UNIVERSITE DE LIEGE

Faculté des Sciences

ANALYSE MATHEMATIQUE

Notes du cours des premiers Bacheliers
en sciences mathématiques ou en sciences physiques

Jean SCHMETS

Année académique 2004–2005

Introduction

Ce livre contient la première partie du cours d'analyse mathématique que j'enseigne en première candidature en sciences mathématiques ou en sciences physiques. La deuxième partie concerne le calcul intégral et fait l'objet d'un volume séparé.

Comme tout cours d'initiation à l'analyse, il développe essentiellement une description de l'espace euclidien $\mathbb{R}^n$ de dimension n ainsi qu'une étude de la continuité, de la dérivabilité et de la primitivabilité des fonctions, et se termine avec la considération des équations différentielles linéaires à coefficients constants et de quelques équations différentielles ordinaires.

En rédigeant ces notes, j'ai désiré rencontrer le souhait émis par les étudiants de disposer d'un texte proche de la matière enseignée. Je n'ai pu cependant m'empêcher d'y inclure quelques compléments théoriques (parfois présentés sous la forme d'exercices).

Ces notes sont complétées par un **Cahier d'Exercices**. C'est la raison pour laquelle elles ne contiennent pas beaucoup d'exemples et exercices, malgré l'importance que je leur accorde.

Les textes placés entre les symboles "$* \rightarrow$" et "$\leftarrow *$" font appel à de la matière ultérieure et sont à réserver pour une deuxième lecture.

J. Schmets

Quelques repères chronologiques de mathématiciens cités

1600 1650 1700 1750 1800 1850 1900

- PASCAL Blaise (1623–1662)
- NEWTON Isaac (1642–1727)
- LEIBNIZ Gottfried Wilhelm (1646–1716)
- ROLLE Michel (1652–1719)
- BERNOULLI Jacques (1654–1705)
- HOSPITAL Guillaume de L' (1661–1704)
- DE MOIVRE Abraham (1667–1754)
- TAYLOR Brook (1685–1731)
- MACLAURIN Colin (1698–1746)
- EULER Leonhard (1707–1783)
- CLAIRAUT Hermann (1713–1765)
- LAGRANGE Joseph (1736–1813)
- LAPLACE Pierre de (1749–1827)
- LEGENDRE Adrien-Marie (1752–1833)
- GAUSS Carl-Friedrich (1777–1855)
- BOLZANO Bernhard (1781–1848)
- CAUCHY Augustin (1789–1857)

1800 1850 1900 1950 2000

ABEL	Niels (1802–1829)
JACOBI	Carl (1804–1851)
MORGAN	Augustus De (1806–1871)
HESSE	Otto (1811–1874)
WEIERSTRASS	Karl (1815–1897)
HEINE	Eduard (1821–1881)
RIEMANN	Bernhard (1826–1866)
HERMITE	Charles (1822–1901)
LEVY	Maurice (1838–1910)
SCHROEDER	Ernst (1841–1902)
SCHWARZ	Herman (1843–1921)
CANTOR	Georg (1845–1918)
LORENTZ	Hendrick (1853–1928)
PICARD	Emile (1856–1941)
CESARO	E. (1859–1906)
JENSEN	Johann (1859–1925)
HŐLDER	Ludwig (1859–1937)
MINKOWSKI	Hermann (1864–1909)
STEINITZ	Ernst (1871–1928)
BOREL	Emile (1871–1956)
BERNSTEIN	Felix (1878–1956)
HALMOS	Paul (1914–)

Chapitre 1

Théorie naïve des ensembles

1.1 Introduction

Les processus fondamentaux des mathématiques sont
a) introduire des *objets* dits *mathématiques*,
b) démontrer que certaines *relations* entre ces objets sont vraies; on dit que ce sont des *théorèmes*.

Les objets mathématiques sont les nombres, les fonctions, les fonctions continues, les fonctions dérivables, les fonctions intégrables, ... Les relations sont les assertions (qui peuvent donc être vraies ou fausses) qu'on peut formuler sur ces objets. Les vraies ou *théorèmes* sont celles qu'on *démontre*, c'est-à-dire qu'on peut déduire logiquement d'un certain nombre d'*axiomes*. Les axiomes sont la formulation mathématique des propriétés "évidentes" des êtres auxquels on désire appliquer les mathématiques.

Il ne faut pas voir dans ce qui précède des définitions correctes du point de vue logique mais seulement une introduction imagée qui se précisera au fur et à mesure des études. En fait, la logique mathématique et la théorie formelle des ensembles constituent des domaines fort abstraits et demandent de longs développements. Il n'est donc pas possible de les voir, en première candidature, comme introduction à un cours d'analyse mathématique.

