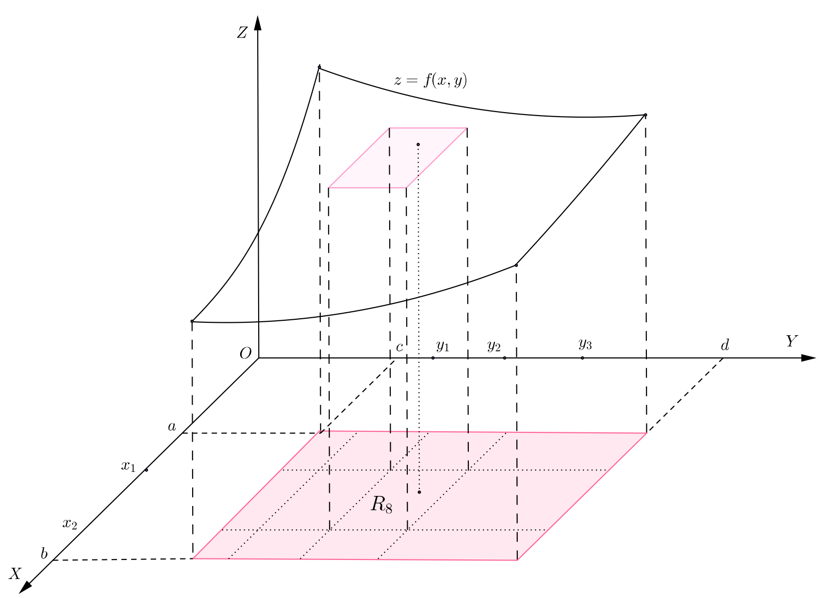
Dans le cas de deux variables, on procède comme dans le cas d'une variable mais à partir d'un rectangle R borné fermé $[a,b] \times [c,d]$. On considère aussi des sous-rectangles, construits à partir de découpages des intervalles $[a,b]$ et $[c,d]$, et dans chacun d'eux, on considère un point.



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.

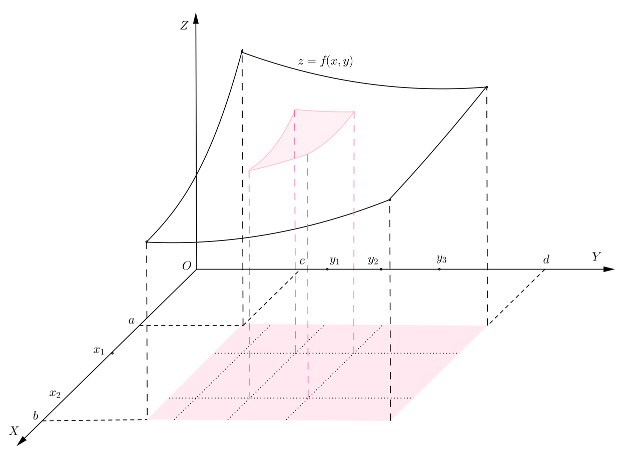
Comme dans le cas d'une variable, on considère des suites de découpages. Lorsque l'on construit une suite de découpages de R en rectangles telle que la suite associée d_n $(n \in \mathbb{N}_0)$ converge vers 0, on examine ici encore le comportement de la suite S_n $(n \in \mathbb{N}_0)$ des sommes de 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 le rectangle R et que son intégrale sur le rectangle R est la limite de la suite S_n $(n \in \mathbb{N}_0)$. L'intégrale de f sur R est notée

$$\iint_R f(x,y)\ dxdy \quad \text{ou} \quad \iint_R f(x,y)\ dydx.$$

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

$$V = \iint_R f(x,y) dxdy.$$



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'admettrons 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 $R = [a,b] \times [c,d]$. On a*

$$\iint_R f(x,y)\ dx\ dy \ = \ \int_a^b \left(\int_c^d f(x,y)\ dy \right)\ dx$$

et

$$\iint_R f(x,y)\ dx\ dy \ = \ \int_c^d \left(\int_a^b f(x,y)\ dx \right)\ dy.$$

On dit que l'on peut permuter l'ordre d'intégration sans changer la valeur de l'intégrale.

Prenons un exemple : on demande de calculer l'intégrale de la fonction f définie explicitement par $f(x,y) = \sin(x+y)$ sur le rectangle $R = [0, \pi/2] \times [-\pi, 0]$.

La fonction f est continue sur $\mathbb{R}^2$ donc elle est intégrable sur le rectangle fermé borné. Cela étant, on a successivement

$$\begin{aligned}
\iint_R \sin(x+y)\, dx\, dy &= \int_0^{\pi/2} \left(\int_{-\pi}^{0} \sin(x+y)\ dy \right)\ dx \\
&= -\int_0^{\pi/2} \left(\int_{-\pi}^{0} D_y(\cos(x+y))\ dy \right)\ dx \\
&= -\int_0^{\pi/2} \Big(\cos(x) - \cos(x-\pi) \Big)\ dx \\
&= -2 \int_0^{\pi/2} \cos(x)\ dx \\
&= -2 \int_0^{\pi/2} D\sin(x)\ dx \\
&= -2\,(\sin(\pi/2) - \sin(0)) \ = \ -2.
\end{aligned}$$

Propriété(s) 2.4.2. *Cas des variables séparées : si $R = [a,b] \times [c,d]$ et si $f(x,y) = f_1(x)\ f_2(y)$ avec f_1 continu sur $[a,b]$ et f_2 continu sur $[c,d]$, alors la fonction f est intégrable sur le rectangle R et on a*

$$\iint_R f_1(x) f_2(y)\ dx\ dy \ = \ \left(\int_a^b f_1(x)\ dx \right) \ \left(\int_c^d f_2(y)\ dy \right).$$

Preuve. C'est direct par le calcul.□

2.4.2 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.*

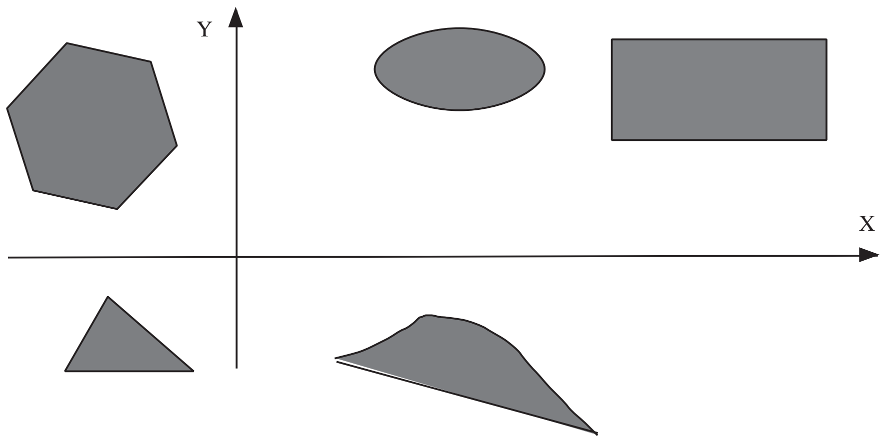
a) On dit que A est parallèle à l'axe Y s'il existe deux fonctions f_1, f_2 continues sur un intervalle $[a,b]$ de $\mathbb{R}$ telles que $f_1 \leq f_2$ sur $[a,b]$ et telles que

$$A = \{(x,y) \in \mathbb{R}^2 \ : \ a \leq x \leq b,\ f_1(x) \leq y \leq f_2(x)\}.$$

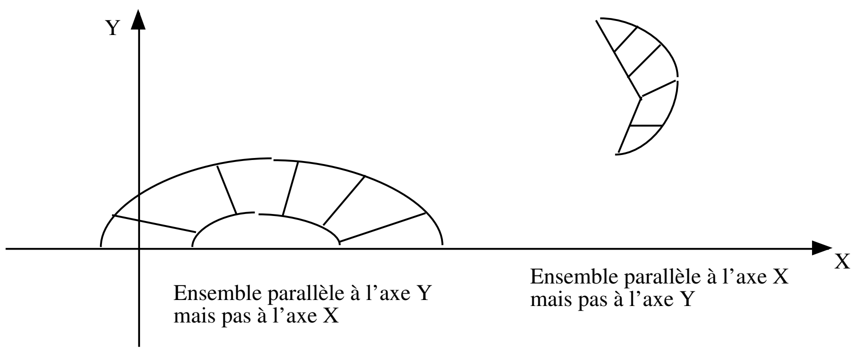
b) On dit que A est parallèle à l'axe X s'il existe deux fonctions g_1, g_2 continues sur un intervalle $[c,d]$ de $\mathbb{R}$ telles que $g_1 \leq g_2$ sur $[c,d]$ et telles que

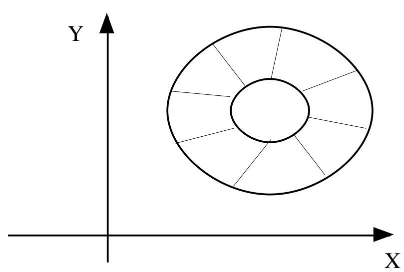
$$A = \{(x,y) \in \mathbb{R}^2 \ : \ g_1(y) \leq x \leq g_2(y),\ c \leq y \leq d\}.$$

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 fermé borné parallèle à l'axe X ou à l'axe Y. Alors f est intégrable sur A et on a*

$$\iint_A f(x,y)\ dxdy \ = \ \int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x,y)dy \right)\ dx$$

si A est parallèle à l'axe Y et

$$\iint_A f(x,y)\ dxdy \ = \ \int_c^d \left(\int_{g_1(y)}^{g_2(y)} f(x,y)dx \right)\ dy$$

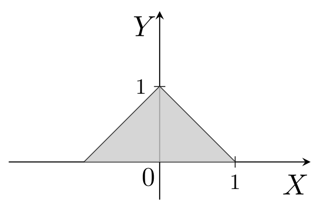
si A est parallèle à l'axe X.

Remarquons que si A est à la fois parallèle à l'axe Y et à l'axe X, on a

$$\begin{aligned}\iint_A f(x,y)\,dxdy &= \int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x,y)dy\right)\,dx\\ &= \int_c^d \left(\int_{g_1(y)}^{g_2(y)} f(x,y)dx\right)\,dy.\end{aligned}$$

On dit que l'on peut *permuter l'ordre d'intégration sans changer la valeur de l'intégrale.*

Par exemple, intégrons la fonction $(x,y)\mapsto f(x,y)=y$, continue sur $\mathbb{R}^2$, sur l'ensemble suivant.



Cet ensemble est fermé et borné, à la fois parallèle à l'axe X et Y. Posons

$$f_1(x)=0\ (x\in[-1,1]),\ f_2(x)=\begin{cases} x+1 & \text{si } x\in[-1,0]\\ 1-x & \text{si } x\in[0,1]\end{cases}$$

et

$$g_1(y)=y-1\ (y\in[0,1]),\ g_2(y)=1-y\ (y\in[0,1]).$$

Les fonctions f_1,f_2 sont continues sur $[-1,1]$; les fonctions g_1,g_2 sont continues sur $[0,1]$. On a

$$A=\{(x,y)\in\mathbb{R}^2 : -1\le x\le 1,\ f_1(x)\le y\le f_2(x)\}=\{(x,y)\in\mathbb{R}^2 : 0\le y\le 1,\ g_1(y)\le x\le g_2(y)\}.$$

Dès lors, comme f est continu sur A, on obtient

$$\begin{aligned}\iint_A f(x,y)\,dxdy &= \int_{-1}^1 \left(\int_{f_1(x)}^{f_2(x)} y\,dy\right)\,dx\\ &= \int_{-1}^0 \left(\int_0^{x+1} y\,dy\right)\,dx\ +\ \int_0^1 \left(\int_0^{1-x} y\,dy\right)\,dx\\ &= \int_{-1}^0 \frac{(x+1)^2}{2}\,dx + \int_0^1 \frac{(1-x)^2}{2}dx\\ &= \frac{1}{2}\left[\frac{(x+1)^3}{3}\right]_{-1}^0\ +\ \frac{1}{2}\left[\frac{(x-1)^3}{3}\right]_0^1\\ &= \frac{1}{2}\times\frac{1}{3}\ +\ \frac{1}{2}\times\frac{1}{3}=\frac{1}{3}.\end{aligned}$$

Calculons cette intégrale dans l'autre ordre. On a

$$\begin{aligned}
\iint_A f(x,y)\ dxdy &= \int_0^1 \left(\int_{g_1(y)}^{g_2(y)} y\ dx\right)\ dy \\
&= \int_0^1 \left(\int_{y-1}^{1-y} y\ dx\right)\ dy \\
&= \int_0^1 y(1-y-(y-1))\ dy \\
&= 2\int_0^1 y(1-y)\ dy \\
&= 2\left[\frac{y^2}{2}-\frac{y^3}{3}\right]_0^1 \\
&= 2\left(\frac{1}{2}-\frac{1}{3}\right) = \frac{1}{3}.
\end{aligned}$$

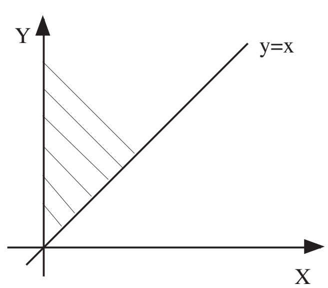
2.4.3 Intégration sur des ensembles non bornés fermés

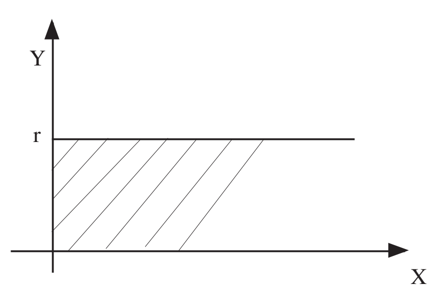
Introduction

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 $g_2(y)=y$.





$$\begin{aligned}
A &= \{(x,y)\in\mathbb{R}^2 : 0\le x,\ x\le y\} \\
&= \{(x,y)\in\mathbb{R}^2 : 0\le x,\ f_1(x)\le y\} \\
&= \{(x,y)\in\mathbb{R}^2 : 0\le y,\ 0\le x\le y\} \\
&= \{(x,y)\in\mathbb{R}^2 : 0\le y,\ g_1(y)\le x\le g_2(y)\}
\end{aligned}$$

$$\begin{aligned}
A &= \{(x,y)\in\mathbb{R}^2 : 0\le x,\ 0\le y\le r\} \\
&= [0,+\infty[\times[0,r]
\end{aligned}$$

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
$$f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2}, \quad (x,y) \in A =]0,1]\times]0,1].$$
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
$$\int_0^1 \frac{-1}{1+y^2} dy = -\operatorname{arctg}(1) = -\frac{\pi}{4}.$$
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(s) 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 $\displaystyle\int_{f_1(x)}^{f_2(x)} |f(x,y)|dy$, $x \in]a,b[$, est intégrable sur $]a,b[$

alors

- pour tout $y \in]c,d[$, la fonction $|f(x,y)|$, $x \in]g_1(y), g_2(y)[$, est intégrable sur l'intevalle $]g_1(y), g_2(y)[$

- la fonction $\displaystyle\int_{g_1(y)}^{g_2(y)} |f(x,y)|dx$, $y \in]c,d[$, est intégrable sur $]c,d[$

et on a

$$\int_a^b \left(\int_{f_1(x)}^{f_2(x)} |f(x,y)|\ dy\right)\ dx = \int_c^d \left(\int_{g_1(y)}^{g_2(y)} |f(x,y)|\ dx\right)\ dy$$

ainsi que *les intégrabilités successives avec f au lieu de son module et*

$$\int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x,y)\ dy\right)\ dx = \int_c^d \left(\int_{g_1(y)}^{g_2(y)} f(x,y)\ dx\right)\ dy.$$

Reprenons l'exemple précédent, à savoir la fonction

$$f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2}, \quad (x,y) \in A =]0,1]\times]0,1].$$

Pour tout $x \in]0,1]$ on a

$$\begin{aligned}
\int_0^1 |f(x,y)|\ dy &= \int_0^x \frac{x^2-y^2}{(x^2+y^2)^2}\ dy\ +\ \int_x^1 \frac{y^2-x^2}{(x^2+y^2)^2}\ dy \\
&= \int_0^x D_y \frac{y}{x^2+y^2}\ dy\ -\ \int_x^1 D_y \frac{y}{x^2+y^2}\ dy \\
&= \frac{1}{2x} - \frac{1}{x^2+1} + \frac{1}{2x} \\
&= \frac{1}{x} - \frac{1}{x^2+1}.
\end{aligned}$$

Comme la fonction $x \mapsto 1/(x^2+1)$ est intégrable en 0^+ et comme la fonction $x \mapsto 1/x$ ne l'est pas, la fonction

$$x \mapsto \int_0^1 |f(x,y)|\ dy$$

ne l'est pas non plus. Les hypothèses du résultat précédent ne sont pas vérifiées et il n'est donc pas étonnant que les calculs après permutation des intégrales ne donnent pas la même valeur.

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 $\displaystyle\int_{f_1(x)}^{f_2(x)} |f(x,y)|dy$, $x \in]a,b[$, est intégrable sur $]a,b[$.

L'intégrale de f sur A est alors

$$\iint_A f(x,y)\ dxdy \ = \ \int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x,y)\ dy\right)\ dx.$$

La définition est analogue (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.

Propriété(s) 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*

$$\iint_A f(x,y)\ dx\ dy = \int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x,y)\ dy\right)\ dx = \int_c^d \left(\int_{g_1(y)}^{g_2(y)} f(x,y)\ dx\right)\ dy.$$

2.4.4 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 :

$$\iint_A f(x,y)\ dxdy \ = \ \iint_{A_1} f(x,y)\ dxdy \ + \ \iint_{A_2} f(x,y)\ dxdy.$$

2.4.5 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 2.4.8. *Soient deux fonctions g_1, g_2 qui sont continûment dérivables sur un ouvert B inclus dans $\mathbb{R}^2$ et établissent une bijection entre A et B (ouvert et inclus aussi dans $\mathbb{R}^2$) telle que la bijection inverse soit aussi formée par deux fonctions continûment dérivables sur A.*

Alors la fonction continue $(x,y) \mapsto f(x,y)$ est intégrable sur A si et seulement si la fonction $(s,t) \mapsto f\left(g_1(s,t), g_2(s,t)\right)\ |D_1g_1(s,t)\, D_2g_2(s,t) - D_1g_2(s,t)\, D_2g_1(s,t)|$ est intégrable sur B et dans ce cas

$$\begin{aligned}&\iint_A f(x,y)\ dx\ dy \\ &= \ \iint_B f\left(g_1(s,t), g_2(s,t)\right)\ |D_1g_1(s,t)\ D_2g_2(s,t)\ - \ D_1g_2(s,t)\ D_2g_1(s,)|\ ds\ dt.\end{aligned}$$

La fonction

$$(s,t) \mapsto D_1g_1(s,t)\ D_2g_2(s,t)\ - \ D_1g_2(s,t)\ D_2g_1(s,t)$$

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 forme

$$x = r\cos\theta,\ y = r\sin\theta$$

où $r > 0$ et $\theta \in [0, 2\pi[$. Si on note $s = r$ et $t = \theta$, on définit (cf le cas général)

$$g_1(r, \theta) = r\cos\theta, \quad g_2(r, \theta) = r\sin\theta.$$

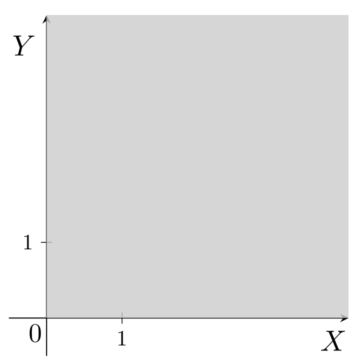
Ces fonctions sont continûment dérivables sur $\mathbb{R}^2$. Elles établissent une bijection entre l'ensemble $]0, +\infty[\times[0, 2\pi[$ et $\mathbb{R}^2 \setminus \{(0, 0)\}$ et s'inversent selon

$$r = \sqrt{x^2 + y^2}, \quad \theta = \begin{cases} \arctan(y/x) & \text{si } x > 0, y \geq 0 \\ \arctan(y/x) + \pi & \text{si } x < 0, y \in \mathbb{R} \\ \arctan(y/x) + 2\pi & \text{si } x > 0, y < 0 \\ \pi/2 & \text{si } x = 0, y > 0 \\ 3\pi/2 & \text{si } x = 0, y < 0 \end{cases}$$

On montre que ces deux fonctions sont aussi continûment dérivables sur $]0, +\infty[\times]0, 2\pi[$.

Par exemple, si

$$A =]0, +\infty[\times]0, +\infty[,$$

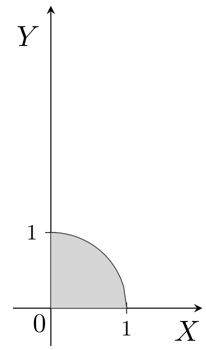


on a

$$B =]0, +\infty[\times]0, \pi/2[.$$

De même si

$$A = \{(x, y) \in \mathbb{R}^2\ :\ x > 0, y > 0, x^2 + y^2 \leq 1\}$$



on a

$$B =]0,1]\times]0,\pi/2[.$$

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

$$\boxed{\iint_A f(x,y)\ dx\ dy\ =\ \iint_B f(r\cos\theta, r\sin\theta)\ r\ dr\ d\theta}$$

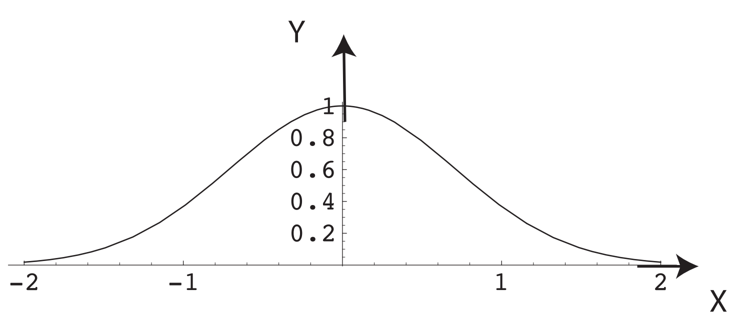
où $A\subset\mathbb{R}^2$ et $B\subset]0,+\infty[\times]0,2\pi[$.