Cependant la logique mathématique et les propriétés de la théorie des ensembles sont fondamentales en mathématiques et tout au long de ce cours, nous allons les utiliser. La méthode utilisée consiste, si cela est possible, à introduire les notions de manière définitive et d'en étudier les propriétés de manière rigoureuse. En cas d'impossibilité, le fait est mentionné clairement, le vocabulaire correct est introduit et les règles d'utilisation sont précisées, réservant la justification à une étude ultérieure.

1.2 Quelques locutions et symboles

En ce qui concerne la logique mathématique, nous allons nous limiter à introduire un vocabulaire correct et les règles d'utilisation de ce vocabulaire.

Soient des relations R, S.

a) La "*négation de R*" est désignée généralement par l'assemblage "$\not R$", c'est-à-dire "R superposé de /". On recourt aussi souvent à des notations différentes telle que "*non R*", "$\neg R$".

Une relation est fausse si sa négation est vraie; elle est vraie si sa négation est fausse.

b) "*R ou S*" est une relation qui est vraie si l'une au moins des relations R, S est vraie.

Par exemple, si R est la relation "5 est strictement inférieur à 6" et si S est la relation "5 est égal à 6", la relation "R ou S" est la relation "5 est inférieur ou égal à 6" et est donc vraie.

En logique et en mathématique, le mot *ou* est toujours pris au sens non disjonctif. Il faut donc recourir à une périphrase pour traduire le ou disjonctif de la langue française.

Ces méthodes fondamentales de construction de relations permettent d'en introduire d'autres qui jouent un rôle tout aussi important.

a) "*R et S*" est une relation qui est vraie si les deux relations R, S sont vraies. En fait, "R et S" est défini comme étant la relation

$$\text{“}R \text{ et } S\text{”} = \text{“}\neg(\neg R \text{ ou } \neg S)\text{”}.$$

Par exemple, si R est la relation "r est un multiple de 2" et si S est la relation "r est un multiple de 3", la relation "R et S" est vraie si "r est un multiple de 6".

Cela étant, on a

$$\text{“}\neg(R \text{ et } S)\text{”} = \text{“}\neg R \text{ ou } \neg S\text{”}$$

et

$$\text{“}\neg(R \text{ ou } S)\text{”} = \text{“}\neg R \text{ et } \neg S\text{”}.$$

b) "$R \Rightarrow S$" qui se lit "*R implique S*" est la relation

$$\text{“}R \Rightarrow S\text{”} = \text{“}S \text{ ou } \neg R\text{”}.$$

Elle exprime que si R est vrai, alors S est vrai.

c) "$R \Leftrightarrow S$" qui se lit "*R si et seulement si S*" est la relation

$$\text{“}R \Leftrightarrow S\text{”} = \text{“}R \Rightarrow S \ \text{ et } \ S \Rightarrow R\text{”}.$$

1.3 Ensembles

1.3.1 Définition

En ce qui concerne les ensembles, nous allons recourir à la "*théorie naïve des ensembles*". Le point de vue naïf consiste à introduire la notion d'ensemble de manière vague, puis d'en donner les propriétés sans démonstration. Cette manière vague peut définir un ensemble comme étant une notion fondamentale qui jouit de propriétés particulières ou comme étant la collection des êtres mathématiques qui vérifient une propriété. (Remarquons de suite que cette deuxième manière de procéder n'est en aucune sorte une définition: elle définirait la notion "ensemble" par une autre "collection" qui n'a pas été définie auparavant.)

Cependant cette notion d'ensemble n'est pas que formelle; elle procède en fait d'une base intuitive. Pour s'en assurer, il suffit de considérer l'ensemble des nombres réels, l'ensemble des nombres complexes, ...

Un ensemble est déterminé par ses éléments qui sont donnés indifféremment

a) d'une manière explicite, c'est-à-dire par un symbole individuel tel que 1, 2, 3, ...

b) par un symbole générique affecté d'indices variant dans des ensembles: on trouve par exemple x_j, $x_{j,k}$, ...

c) par un symbole générique seulement s'il n'est pas nécessaire de les distinguer.

Un ensemble est donné indifféremment

a) de manière explicite en donnant la liste complète de ses éléments placés entre accolades et séparés par un symbole approprié (très souvent une virgule): par exemple $\{1, 2, 3\}$. Bien sûr, cette manière explicite ne peut être utilisée que pour les ensembles "finis"; aussi on accepte également de suggérer la liste des éléments de l'ensemble en recourant aux trois points de suspension. Ainsi, $\{a, b, \ldots, z\}$ représente l'ensemble des lettres de l'alphabet et $\{1, 2, 3, \ldots\}$ représente l'ensemble des nombres entiers supérieurs ou égaux à 1. (Les trois points de suspension doivent évidemment avoir une signification claire.)

Un *singleton* est un ensemble contenant un et un seul élément. Si a est cet élément, le singleton peut donc être noté $\{a\}$.

b) en plaçant entre accolades le symbole générique suivi d'un symbole approprié (très souvent ":") puis la propriété qui caractérise ses éléments. On obtient de la sorte une formule du genre $\{\, x : P\}$ qui se lit "ensemble des x tels que P";

c) par un symbole (généralement une lettre majuscule) s'il n'est pas nécessaire d'en détailler les éléments. En particulier, certains symboles réfèrent à des ensembles précis; on trouve notamment:

$\mathbb{N}$ = ensemble des nombres entiers supérieurs ou égaux à 0,

$\mathbb{N}_0$ = ensemble des nombres entiers supérieurs ou égaux à 1,

$\mathbb{Z}$ = ensemble des nombres entiers positifs, nul ou négatifs,
$\mathbb{R}$ = ensemble des nombres réels,
$\mathbb{Q}$ = ensemble des nombres réels rationnels,
$\mathbb{C}$ = ensemble des nombres complexes.

1.3.2 Relations entre éléments et parties d'un ensemble

Soit A un ensemble.

a) **Appartenance.** Nous écrivons $a \in A$ pour signaler que a est un élément de A. La formule $a \in A$ se lit "*a est un élément de A*" ou "*a appartient à A*". On trouve aussi l'écriture $A \ni a$ qui se lit "*A contient a*".

Si a n'est pas élément de A, nous écrivons $a \notin A$, ce qui se lit "*a n'est pas élément de A*" ou "*a n'appartient pas à A*". On trouve aussi $A \not\ni a$ qui se lit "*A ne contient pas a*".

b) **Inclusion.** Si B est un ensemble, nous écrivons $B \subset A$ pour signaler que tout élément de B appartient à A. La formule $B \subset A$ se lit "*B est inclus dans A*" ou "*B est un sous-ensemble de A*" ou "*B est une partie de A*". On trouve aussi la notation $A \supset B$ qui se lit "*A contient B*".

Sinon nous écrivons $B \not\subset A$, ce qui se lit "*B n'est pas inclus dans A*".

c) **Egalité.** Si a et b sont deux éléments de A, nous écrivons $a = b$ pour signaler qu'il s'agit du même élément. La formule $a = b$ se lit "*a est égal à b*". De même, si A et B sont des ensembles, nous écrivons $A = B$ si tout élément de A est élément de B et inversement. Cette notation $A = B$ se lit "*A est égal à B*". Elle a donc lieu si et seulement si on a $A \subset B$ et $B \subset A$.

Si a, b désignent deux éléments distincts de A, nous écrivons $a \neq b$, ce qui se lit "*a diffère de b*". Si les ensembles A, B ne sont pas égaux, nous écrivons $A \neq B$, ce qui se lit "*A diffère de B*" ou "*A n'est pas égal à B*".

d) **Ensemble vide.** Tout ensemble A contient trivialement deux parties, à savoir A lui-même et l'ensemble vide, noté $\emptyset$, ensemble conventionnel qui ne contient pas d'élément.

e) **Ensemble des parties.** Etant donné un ensemble A, $\wp(A)$ désigne l'*ensemble des parties de A*.

1.3.3 Ensembles associés à des ensembles

A des parties A et B de l'ensemble X, on associe les parties suivantes de X.

a) **Union.** L'*union $A \cup B$ de A et B* est l'ensemble des éléments qui appartiennent à A ou à B, c'est-à-dire à l'un au moins des ensembles A, B. La notation $A \cup B$ se lit "*A union B*".

b) **Intersection.** L'*intersection* $A \cap B$ *de* A *et* B est l'ensemble des éléments qui appartiennent à A et à B. La notation $A \cap B$ se lit "A *inter* B".

c) **Différence, complémentaire.** La *différence* $A \setminus B$ *de* A *et* B est l'ensemble des éléments de A qui n'appartiennent pas à B. La notation $A \setminus B$ se lit "A *moins* B". Si B est inclus dans A, on écrit parfois $\complement_A B$ à la place de $A \setminus B$. La notation $\complement_A B$ se lit "*complémentaire de* B *dans* A".