2.4.6 Applications

Calcul de $\boxed{\int_0^{+\infty} e^{-ax^2}dx}$ **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 $\lim_{x\to+\infty} x^2e^{-ax^2} = \lim_{x\to-\infty} x^2e^{-ax^2} = 0$.

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 donnée explicitement par

$$F(x,y) = e^{-a(x^2+y^2)} = f(x)\ f(y).$$

En utilisant le changement de variables polaires, on a

$$\iint_A F(x,y)\ dx\ dy \ = \ I_a^2 = \iint_{A'} e^{-ar^2} r\ dr\ d\theta$$

avec

$$A =]0,+\infty[\times]0,+\infty[\quad A' =]0,+\infty[\times]0,\frac{\pi}{2}[.$$

L'intégrale de droite se calcule aisément :

$$\begin{aligned}
\iint_{A'} e^{-ar^2} r\ dr\ d\theta \ &= \ \left(\int_0^{\pi/2} 1\ d\theta\right)\left(\int_0^{+\infty} re^{-ar^2}\ dr\right) \\
&= \ \frac{\pi}{2}\int_0^{+\infty} \frac{-1}{2a}\ D_r e^{-ar^2}\ dr = \frac{-\pi}{4a}\ \left[e^{-ar^2}\right]_0^{+\infty} \\
&= \ \frac{\pi}{4a}.
\end{aligned}$$

Ainsi, puisque $I_a > 0$, on obtient

$$\boxed{I_a = \frac{1}{2}\sqrt{\frac{\pi}{a}}.}$$

Calcul de $\boxed{\int_0^{\rightarrow+\infty} \frac{\sin(x)}{x} dx}$

Cette intégrale ressemble aux intégrales *de Fresnel* qui interviennent dans la diffraction [4]. 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 ($x \mapsto \sin(x)/x$ 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

$$\lim_{t\to+\infty} \int_0^t \frac{\sin(x)}{x} dx$$

existe, est finie et calculer sa valeur.

Remarquons d'abord que, pour tout $t > 0$, la fonction $x \mapsto \sin(x)/x$ est continue sur $]0,t]$ et admet une limite finie en 0. Elle est donc intégrable sur l'intervalle $]0,t]$. Notons

$$I_t = \int_0^t \frac{\sin(x)}{x}\ dx.$$

4. Intégrales de Fresnel : $(1/\sqrt{2\pi}) \int_0^{x^2} \cos(t)/\sqrt{t}\ dt = \sqrt{2/\pi} \int_0^x \cos(t^2)\ dt$ et $(1/\sqrt{2\pi}) \int_0^{x^2} \sin(t)/\sqrt{t}\ dt = \sqrt{2/\pi} \int_0^x \sin(t^2)\ dt$.

Cela étant, pour tout $x > 0$, on a

$$\frac{1}{x} = \int_0^{+\infty} e^{-sx}\ ds.$$

On a donc, pour $t > 0$:

$$I_t = \int_0^t \frac{\sin(x)}{x}\ dx\ =\ \int_0^t \sin(x) \left(\int_0^{+\infty} e^{-xs} ds \right)\ dx. \tag{2.1}$$

Pour $t > 0$, on considère alors la fonction

$$|\sin(x)|\ e^{-sx}, \quad s \in]0, +\infty[,\ x \in]0, t].$$

On a

$$\int_0^{+\infty} |\sin(x)|\ e^{-sx}\ ds = \frac{|\sin(x)|}{x}$$

et $x \mapsto |\sin(x)|/x$ 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

$$\begin{aligned}
\int_0^t \sin(x)\, e^{-xs}\ dx \quad &= \quad -\int_0^t D_x \cos(x)\, e^{-xs}\ dx \\
&= \quad -\left[\cos(x)\, e^{-xs}\right]_{x=0}^{x=t}\ -\ s \int_0^t \cos(x)\, e^{-xs}\ dx \\
&= \quad -\cos(t)\, e^{-ts} + 1\ -\ s \int_0^t D_x \sin(x)\, e^{-xs}\ dx \\
&= \quad -\cos(t)\ e^{-ts} + 1\ -\ s \left(\left[\sin(x)\, e^{-xs}\right]_{x=0}^{x=t} + s \int_0^t \sin(x)\, e^{-xs}\ dx \right) \\
&= \quad -\cos(t)\, e^{-ts} + 1\ -\ s \sin(t)\ \ e^{-ts} - s^2 \int_0^t \sin(x)\, e^{-xs}\ dx
\end{aligned}$$

donc

$$\int_0^t \sin(x)\, e^{-xs}\ dx = \frac{1}{1+s^2} \left(1 - \cos(t)\, e^{-st} - s \sin(t)\, e^{-st}\right).$$

Dès lors,

$$\begin{aligned}
I_t \quad &= \quad \int_0^{+\infty} \frac{1}{1+s^2} (1 - \cos(t)\, e^{-st} - s \sin(t)\, e^{-st})\ ds \\
&= \quad \int_0^{+\infty} \frac{1}{1+s^2}\ ds\ -\ \cos(t) \int_0^{+\infty} \frac{e^{-st}}{1+s^2}\ ds\ -\ \sin(t) \int_0^{+\infty} \frac{se^{-st}}{1+s^2}\ ds.
\end{aligned}$$

On a

$$\int_0^{+\infty} \frac{1}{1+s^2} = [\arctan(s)]_0^{+\infty} = \frac{\pi}{2};$$

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

$$\left|\sin(t)\int_0^{+\infty}\frac{se^{-st}}{1+s^2}\,ds\right| \leq \int_0^{+\infty} e^{-st}\,ds \;=\; \frac{1}{t}$$

on obtient que

$$\lim_{t\to+\infty}\left(\cos(t)\int_0^{+\infty}\frac{e^{-st}}{1+s^2}\,ds\right) \;=\; 0, \quad \lim_{t\to+\infty}\left(\sin(t)\int_0^{+\infty}\frac{se^{-st}}{1+s^2}\,ds\right) \;=\; 0$$

donc finalement

$$\boxed{\lim_{t\to+\infty} I_t = \int_0^{\to+\infty}\frac{\sin(x)}{x}\,dx \;=\; \frac{\pi}{2}.}$$

2.4.7 Intégrales triples

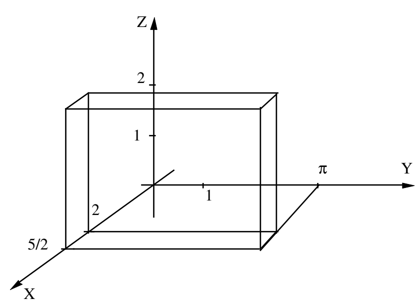
Cas général

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 $x = 2, x = 5/2, y = 0, y = \pi$ et soit C le parallélépipède situé entre les plans d'équation $z = 0, z = 2$ sur R. Calculer

$$\iiint_C zx\sin(xy)\;dxdydz.$$

Cela a bien un sens car il s'agit d'intégrer une fonction continue sur un ensemble fermé borné.



L'ensemble d'intégration C s'écrit

$$C = \Big\{(x,y,z)\in\mathbb{R}^3 :\; 2\leq x\leq 5/2,\; 0\leq y\leq\pi,\; 0\leq z\leq 2\Big\} = [2,5/2]\times[0,\pi]\times[0,2].$$

On a donc

$$\begin{aligned}
\iiint_C zx\sin(xy)\;dxdydz &= \int_2^{5/2}dx\int_0^{\pi}dy\int_0^{2} zx\sin(xy)\;dz \\
&= \int_2^{5/2}dx\int_0^{\pi}dy\left[\frac{z^2}{2}x\sin(xy)\right]_{z=0}^{z=2} \\
&= \int_2^{5/2}dx\int_0^{\pi}2x\sin(xy)dy = \int_2^{5/2}dx\int_0^{\pi}2D_y(-\cos(xy))dy \\
&= 2\int_2^{5/2}dx\,[-\cos(xy)]_{y=0}^{y=\pi} = 2\int_2^{5/2}dx(1-\cos(x\pi)) \\
&= 2\left[x-\frac{1}{\pi}\sin(\pi x)\right]_2^{5/2} = 2\;\left(\frac{5}{2}-\frac{1}{\pi}\sin(\frac{5\pi}{2})-2+\frac{1}{\pi}\sin(2\pi)\right) \\
&= 1-\frac{2}{\pi}.
\end{aligned}$$

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

$$V = \iint_R f(x,y)\ dx\ dy.$$

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

$$C = \{(x,y,z) \in \mathbb{R}^3\ :\ 0 \leq z \leq f(x,y)\}$$

l'intégrale donnant le volume s'écrit

$$V = \iint_R f(x,y)\ dxdy\ =\ \iint_R \left(\int_0^{f(x,y)} 1dz\right)\ dxdy\ =\ \iiint_C 1\ dxdydz.$$

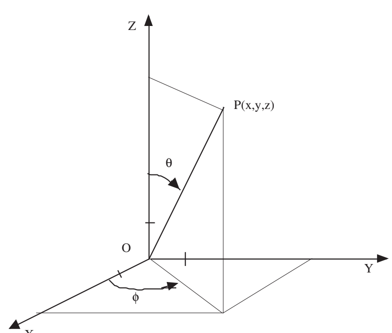
Définition 2.4.9. *Soit C un corps borné fermé de l'espace. Le volume de C est*

$$V = \iiint_C 1\ dxdydz.$$

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

$$\left\{\begin{array}{lcl} x & = & r\cos\varphi\sin\theta \\ y & = & r\sin\varphi\sin\theta \\ z & = & r\cos\theta. \end{array}\right.$$

Ici encore, on définit un jacobien. La formule de changement de variables dans les intégrales s'écrit alors

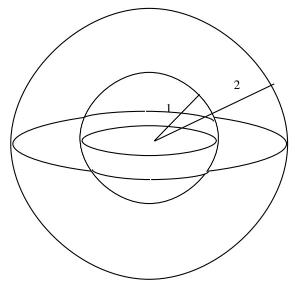
$$\boxed{\iiint_A f(x,y,z)\ dxdydz \ = \ \iiint_{A'} f(r\cos\varphi\sin\theta, r\sin\varphi\sin\theta, r\cos\theta)\ r^2\sin\theta\ drd\varphi d\theta}$$

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

$$\iiint_C z^2\ dxdydz$$

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

$$\begin{aligned} &\iiint_A z^2\ dxdydz \\ &= \iiint_{A'} (r\cos\theta)^2\ r^2\sin\theta\ drd\varphi d\theta \\ &= \int_0^\pi \cos^2\theta\sin\theta\ d\theta \quad \int_1^2 r^4\ dr \quad \int_0^{2\pi} 1\ d\varphi \\ &= \left[-\frac{1}{3}\cos^3\theta\right]_0^\pi \quad \left[\frac{r^5}{5}\right]_1^2 \quad 2\pi \\ &= \frac{124}{15}\pi. \end{aligned}$$

Chapitre 3

Approximations polynomiales et séries de puissances

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, par définition on a

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} \;=\; Df(x_0)$$

donc

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)-(x-x_0)Df(x_0)}{x-x_0}=0.$$

Ainsi le le polynôme

$$x\mapsto P(x-x_0)=f(x_0)+(x-x_0)Df(x_0)$$

vérifie

$$\lim_{x\to x_0}\frac{f(x)-P(x-x_0)}{x-x_0}=0.$$

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

$$\lim_{x\to x_0}\frac{R(x)}{x-x_0}=0.$$

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 polynomiale de f à l'ordre n en x_0 lorsque[1]

$$\lim_{x\to x_0, x\neq x_0} \frac{f(x)-P(x-x_0)}{(x-x_0)^n} = 0. \qquad (*)$$

Le reste de l'approximation est la fonction

$$x \mapsto R(x) = f(x) - P(x-x_0).$$

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(s) 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

$$x \mapsto f(x_0) + (x-x_0)Df(x_0).$$

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.□

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

$$x \mapsto f(x_0) + (x-x_0)Df(x_0).$$

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 $x - x_0$.

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]

$$\lim_{x\to x_0} \frac{f(x) - \displaystyle\sum_{j=0}^{n} \frac{D^j f(x_0)}{j!}(x-x_0)^j}{(x-x_0)^n} = 0.$$

L'approximation à l'ordre n de f en x_0 est donc

$$x \mapsto P(x-x_0) \;=\; f(x_0) + (x-x_0)Df(x_0) + \frac{(x-x_0)^2}{2!}D^2 f(x_0) + \ldots + \frac{(x-x_0)^n}{n!}D^n f(x_0)$$

ou encore

$$\boxed{x \mapsto P(x-x_0) \;=\; \sum_{j=0}^{n} \frac{D^j f(x_0)}{j!}(x-x_0)^j}$$

Preuve. Cas où f est à valeurs réelles (application du théorème de l'Hospital).
Si $n=1$, on a directement

$$\lim_{x\to x_0} \frac{f(x) - f(x_0) - (x-x_0)Df(x_0)}{x-x_0} = \lim_{x\to x_0}\left(\frac{f(x)-f(x_0)}{x-x_0}\right) - Df(x_0) = 0.$$

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

$$\lim_{x\to x_0}(f(x)-g(x)) = 0 \quad \text{et} \quad \lim_{x\to x_0}(x-x_0)^2 = 0.$$

Comme

$$\begin{aligned}
\lim_{x\to x_0} \frac{Df(x)-Dg(x)}{D(x-x_0)^2} &= \lim_{x\to x_0} \frac{Df(x) - Df(x_0) - (x-x_0)D^2 f(x_0)}{2(x-x_0)} \\
&= \frac{1}{2}\lim_{x\to x_0} \frac{Df(x)-Df(x_0)}{x-x_0} - \frac{1}{2}D^2 f(x_0) \\
&= 0
\end{aligned}$$

le théorème de l'Hospital donne

$$\lim_{x\to x_0} \frac{f(x)-g(x)}{(x-x_0)^2} = 0$$

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.

Si f est à valeurs complexes, on procède avec ses parties réelle et imaginaire pour arriver au résultat. □

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 $x_0 \in]a,b[$

Cas des approximations d'ordre 1 (approximations dites linéaires)

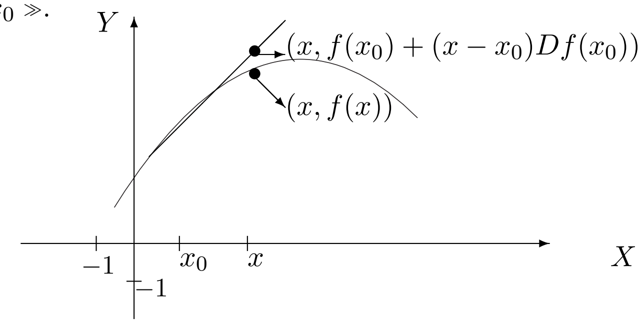
Pour une fonction f vérifiant les hypothèses du résultat précédent avec $n = 1$, rappelons que l'on dit que

$$x \mapsto P(x - x_0) = f(x_0) + (x - x_0)Df(x_0)$$

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 ».



Graphique de f et de son approximation linéaire en x_0

Cas des approximations d'ordre 2 (approximations dites quadratiques)

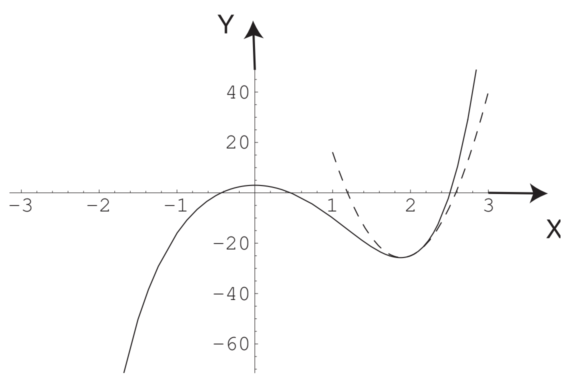
Pour une fonction f vérifiant les hypothèses du résultat précédent avec $n = 2$, on dit que

$$x \mapsto P(x - x_0) = f(x_0) + (x - x_0)Df(x_0) + \frac{(x - x_0)^2}{2} D^2 f(x_0)$$

est l'*approximation quadratique* de f en x_0. Le graphique de

$$g(x) = f(x_0) + (x - x_0)Df(x_0) + \frac{(x - x_0)^2}{2} D^2 f(x_0),$$

lorsque $D^2 f(x_0) \neq 0$, 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 $y = f(x_0) + (x - x_0)Df(x_0) + (x - x_0)^2 D^2 f(x_0)/2$ 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 $x_0 = 2$ (repère orthogonal non normé)

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 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

$$D\sin(x) = \cos(x),\ D^2\sin(x) = -\sin(x), \quad D^3\sin(x) = -\cos(x)$$

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

$$\boxed{\text{les approximations à l'ordre } 0,1,2,3 \text{ de sin}}$$

sont successivement

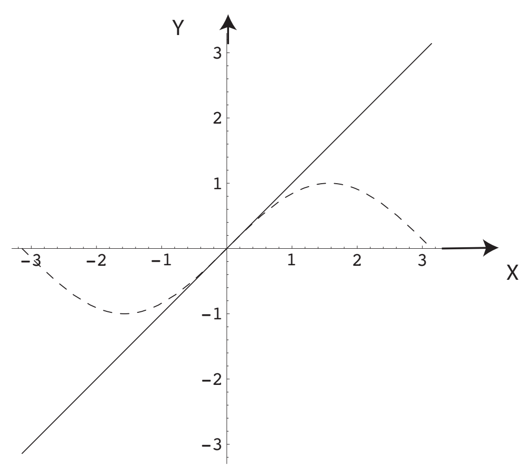
$$\boxed{P_0(x) = 0\ (x \in \mathbb{R}), \quad P_1(x) = x\ (x \in \mathbb{R}), \quad P_2(x) = P_1(x) = x\ (x \in \mathbb{R}), \quad P_3(x) = x - x^3/6\ (x \in \mathbb{R})}$$

et que

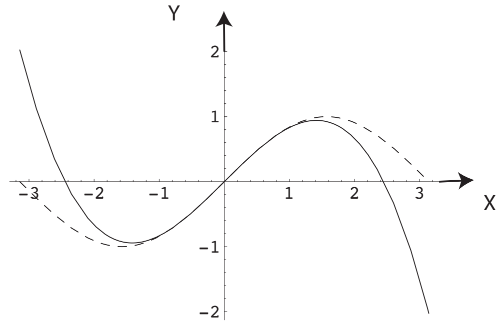
$$\boxed{\text{les approximations à l'ordre } 0,1,2,3 \text{ de cos}}$$

sont successivement

$$\boxed{P_0(x) = 1\ (x \in \mathbb{R}), \quad P_1(x) = 1\ (x \in \mathbb{R}), \quad P_2(x) = 1 - x^2/2\ (x \in \mathbb{R}), \quad P_3(x) = 1 - x^2/2 = P_2(x)\ (x \in \mathbb{R}).}$$



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

$$D^{2n} \sin(x) = (-1)^n \sin x, \qquad D^{2n+1} \sin(x) = (-1)^n \cos(x)$$

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 explicitement

$$P_{2n+1}(x) = \sum_{k=0}^{n} \frac{x^{2k+1}}{(2k+1)!} D^{2k+1} \sin(0) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$$

et l'approximation à l'ordre $2n$ de sin en 0 est un polynôme de degré $2n-1$, à savoir explicitement

$$P_{2n}(x) = \sum_{k=0}^{n-1} \frac{x^{2k+1}}{(2k+1)!} D^{2k+1} \sin(0) = \sum_{k=0}^{n-1} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = P_{2n-1}(x).$$

On procède de même avec la fonction cos. On a

$$D^{2n} \cos x = (-1)^n \cos x, \qquad D^{2n+1} \cos x = (-1)^{n+1} \sin x$$

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 explicitement

$$P_{2n}(x) = \sum_{k=0}^{n} \frac{x^{2k}}{(2k)!} D^{2k} \cos(0) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!}$$

et l'approximation à l'ordre $2n+1$ de cos en 0 est un polynôme de degré $2n$, à savoir explicitement

$$P_{2n+1}(x) = \sum_{k=0}^{n} \frac{x^{2k}}{(2k)!} D^{2k} \cos(0) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} = P_{2n}(x).$$

3.1.5 Estimation du reste

Le résultat qui suit est une généralisation du théorème des accroissements finis ; il correspond en effet à ce théorème lorsque $n = 1$. C'est un **rappel** du cours de première année du premier quadrimestre MATH2007.

Théorème 3.1.4 (Développement limité de Taylor)**.** *Soient n un naturel strictement positif et f une fonction à valeurs réelles 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).

Sous les hypothèses du résultat précédent (développement limité de Taylor), on obtient donc, si $x \mapsto P_n(x - x_0)$ désigne l'approximation à l'ordre n de f en x_0 :

$$f(x) = P_n(x - x_0) \ + \ \frac{(x - x_0)^n}{n!} \left(D^n f(u_0) - D^n f(x_0)\right).$$

Le reste $x \mapsto R_n(x) = f(x) - P(x - x_0)$ de l'approximation à l'ordre n s'exprime donc sous la forme

$$\boxed{x \mapsto R_n(x) = \frac{(x - x_0)^n}{n!} \left(D^n f(u_0) - D^n f(x_0)\right)}$$

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

$$f(x) = f(x_0) + (x - x_0)Df(x_0) + \ldots + \frac{(x - x_0)^n}{n!} D^n f(x_0) + \frac{(x - x_0)^{n+1}}{(n + 1)!} D^{n+1} f(u_1)$$

donc le reste de l'approximation à l'ordre n est

$$\boxed{x \mapsto R_n(x) = \frac{(x - x_0)^{n+1}}{(n + 1)!} D^{n+1} f(u_1)}$$

ce qui donne aussi lieu a des estimations.