d) **Différence symétrique.** La *différence symétrique* $A \triangle B$ *de* A *et* B est l'ensemble des éléments de $A \cup B$ qui n'appartiennent pas à $A \cap B$. On a donc

$$A \triangle B = B \triangle A = (A \setminus B) \cup (B \setminus A).$$

1.3.4 Inclusions et identités remarquables

Les opérations $\cup$, $\cap$, $\setminus$, $\complement$ et $\triangle$ se combinent entre elles pour donner lieu à un véritable calcul entre ensembles. Certaines de ces opérations constituent des inclusions et identités remarquables qui permettent souvent d'alléger les autres opérations. Voici les plus importantes d'entre elles.

Proposition 1.3.4.1 *Si A, B, C sont des parties de l'ensemble X, on a les propriétés suivantes:*

a) $\complement_X X = \emptyset$, $\complement_X \emptyset = X$,

b) $\complement_X \complement_X A = A$,

c) $A \cup \complement_X A = X$, $A \cap \complement_X A = \emptyset$,

d) $A \subset B \Rightarrow (A \cup C \subset B \cup C)$,
$A \subset B \Rightarrow (A \cap C \subset B \cap C)$,
$(A \subset B, B \subset C) \Rightarrow A \subset C$,

e) $A \cap B \subset A \subset A \cup B$,

f) $$\begin{aligned} A \subset B &\Leftrightarrow (A \cup B = B) \quad \Leftrightarrow \quad (A \cap B = A) \\ &\Leftrightarrow \complement_X A \supset \complement_X B \\ &\Leftrightarrow (A \cap \complement_X B = \emptyset) \quad \Leftrightarrow \quad ((\complement_X A) \cup B = X). \end{aligned}$$

En particulier, on a $A \cup A = A$, $A \cap A = A$, $\emptyset \cup A = A$, $\emptyset \cap A = \emptyset$, $A \cup X = X$ et $A \cap X = A$.

g) $(C \subset A, C \subset B) \Rightarrow (C \subset A \cap B)$,
$(A \subset C, B \subset C) \Rightarrow (A \cup B \subset C)$,

h) $A \cup B = B \cup A$, $A \cup (B \cup C) = (A \cup B) \cup C$,
$A \cap B = B \cap A$, $A \cap (B \cap C) = (A \cap B) \cap C$,

i) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,

j) $\complement_X(A \cup B) = (\complement_X A) \cap (\complement_X B)$,
$\complement_X(A \cap B) = (\complement_X A) \cup (\complement_X B)$.■

Quelques remarques s'imposent.

a) Les formules h) permettent de donner un sens aux notations $A \cup B \cup C$ et $A \cap B \cap C$, à savoir

$$A \cup B \cup C = A \cup (B \cup C) = (A \cup B) \cup C$$

et

$$A \cap B \cap C = A \cap (B \cap C) = (A \cap B) \cap C.$$

b) Par contre, les formules i) établissent clairement que des notations telles que $A \cup B \cap C$ et $A \cap B \cup C$ sont totalement dénuées de sens et doivent être proscrites.

c) Ces propriétés donnent lieu à la *loi de dualité* ou *loi de Morgan: si on a une relation* $A \subset$ (resp. $=; \supset$) E_1 *où* A *est une partie de l'ensemble* X *et où* E_1 *est une expression qui ne fait intervenir que des parties de* X *et les symboles* $\cup$ *et* $\cap$*, on a également* $X \setminus A \supset$ (resp. $=; \subset$) E_2 *où* E_2 *désigne l'expression qu'on obtient à partir de* E_1 *en remplaçant chaque partie de* X *par son complémentaire dans* X*,* $\cup$ *par* $\cap$ *et* $\cap$ *par* $\cup$ *respectivement.*

Exercice. Vérifier que

a) $A \bigtriangleup B = B \bigtriangleup A$

b) $A \bigtriangleup \emptyset = A$

c) $A \bigtriangleup X = X \setminus A$

d) $A \bigtriangleup A = \emptyset$

e) $A \bigtriangleup (X \setminus A) = X$

f) $(A \bigtriangleup B) \bigtriangleup C = (B \bigtriangleup C) \bigtriangleup A = (C \bigtriangleup A) \bigtriangleup B$ et cet ensemble est égal à

$$\Big(A \setminus (B \cup C)\Big) \bigcup \Big(B \setminus (C \cup A)\Big) \bigcup \Big(C \setminus (A \cup B)\Big) \bigcup \Big(A \cap B \cap C\Big).$$

1.3.5 Union et intersection de plusieurs ensembles

Soit X un ensemble.