Par exemple, pour la fonction sin,
- le reste de l'approximation à l'ordre $2n$ en 0 s'écrit explicitement

$$R_{2n}(x) = \frac{x^{2n+1}}{(2n + 1)!} D^{2n+1} \sin(u_0) = \frac{x^{2n+1}}{(2n + 1)!} (-1)^n \cos(u_0)$$

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 explicitement

$$R_{2n+1}(x) = \frac{x^{2n+2}}{(2n + 2)!} D^{2n+2} \sin(u_0) = \frac{x^{2n+2}}{(2n + 2)!} (-1)^{n+1} \sin(u_0);$$

comme les valeurs de la fonction sin sont petites 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

$$|R_{2n+1}(x)| \leq \frac{|x|^{2n+3}}{(2n + 3)!}.$$

3.1.6 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 3.1.5. *Soit $P: \; x \mapsto P(x) = a_0 + a_1 x + \ldots + a_n x^n$ un polynôme de degré*[3] *n. Pour tout naturel N plus grand ou égal à n, l'approximation polynomiale à l'ordre N en x_0 est le polynôme P lui-même et on a*

$$P(x) = P(x_0) + (x - x_0)DP(x_0) + \ldots + \frac{(x-x_0)^n}{n!} D^n P(x_0), \; x \in \mathbb{R}.$$

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

$$P(x) = P(x_0) + (x - x_0)DP(x_0) + \ldots + \frac{(x-x_0)^N}{N!} D^N P(x_0) + \frac{(x-x_0)^{N+1}}{(N+1)!} D^{N+1} P(u) \qquad (*)$$

Par ailleurs, comme $D^k P$ est la fonction nulle lorsque k est strictement supérieur à n, on a toujours

$$\left(D^{N+1} P\right)(u) = 0$$

et lorsque $N > n$, les derniers termes[4] de cette expression (*) sont nuls également.

On obtient donc

$$P(x) = P(x_0) + (x - x_0)DP(x_0) + \ldots + \frac{(x-x_0)^n}{n!} D^n P(x_0), \; x \in \mathbb{R}.$$

□

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 3.1.6. *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*

$$D^k P(x_0) = 0, \; k = 0, \ldots, \alpha - 1, \quad D^\alpha P(x_0) \neq 0.$$

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 $k \in \{0, \ldots, \alpha - 1\}$ et en considérant cette relation en x_0, on obtient

$$D^k P(x_0) = 0$$

3. on a donc $a_n \neq 0$.
4. le nombre de ces termes dépend bien sûr de N

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 3.1.5), on a

$$P(x) = \frac{(x - x_0)^\alpha}{\alpha!} D^\alpha P(x_0) + \ldots + \frac{(x - x_0)^n}{n!} D^n P(x_0) = (x - x_0)^\alpha Q(x)$$

avec

$$Q(x_0) = \frac{D^\alpha P(x_0)}{\alpha!} \neq 0.$$

□

3.2 Développements illimités-Séries de puissances

3.2.1 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

$$\lim_{N \to +\infty} R_N(x_0, x) = 0 \qquad (*)$$

on a

$$f(x) = \lim_{N \to +\infty} \left(\sum_{n=0}^{N} \frac{(x - x_0)^n}{n!} D^n f(x_0) \; + \; R_N(x_0, x) \right) = \lim_{N \to +\infty} \sum_{n=0}^{N} \frac{(x - x_0)^n}{n!} D^n f(x_0)$$

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 analytiques [5].

5. La fonction

$$f(x) = \begin{cases} \exp(-1/x) = e^{-1/x} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

3.3 Séries de puissances

Soit $a_m (m \in \mathbb{N})$ une suite de complexes et soit x_0 un réel (ou même un complexe). La suite de sommes partielles

$$\sum_{m=0}^{M} a_m (x - x_0)^m, \quad M \in \mathbb{N}$$

est appelée *série de puissances* de $x - x_0$. 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 suite converge vers une limite finie. En tout x où la suite converge vers une limite finie, on appelle celle-ci la *somme* de la série et on la note

$$\sum_{m=0}^{+\infty} a_m (x - x_0)^m.$$

Par abus de langage, on dit aussi *série de puissances* pour l'expression $\sum_{m=0}^{+\infty} a_m (x - x_0)^m$.

Pour l'étude des séries de puissances, on a les résultats suivants, que nous ne démontrons pas dans le cadre de ce cours.

Théorème 3.3.1. *On considère la série de puissances* $\sum_{m=0}^{M} a_m (x - x_0)^m$ $(M \in \mathbb{N})$ *avec* x_0 *et* x *réels.*

• *Si*

$$\lim_{M \to +\infty} \sum_{m=0}^{M} a_m (x - x_0)^m \ \in \mathbb{C}$$

alors

$$\lim_{m \to +\infty} a_m (x - x_0)^m \ = \ 0.$$

• *Si*

$$\lim_{m \to +\infty} \frac{|a_{m+1}|}{|a_m|} = r \in \mathbb{R}_0$$

alors la série de puissances converge en tout x *de l'intervalle* $]x_0 - 1/r, x_0 + 1/r[$.

• *Si*

$$\lim_{m \to +\infty} \frac{|a_{m+1}|}{|a_m|} = 0$$

alors la série de puissances converge en tout réel x*.*

• *Si*

$$\lim_{m \to +\infty} \frac{|a_{m+1}|}{|a_m|} = +\infty$$

alors la série de puissances ne converge qu'en x_0*.*

Théorème 3.3.2. *On considère la série de puissances*

$$\sum_{m=0}^{+\infty} a_m (x - x_0)^m.$$

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.

Supposons que la série converge pour tout $x \in]x_0 - R, x_0 + R[$ *où* R *est un réel strictement positif. Alors la fonction*

$$x \mapsto \; S(x) = \sum_{m=0}^{+\infty} a_m (x - x_0)^m$$

est indéfiniment continûment dérivable dans l'intervalle $]x_0 - R, x_0 + R[$ *et y est dérivable terme à terme c'est-à-dire*

$$DS(x) = D \sum_{m=0}^{+\infty} a_m (x - x_0)^m = \sum_{m=0}^{+\infty} D(a_m (x - x_0)^m) = \sum_{m=1}^{+\infty} a_m \, m \, (x - x_0)^{m-1}$$

et ainsi de suite pour les dérivées d'ordre supérieur.

Il est important de bien comprendre ce que cela veut dire : il ne s'agit en aucune manière de dériver une somme finie puisque des limites interviennent. Ce résultat exprime en fait que l'on peut permuter deux limites, à savoir la limite d'une suite de sommes partielles et la limite qui se trouve dans la notion de dérivée.

3.4 Fonction exponentielle (définie par une série de puissances)

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.

3.4.1 Définition

Définition 3.4.1. *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 de puissances

$$\sum_{m=0}^{+\infty} \frac{x^m}{m!}$$

est une suite de sommes partielles qui converge vers une limite finie quel que soit le réel x. Appliquons les résultats énoncés précédemment, avec $x_0 = 0$ et $a_m = 1/m!$. On a

$$\lim_{m \to +\infty} \frac{|a_{m+1}|}{|a_m|} = \lim_{m \to +\infty} \frac{m!}{(m+1)!} = \lim_{m \to +\infty} \frac{1}{m+1} = 0.$$

On a donc la convergence annoncée.

3.4.2 Propriétés fondamentales

Nous sommes à présent en mesure de démontrer les propriétés fondamentales de la fonction exponentielle, annoncées dans le cours MATH2007.

Théorème 3.4.2. *1. La fonction* exp *est indéfiniment continûment dérivable sur* $\mathbb{R}$ *et*[6] *on a*

$$D_x^k \exp(x) = \exp(x),\ x \in \mathbb{R},\ k \in \mathbb{N}; \qquad \exp(0) = 1.$$

2. On a l'estimation

$$e = \exp(1) = 2.71828182\ldots$$

3. Pour tous $x, y \in \mathbb{R}$*, on a*

$$\exp(x + y) = \exp(x) \times \exp(y).$$

4. Pour tout $x \in \mathbb{R}$*, on a*

$$\exp(x) > 0.$$

5. On a $\lim_{x\to+\infty} \exp(x) = +\infty$ *et* $\lim_{x\to-\infty} \exp(x) = 0$*. Plus généralement, pour tout* $p \in \mathbb{N}$*, on a*

$$\lim_{x\to+\infty} \frac{\exp(x)}{x^p} = +\infty, \ \text{et} \ \lim_{x\to-\infty} x^p \exp(x) = 0$$

(« 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 3.3.2. On a donc

$$D \exp(x) = \sum_{m=1}^{+\infty} \frac{m x^{m-1}}{m!} = \sum_{m=0}^{+\infty} \frac{x^m}{m!}$$

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) Admis.

3) 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 (constante par rapport à x mais qui dépend de y puisque f en dépend) telle que

$$f(x) = \exp(x + y)\ \exp(-x) = r, \quad \forall x \in \mathbb{R}.$$

6. 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.

En prenant cette égalité pour $x = 0$, on obtient $r = \exp(y)$. En conclusion, on a obtenu

$$\exp(x+y)\ \exp(-x) = \exp(y) \quad \forall x, y \in \mathbb{R}. \qquad (*)$$

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

$$\exp(x+y) = \exp(x)\ \exp(y) \quad \forall x, y \in \mathbb{R}.$$

4) Pour tout $x \in \mathbb{R}$, on a

$$\exp(x) = \exp\left(\frac{x}{2} + \frac{x}{2}\right) = \left(\exp\left(\frac{x}{2}\right)\right)^2,$$

ce qui montre que $\exp(x) \geq 0$. Par ailleurs, on a $\exp(x)\ \exp(-x) = \exp(x-x) = \exp(0) = 1$, quel que soit x ; dès lors $\exp(x) \neq 0$ quel que soit x.

5) 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

$$\lim_{x \to +\infty} \frac{\exp(x)}{x^p} = +\infty.$$

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

$$\lim_{x \to -\infty} \frac{\exp(-x)}{(-x)^p} = \lim_{y \to +\infty} \frac{\exp(y)}{y^p} = +\infty$$

on obtient

$$\lim_{x \to -\infty} x^p \exp(x) = 0.$$

□

Remarques 1) Le nombre réel $\exp(1)$ est noté e. On démontre que c'est un nombre irrationnel.

2) Comme la fonction exponentielle vérifie $\exp(x+y) = \exp(x)\exp(y)$, $(x, y \in \mathbb{R})$ on a, pour tout naturel m

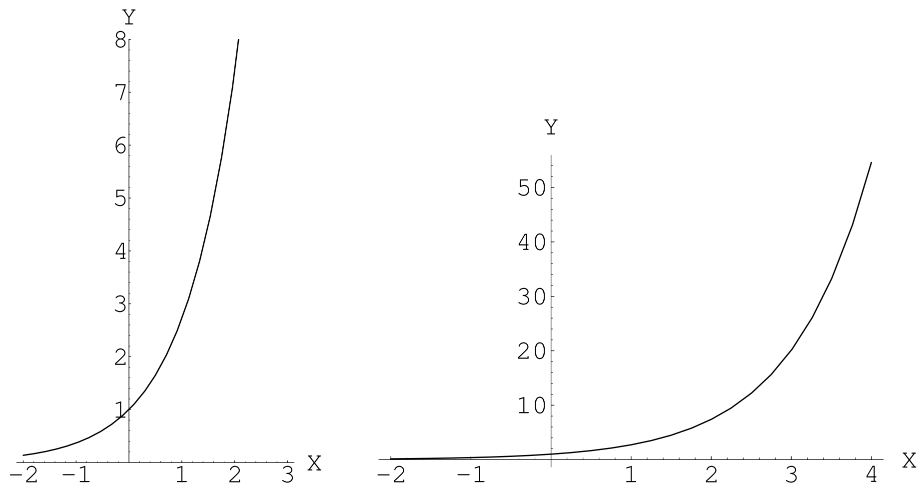
$$\exp(m) = \exp(1 + \ \ldots + 1) = \underbrace{\exp(1)\ \exp(1)\ \ldots \exp(1)}_{m \text{ facteurs}} = \underbrace{e\ e \ldots e}_{m \text{ facteurs}} = e^m;$$

la fonction exponentielle est donc tout naturellement également notée

$$\exp(x) = e^x, \quad x \in \mathbb{R}.$$

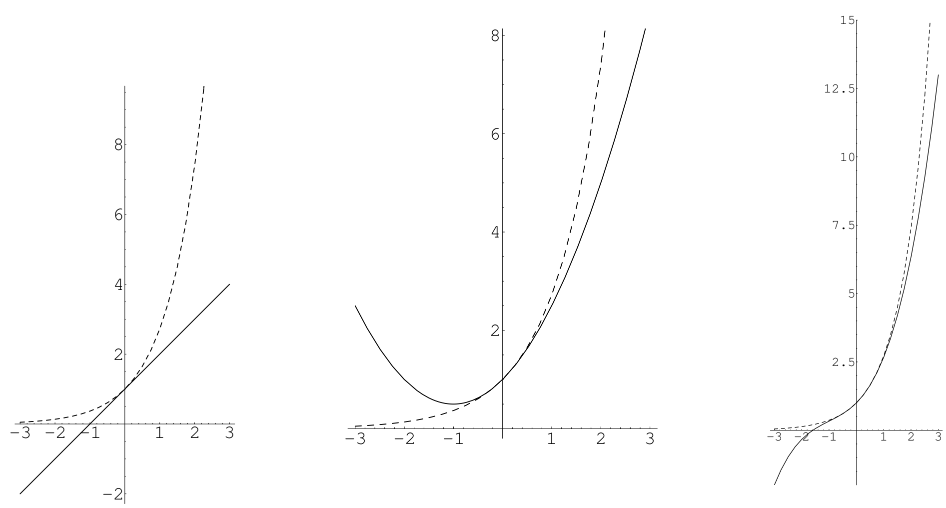
2) A partir des propriétés 1), 5) et 6), on obtient bien sûr que la fonction exp est une fonction bijective de $\mathbb{R}$ sur $]0, +\infty[$.

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

$$P_1(x) = 1 + x \ (x \in \mathbb{R}), \ P_2(x) = 1 + x + \frac{x^2}{2} \ (x \in \mathbb{R}), \ P_3(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} \ (x \in \mathbb{R}).$$



3.5 Exponentielle complexe

Terminons cette partie consacrée à la fonction exponentielle par **une introduction à l'exponentielle complexe (cf aussi le cours math2007 avant les équations différentielles)**, 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, c'est-à-dire que si on remplace $z^m/m!$ par $|z^m/m!| = |z|^m/m!$, on a encore la convergence. 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(s) 3.5.1. *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.□

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 3.5.2. *On définit*

$$\cos(x) = \Re(e^{ix}), \quad \sin(x) = \Im(e^{ix}), \quad x \in \mathbb{R}$$

On en <u>déduit</u> 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 $\Re z = (z+\bar{z})/2$ et $\Im z = (z-\bar{z})/(2i)$ pour tout complexe z, on <u>déduit</u> 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(s) 3.5.3. *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

$$(\cos(x) + i\sin(x))^m = \left(e^{ix}\right)^m = e^{imx} = \cos(mx) + i\sin(mx).$$

4) Résultat admis.□

3.6 Développement de fonctions en séries de puissances

A titre d'exemples de développements de fonctions en séries de puissances, recherchons les développements de quelques fonctions élémentaires.

Rappelons tout d'abord que nous avons démontré les résultats suivants (cf cours MATH2007) :

$$\lim_{M\to+\infty} \sum_{m=1}^{M} \frac{1}{m} = +\infty, \qquad \lim_{M\to+\infty} \sum_{m=1}^{M} \frac{(-1)^{m-1}}{m} = \ln(2).$$

Propriété(s) 3.6.1. *On a*

$$\begin{aligned}
\cos(x) &= \sum_{m=0}^{+\infty} (-1)^m \frac{x^{2m}}{(2m)!}, \quad x \in \mathbb{R} \\
\sin(x) &= \sum_{m=0}^{+\infty} (-1)^m \frac{x^{2m+1}}{(2m+1)!}, \quad x \in \mathbb{R} \\
\frac{1}{1-x} &= \sum_{m=0}^{+\infty} x^m, \quad x \in]-1,1[\\
\frac{1}{1+x} &= \sum_{m=0}^{+\infty} (-1)^m x^m, \quad x \in]-1,1[\\
\ln(1+x) &= \sum_{m=1}^{+\infty} (-1)^{m-1} \frac{x^m}{m}, \quad x \in]-1,1] \\
\ln(1-x) &= -\sum_{m=1}^{+\infty} \frac{x^m}{m}, \quad x \in [-1,1[
\end{aligned}$$

On a donc

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \dots \quad (x \in \mathbb{R}), \qquad \sin(x) = x - \frac{x^3}{6} + \frac{x^5}{120} - \dots \quad (x \in \mathbb{R}),$$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots \quad (x \in]-1,1[), \qquad \frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots \quad (x \in]-1,1[),$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \quad (x \in]-1,1]), \qquad -\ln(1-x) = x + \frac{x^2}{2} + \frac{x^3}{3} + \dots \quad (x \in [-1,1[),$$

avec l'interprétation des pointillés comme étant un passage à la limite, vu la signification de $\sum_{m=0}^{+\infty}$.

Preuve. • *Fonctions cosinus et sinus.*

Quel que soit $x \in \mathbb{R}$, par définition on a

$$\cos(x) = \Re(e^{ix}), \quad \sin(x) = \Im(e^{ix})$$

et

$$e^{ix} = \sum_{m=0}^{+\infty} \frac{(ix)^m}{m!}.$$

Pour trouver les développements des fonctions cosinus et sinus, il suffit donc d'exprimer le second membre de l'égalité de la définition de l'exponentielle comme étant $a + ib$ avec a, b réels.

Cela étant, comme $i^2 = -1$, quel que soit le naturel m on a

$$i^{2m} = (-1)^m \quad \text{et} \quad i^{2m+1} = (-1)^m \, i.$$

Il s'ensuit que l'on obtient directement

$$\begin{aligned} e^{ix} = \sum_{m=0}^{+\infty} \frac{(ix)^m}{m!} &= \sum_{m=0}^{+\infty} \frac{(ix)^{2m}}{(2m)!} + \sum_{m=0}^{+\infty} \frac{(ix)^{2m+1}}{(2m+1)!} \\ &= \sum_{m=0}^{+\infty} (-1)^m \frac{x^{2m}}{(2m)!} + i \sum_{m=0}^{+\infty} (-1)^m \frac{x^{2m+1}}{(2m+1)!}, \end{aligned}$$

donc

$$\cos(x) = \sum_{m=0}^{+\infty} (-1)^m \frac{x^{2m}}{(2m)!}, \quad \sin(x) = \sum_{m=0}^{+\infty} (-1)^m \frac{x^{2m+1}}{(2m+1)!}.$$

• *Cas des fractions $x \mapsto 1/(1-x)$ et $x \mapsto 1/(1+x)$.*

Ces fonctions appartiennent respectivement à $C_\infty(\mathbb{R} \setminus \{1\})$ et $C_\infty(\mathbb{R} \setminus \{-1\})$.

Cela étant, on sait que quel que soit le réel x différent de 1 et quel que soit le naturel M, on a

$$\sum_{m=0}^{M} x^m = \frac{1 - x^{M+1}}{1 - x}$$

donc

$$\frac{1}{1-x} = \frac{x^{M+1}}{1-x} + \sum_{m=0}^{M} x^m.$$

◦ Lorsque $x \in]-1, 1[$ et $x \neq 0$, on a

$$\lim_{M \to +\infty} \left| x^{M+1} \right| = \lim_{M \to +\infty} e^{(M+1) \ln(|x|)} = 0$$

et ainsi

$$\frac{1}{1-x} = \lim_{M \to +\infty} \left(\frac{x^{M+1}}{1-x} + \sum_{m=0}^{M} x^m \right) = \sum_{m=0}^{+\infty} x^m.$$

◦ Pour $x = 0$, il est clair que l'égalité est encore vérifiée.

◦ Pour $x = -1$ on a

$$\sum_{m=0}^{M} (-1)^m = \frac{1 - (-1)^{M+1}}{2}$$

donc il est clair que

$$\lim_{M\to+\infty} \sum_{m=0}^{M} (-1)^m$$

n'existe pas.

∘ Enfin, pour $x \in]-\infty, 1[\cup]1, +\infty[$, c'est-à-dire pour x de module strictement supérieur à 1, on a

$$\lim_{M\to+\infty} \left|x^{M+1}\right| = \lim_{M\to+\infty} e^{(M+1)\ln(|x|)} = +\infty$$

donc

$$\lim_{M\to+\infty} \sum_{m=0}^{M} x^m \;=\; \lim_{M\to+\infty} \left(\frac{1-x^{M+1}}{1-x}\right) = \infty.$$

Le cas de la fonction $x \mapsto 1/(1+x)$ se déduit bien sûr du cas précédent.

• *Cas des fonctions logarithmes.*

Faisons le raisonnement pour le cas de $x \mapsto \ln(1+x)$; l'autre cas s'en déduit directement.

Pour obtenir le développement, on pourrait penser à utiliser le développement limité de Taylor ; pour obtenir le résultat, il suffirait alors de montrer que le reste tend vers 0 pour les valeurs de x annoncées, et ne tend pas vers 0 pour les autres. Ceci peut se faire, mais pas pour toutes les valeurs de x (cf une annexe du syllabus de MATH 2007).

Nous allons utiliser une autre méthode, assez standard en fait, et qui a le mérite de pouvoir être utilisée pour d'autres fonctions, notamment pour la fonction arctan.

La fonction $x \mapsto \ln(1+x)$ appartient à $C_\infty(]-1, +\infty[)$ et on a

$$D\ln(1+x) = \frac{1}{1+x} = \sum_{m=0}^{+\infty} (-1)^m x^m, \quad \forall x \in]-1, 1].$$

En utilisant le théorème 3.3.2 (dérivation « terme à terme » d'une série de puissances), on obtient

$$\begin{aligned} D\ln(1+x) = \frac{1}{1+x} \;&=\; \sum_{m=0}^{+\infty} (-1)^m x^m \\ &=\; \sum_{m=0}^{+\infty} (-1)^m D\frac{x^{m+1}}{m+1} \\ &=\; D\sum_{m=0}^{+\infty} (-1)^m \frac{x^{m+1}}{m+1} \end{aligned}$$

puisque

$$S(x) = \sum_{m=0}^{+\infty} (-1)^m \frac{x^{m+1}}{m+1} = \sum_{m=1}^{+\infty} (-1)^{m-1} \frac{x^m}{m}$$

converge pour tout $x \in]-1, 1[$ par utilisation du théorème 3.3.1. Ainsi, comme les fonctions $x \mapsto \ln(1+x)$ et $x \mapsto S(x)$ sont dérivables sur $]-1, 1[$ et à valeurs réelles, l'égalité de leur dérivée sur cet intervalle implique l'existence l'un réel r tel que

$$\ln(1+x) = S(x) + r \quad \forall x \in]-1, 1[.$$

En prenant cette égalité en $x = 0$, on obtient $r = 0$ et finalement on conclut.

3.7 Annexe : Approximations polynomiales

3.7.1 Remarques

1) Si $x \mapsto P(x) = a_N x^N + \ldots + a_n x^n + \ldots + a_1 x + a_0$ est un polynôme de degré strictement supérieur à n et vérifie
$$\lim_{x \to x_0, x \neq x_0} \frac{f(x) - P(x - x_0)}{(x - x_0)^n} = 0$$
alors le polynôme $x \mapsto P_1(x) = a_n x^n + \ldots + a_1 x + a_0$ vérifie
$$\lim_{x \to x_0, x \neq x_0} \frac{f(x) - P_1(x - x_0)}{(x - x_0)^n} = 0.$$
En effet, on a
$$\lim_{x \to x_0, x \neq x_0} \frac{P(x - x_0) - P_1(x - x_0)}{(x - x_0)^n} = \sum_{j=n+1}^{N} \lim_{x \to x_0, x \neq x_0} a_j (x - x_0)^{j-n} = 0.$$

2) Si f est continu en x_0 alors
$$\lim_{x \to x_0} (f(x) - f(x_0)) = \lim_{x \to x_0, x \neq x_0} (f(x) - f(x_0)) = 0.$$
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
$$\lim_{x \to x_0, x \neq x_0} \frac{f(x) - P(x - x_0)}{(x - x_0)^n} = 0$$
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
$$\lim_{x \to x_0, x \neq x_0} (f(x) - P(x - x_0)) = \lim_{x \to x_0} (f(x) - P(x - x_0)) \; (= f(x_0) - P(0)).$$
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
$$\lim_{x \to x_0, x \neq x_0} \frac{f(x) - P(x - x_0)}{(x - x_0)^n} = \lim_{x \to x_0} \frac{f(x) - P(x - x_0)}{(x - x_0)^n} = 0.$$

3.7.2 Propriétés

Propriété(s) 3.7.1. *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
$$x \mapsto f(x_0) + (x - x_0) Df(x_0).$$

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

$$\lim_{x\to x_0, x\neq x_0} \frac{f(x)-P(x-x_0)}{(x-x_0)^n} = 0$$

alors, pour tout $k \in \{0,\ldots,n-1\}$, on a

$$\begin{aligned}
\lim_{x\to x_0, x\neq x_0} \frac{f(x)-P(x-x_0)}{(x-x_0)^k} &= \lim_{x\to x_0, x\neq x_0} \left(\frac{f(x)-P(x-x_0)}{(x-x_0)^n}\,(x-x_0)^{n-k}\right) \\
&= \lim_{x\to x_0, x\neq x_0} \left(\frac{f(x)-P(x-x_0)}{(x-x_0)^n}\right) \lim_{x\to x_0, x\neq x_0} (x-x_0)^{n-k} \\
&= 0
\end{aligned}$$

On conclut alors en utilisant la première remarque ci-dessus.

2) Soit $x \mapsto P(x) = a_0 + a_1 x + \ldots + a_n x^n$ un polynôme de degré inférieur ou égal à n tel que

$$\lim_{x\to x_0} \frac{f(x)-P(x-x_0)}{(x-x_0)^n} = 0.$$

Recherche du coefficient a_0. Vu 1) avec $k=0$, on a

$$\lim_{x\to x_0, x\neq x_0} (f(x)-P(x-x_0)) = 0.$$

Il s'ensuit que

$$\lim_{x\to x_0, x\neq x_0} f(x) = \lim_{x\to x_0, x\neq x_0} (f(x)-P(x-x_0)) + \lim_{x\to x_0, x\neq x_0} P(x-x_0) = a_0.$$

Le coefficient a_0 est donc unique.

Recherche du coefficient a_1. Vu 1) avec $k=1$, on a

$$\lim_{x\to x_0, x\neq x_0} \frac{f(x)-P(x-x_0)}{x-x_0} = 0.$$

Il s'ensuit que

$$\lim_{x\to x_0, x\neq x_0} \frac{f(x)-a_0}{x-x_0} = \lim_{x\to x_0, x\neq x_0} \frac{f(x)-P(x-x_0)}{x-x_0} + \lim_{x\to x_0, x\neq x_0} \frac{P(x-x_0)-a_0}{x-x_0} = a_1.$$

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 $x \mapsto P(x-x_0)$ 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

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{x\to x_0} \frac{f(x)-P(x-x_0)}{x-x_0} \ + \ a = a$$

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

$$\lim_{x \to 0} \frac{3x^2 \sin(1/x) - x\cos(1/x)}{x}$$

n'existe pas.

Cependant, la fonction f admet une approximation à l'ordre 2 en $x_0 = 0$ car

$$\lim_{x \to 0} \frac{f(x) - 0}{x^2} = \lim_{x \to 0} x \sin(1/x) = 0.$$

Annexe A

Petit formulaire pour les mathématiques et les sciences

A.1 L'alphabet grec

alpha	α	iota	ι	rhô	ρ
bêta	β	kappa	κ	sigma	σ, Σ
gamma	γ, Γ	lambda	λ, Λ	tau	τ
delta	δ, Δ	mu	μ	upsilon	υ, Υ
epsilon	ϵ, ε	nu	ν	phi	ϕ, φ, Φ
zêta, dzêta	ζ	xi, ksi	ξ, Ξ	khi	χ
êta	η	omicron	o	psi	ψ, Ψ
thêta	$\theta, \vartheta, \Theta$	pi	π, Π	omega	ω, Ω

A.2 Symboles usuels du langage mathématique

Notations habituelles pour les ensembles classiques de nombres

$\mathbb{N}$	ensemble des naturels positifs ou nul
$\mathbb{N}_0$	ensemble des naturels strictement positifs
$\mathbb{Z}$	ensemble des nombres entiers
$\mathbb{Z}_0$	ensemble des nombres entiers non nuls
$\mathbb{Q}$	ensemble des nombres rationnels
$\mathbb{Q}_0$	ensemble des nombres rationnels non nuls
$\mathbb{R}$	ensemble des nombres réels
$\mathbb{R}_0$	ensemble des nombres réels non nuls
$\mathbb{C}$	ensemble des nombres complexes
$\mathbb{C}_0$	ensemble des nombres complexes non nuls

Notations relevant de la théorie des ensembles

Un ensemble est désigné soit explicitement, en notant ses éléments entre accolades, soit de façon générique en utilisant (le plus souvent) une lettre majuscule. Ainsi, l'ensemble dont les éléments sont a, b, c, d, e est noté explicitement $\{a, b, c, d, e\}$. Lorsque l'ensemble contient une infinité d'éléments, on adapte cette notation.

Dans ce qui suit, A, B désignent deux ensembles.

Notation	Signification
$a \in A$	a appartient à l'ensemble A ou a est un élément de A
$A \subset B$	l'ensemble A est inclus dans l'ensemble B
	c'est-à-dire tout élément de A est un élément de B
$A = B$	les ensembles A et B sont les mêmes
	c'est-à-dire tout élément de A est élément de B et tout élément de B est élément de A
	c'est-à-dire $A \subset B$ et $B \subset A$
$A \cap B$	ensemble intersection de A et de B
	c'est-à-dire l'ensemble des élements qui appartiennent à la fois à A et à B
$A \cup B$	ensemble union de A et de B
	c'est-à-dire l'ensemble des éléments qui appartiennent à A ou à B
	c'est-à-dire l'ensemble des éléments qui appartiennent soit à A et pas à B, soit à B et pas à A, soit à A et à B
$\emptyset$	ensemble vide
	c'est-à-dire l'ensemble qui ne contient aucun élément
$A \setminus B$	ensemble A moins B
	c'est-à-dire l'ensemble des éléments de A qui n'appartiennent pas à B

Par exemple, l'ensemble des réels en lesquels la fonction cosinus s'annule est l'ensemble des réels qui sont égaux à $\pi/2$ auquel on ajoute un multiple entier de π ; cet ensemble est noté

$$\left\{\frac{\pi}{2} + k\pi, \quad k \in \mathbb{Z}\right\}.$$

L'ensemble de définition de la fonction tangente, quotient de la fonction sinus par la fonction cosinus, est l'ensemble des réels pour lesquels le cosinus ne s'annule pas ; il s'agit donc de l'ensemble

$$\mathbb{R} \setminus \left\{\frac{\pi}{2} + k\pi, \quad k \in \mathbb{Z}\right\}.$$

Notations relevant de la logique élémentaire

Soient P, Q deux propositions

Notation	Signification
$P \Rightarrow Q$	si la proposition P est vraie, alors la proposition Q est vraie ;
	on dit aussi
	- *il suffit que la proposition P soit vraie pour que Q le soit aussi,*
	- *il est nécessaire que la proposition Q soit vraie pour que P soit vrai,*
	- *pour que la proposition Q soit vraie, il est suffisant que P soit vrai*
	- *pour que la proposition P soit vraie, il est nécessaire que Q soit vrai*
$P \Leftrightarrow Q$	P et Q sont des propositions équivalentes
	c'est-à-dire $P \Rightarrow Q$ et $Q \Rightarrow P$
$\forall$	pour tout
$\forall x \in A$ on a ...	pour tout (ou quel que soit) l'élément x de l'ensemble A, on a ...
$\exists$	il existe
$\exists x \in A$ tel que ...	il existe un élément x de l'ensemble A tel que ...

A.3 Rappels sur les triangles et les angles

Cas d'égalité des triangles

Deux triangles sont dits égaux s'ils sont "superposables" c'est-à-dire si on obtient l'un à partir de l'autre par un déplacement dans le plan (qui n'affecte pas leur rigidité) ou encore si on obtient l'un à partir de l'autre par une translation suivie d'une rotation.

Deux triangles sont égaux dans chacun des cas suivants :
- ils ont un côté égal adjacent à deux angles égaux chacun à chacun
- ils ont un angle égal compris entre deux côtés égaux chacun à chacun
- ils ont les trois côtés égaux chacun à chacun.

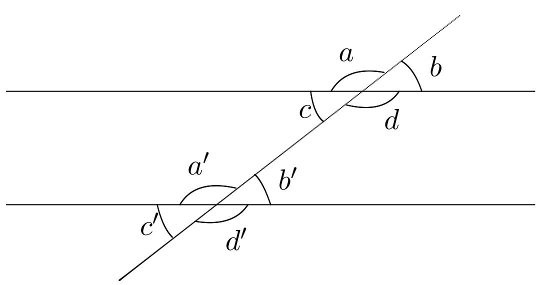
Cas de similitude des triangles

Deux triangles sont dits semblables si on obtient l'un à partir de l'autre par une similitude. (En géométrie, une similitude est une transformation qui conserve les rapports de distances.)

Deux triangles sont semblables dans chacun des cas suivants :
- ils ont deux angles égaux chacun à chacun
- ils ont un angle égal compris entre des côtés proportionnels
- ils ont les trois côtés proportionnels
- ils ont leurs côtés parallèles chacun à chacun
- ils ont leurs côtés perpendiculaires chacun à chacun.

Cas d'égalité des angles

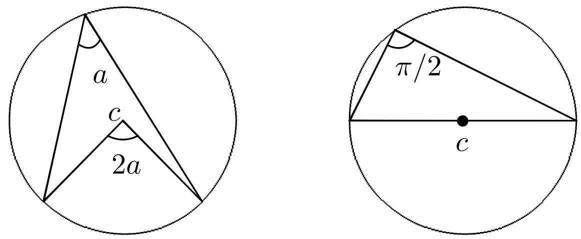
Considérons deux droites parallèles distinctes et une sécante.



Les angles alternes internes $c,\ b'$ (resp. $d,\ a'$) sont égaux.
Les angles alternes externes $a,\ d'$ (resp. $b,\ c'$) sont égaux.
Les angles opposés par le sommet b et c (resp. a et d, b' et c', a' et d') sont égaux.
Les angles correspondants a et a'(resp. b et b', c et c', d et d') sont égaux.

Angles et cercle

Un angle inscrit dans un cercle a la mesure de la moitié de l'angle au centre qui intercepte le même arc.



A.4 Quelques relations fondamentales de trigonométrie

Les fonctions sinus et cosinus sont définies sur $\mathbb{R}$ et périodiques de période 2π. On a

$$\sin x = 0 \;\Leftrightarrow\; \exists k \in \mathbb{Z} : x = k\pi; \qquad \cos x = 0 \;\Leftrightarrow\; \exists k \in \mathbb{Z} : x = \frac{\pi}{2} + k\pi.$$

Pour tout réel x qui n'annule pas le dénominateur, on a

$$\tan x = \frac{\sin x}{\cos x}, \quad \operatorname{cotan} x = \frac{\cos x}{\sin x}.$$

On a les relations suivantes (et de nombreuses conséquences !) pour tous réels x, y

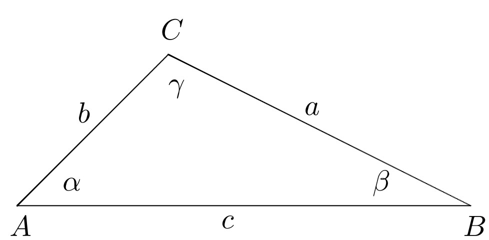
$\cos(-x) = \cos x$	$\sin(-x) = -\sin x$
$\cos\left(\frac{\pi}{2} - x\right) = \sin x$	$\cos(\pi - x) = -\cos x$ et $\sin(\pi - x) = \sin x$
$\cos(x + y) = \cos x \cos y - \sin x \sin y$	$\sin(x + y) = \sin x \cos y + \cos x \sin y$

Relations dans les triangles

On désigne par A, B, C les sommets d'un triangle et par a, b, c les longueurs des côtés opposés respectivement à ces sommets. Enfin, les mesures des angles (orientés positivement) de ce triangle sont respectivement appelées α, β, γ.

Triangle quelconque

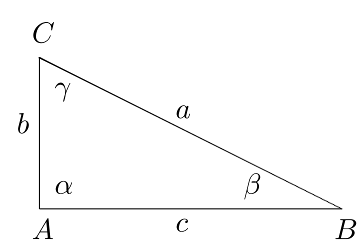
On a les formules suivantes.
$\alpha + \beta + \gamma = \pi$
$a^2 = b^2 + c^2 - 2bc \cos\ \alpha$
$b^2 = a^2 + c^2 - 2ac \cos\ \beta$
$c^2 = a^2 + b^2 - 2ab \cos\ \gamma$
$\dfrac{a}{\sin\ \alpha} = \dfrac{b}{\sin\ \beta} = \dfrac{c}{\sin\ \gamma}$



Triangle rectangle

Dans le cas particulier des triangles rectangles, les relations ci-dessus se simplifient de la manière suivante.

Le côté opposé à l'angle droit (ici α) se nomme hypoténuse.
On a les formules suivantes :
$\alpha + \beta + \gamma = \pi$ avec un des angles égal à $\pi/2$
$b = a\ \sin\beta = a\ \cos\gamma = c\ \tan\ \beta = c\ \operatorname{cotan}\ \gamma$
$c = a\ \sin\gamma = a\ \cos\beta = b\ \tan\ \gamma = b\ \operatorname{cotan}\ \beta$
$a^2 = b^2 + c^2$

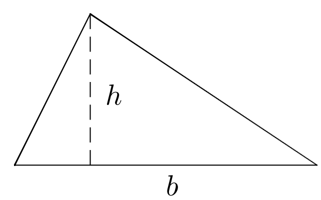


Dans un triangle rectangle, la longueur d'un côté de l'angle droit est égale à
- la longueur de l'hypoténuse multipliée par le sinus de l'angle opposé ou le cosinus de l'angle adjacent
- la longueur de l'autre côté multipliée par la tangente de l'angle opposé ou la cotangente de l'angle adjacent.

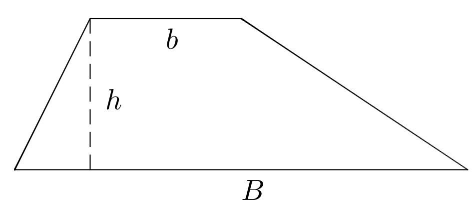
Dans un triangle rectangle, le carré de la longueur de l'hypoténuse est égal à la somme des carrés des longueurs des deux autres côtés.

A.5 Aires et volumes

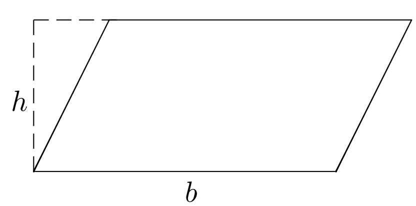
Aire d'un triangle =
la moitié du produit de la longueur d'un côté (b) et de la hauteur correspondante (h) c'est-à-dire $bh/2$



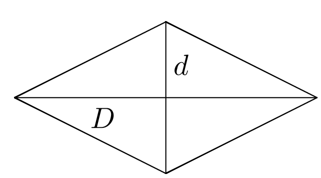
Aire d'un trapèze =
la moitié du produit de sa hauteur (h) par la somme des longueurs de ses bases (B et b) c'est-à-dire $(B+b)h/2$



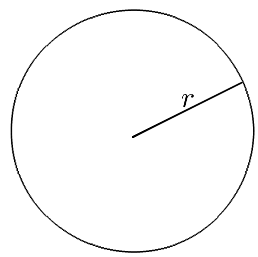
Aire d'un parallélogramme =
le produit de la longueur d'un côté (b) par la hauteur correspondante (h) c'est-à-dire bh



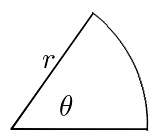
Aire d'un losange =
la moitié du produit des longueurs de ses diagonales (D et d) c'est-à-dire $Dd/2$



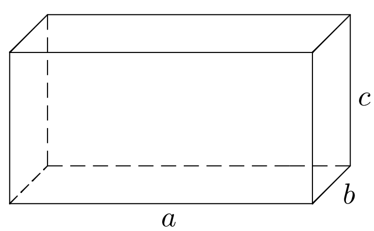
Aire d'un disque de rayon de longueur r =
le produit de π par le carré de la longueur du rayon (r) c'est-à-dire πr^2
Longueur de la circonférence (cercle) =
le double du produit de π par la longueur de son rayon (r) c'est-à-dire $2\pi r$



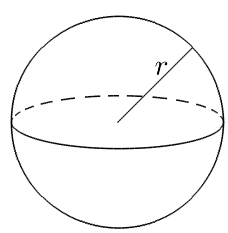
Aire d'une partie de disque de rayon r =
la moitié du produit de la mesure de l'angle en radian (θ) par le carré de la longueur du rayon (r) c'est-à-dire $\theta r^2/2$.
Longueur d'une partie de circonférence (cercle) =
le produit de la mesure de l'angle en radian (θ) par la longueur du rayon (r) c'est-à-dire θr



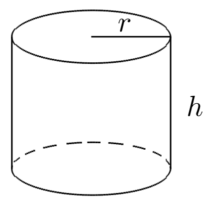
Volume d'un parallélépipède (dont les arêtes ont pour longueur a, b, c) = abc
Aire totale des 6 faces d'un parallélépipède = $2ab + 2ac + 2bc$



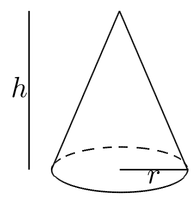
Volume d'une boule dans l'espace (volume sphérique) de rayon r =
le produit du cube de la longueur du rayon (r) par quatre tiers de π c'est-à-dire $4\pi r^3/3$
Aire d'une sphère =
le quadruple du produit du carré de la longueur du rayon (r) par π c'est-à-dire $4\pi r^2$



Volume d'un corps cylindrique de rayon r et de hauteur $h =$
le produit de l'aire du disque par la hauteur (h) du cylindre c'est-à-dire $\pi r^2 h$
Aire latérale d'un cylindre (surface cylindrique), sans compter les disques des bases $=$
le produit de la longueur du cercle par la hauteur (h) du cylindre c'est-à-dire $2\pi r h$



Volume d'un corps conique de hauteur h et dont la base a un rayon $r =$
le tiers du volume du cylindre de hauteur h et de base de même rayon c'est-à-dire $\pi r^2 h/3$
Aire latérale d'un cône, sans compter le disque de base $= \pi r\sqrt{r^2 + h^2}$



A.6 Dérivées des fonctions élémentaires

Dans ce qui suit, x désigne une variable réelle, m désigne un naturel strictement positif et r désigne un réel. Certaines dérivées peuvent être obtenues à partir d'autres ; il y a également de nombreuses autres expressions que l'on peut obtenir à partir de celles-ci !

Expression fonction	Domaine de définition et de continuité	Domaine de dérivabilité	Expression dérivée
r	$\mathbb{R}$	$\mathbb{R}$	0
x^m	$\mathbb{R}$	$\mathbb{R}$	mx^{m-1}
$\sin(x)$	$\mathbb{R}$	$\mathbb{R}$	$\cos(x)$
$\cos(x)$	$\mathbb{R}$	$\mathbb{R}$	$-\sin(x)$
$\tan(x)$	$\mathbb{R} \setminus \{\pi/2 + k\pi \ : \ k \in \mathbb{Z}\}$	$\mathbb{R} \setminus \{\pi/2 + k\pi \ : \ k \in \mathbb{Z}\}$	$1/\cos^2(x)$
$\text{cotan}(x)$	$\mathbb{R} \setminus \{k\pi \ : \ k \in \mathbb{Z}\}$	$\mathbb{R} \setminus \{k\pi \ : \ k \in \mathbb{Z}\}$	$-1/\sin^2(x)$
$\exp(x)$	$\mathbb{R}$	$\mathbb{R}$	$\exp(x)$
$\arcsin x$	$[-1, 1]$	$]-1, 1[$	$1/\sqrt{1-x^2}$
$\text{arcos}(x)$	$[-1, 1]$	$]-1, 1[$	$-1/\sqrt{1-x^2}$
$\arctan(x)$	$\mathbb{R}$	$\mathbb{R}$	$1/(1+x^2)$
$\text{arcotan}(x)$	$\mathbb{R}$	$\mathbb{R}$	$-1/(1+x^2)$
$\ln(x)$	$]0, +\infty[$	$]0, +\infty[$	$1/x$
x^r	$]0, +\infty[$	$]0, +\infty[$	rx^{r-1}