Nous venons de donner un sens à des notations telles que $A \cup B \cup C$ et $A \cap B \cap C$, si A, B, C sont des parties de X.

Plus généralement, si J est un ensemble et si, pour tout $j \in J$, A_j est une partie de X, on introduit

a) l'*union* $\cup_{j \in J} A_j$ *des* A_j comme étant l'ensemble des éléments qui appartiennent à l'un au moins des ensembles A_j,

b) l'*intersection* $\cap_{j\in J} A_j$ *des* A_j comme étant l'ensemble des éléments qui appartiennent à chacun des ensembles A_j.

En particulier (et ce sera souvent ce cas qui arrivera dans la suite), J peut être une partie de $\mathbb{N}$. Nous utilisons alors les notations suivantes: si p et q sont des entiers tels que $p \leq q$, chacune des notations

$$\bigcup_{j=p}^{q} A_j, \qquad A_p \cup \ldots \cup A_q$$

désigne l'union des ensembles A_p, ..., A_q tandis que chacune des notations

$$\bigcap_{j=p}^{q} A_j, \qquad A_p \cap \ldots \cap A_q$$

désigne leur intersection. Si p est un entier, $\cup_{j=p}^{\infty} A_j$ désigne l'union des ensembles A_p, A_{p+1}, ... et $\cap_{j=p}^{\infty} A_j$ leur intersection.

Deux parties A, B d'un même ensemble X sont *disjointes* si leur intersection est vide. Plus généralement, des parties $(A_j)_{j\in J}$ d'un même ensemble X sont *disjointes deux à deux* si $A_k \cap A_l = \emptyset$ pour tous k, $l \in J$ tels que $k \neq l$.

Cela étant, des parties $(A_j)_{j\in J}$ d'un même ensemble X constituent une *partition de* X si elles sont disjointes deux à deux et constituent un *recouvrement de* X (c'est-à-dire que leur union contient X).

1.3.6 Produits finis d'ensembles

Si A_1, ..., A_J sont des ensembles en nombre fini, leur *produit*, noté indifféremment

$$A_1 \times \cdots \times A_J \quad \text{ou} \quad \prod_{j=1}^{J} A_j$$

est l'ensemble dont les éléments sont les J-uples ordonnés $(a_1, \ldots, a_J)$ tels que $a_1 \in A_1$, ..., $a_J \in A_J$.

Si on a $A_1 = \ldots = A_J = A$, il est plutôt noté A^J.

1.4 Quantificateurs

Signalons à présent deux quantificateurs qui sont introduits en logique mathématique:

a) $\forall$ qui se lit "*pour tout*",

b) $\exists$ qui se lit "*il existe*".

Ils sont à la base de la plupart des raisonnements et apparaissent le plus souvent dans des assertions telles que

$$\text{“}\forall x \in A,\ \text{on a } R\text{”} \quad \text{et} \quad \text{“}\exists x \in A \ \text{ tel que } \ R\text{”}.$$

Ces deux assertions ne sont pas indépendantes car la négation de la première est "$\exists x \in A$ tel que $\neg R$" et celle de la seconde "$\forall x \in A$, on a $\neg R$".

Dès lors, par exemple, la négation de l'assertion

$$\text{“}\forall x \in A, \exists y \in B \text{ tel que } R\text{”}$$

est

$$\text{“}\exists x \in A \text{ tel que, } \forall y \in B,\ \text{on a } \neg R\text{”}.$$

1.5 Applications

1.5.1 Définition

Définitions. Soient A, B deux ensembles. Une *application f de A dans B* est une loi qui, à tout élément a de A, associe un élément $f(a)$ de B. On écrit explicitement

$$f\colon A \to B \quad a \mapsto f(a)$$

mais parfois on se contente d'une des notations moins explicites suivantes

$$f\colon A \to B \ \text{ ou } \ f\colon a \mapsto f(a)$$

si aucune ambiguïté ne peut en résulter.

Une *fonction définie sur A* est une application de A dans $\mathbb{C}$.

Exemple. Etant donné un ensemble A, *l'application identité de A*, notée id_A, est définie par

$$\mathrm{id}_A\colon A \to A \quad a \mapsto a.$$

Définition. Le *graphe* de l'application f de A dans B, noté $\mathrm{G}(f)$ ou même G si aucune ambiguïté sur f n'est possible est l'ensemble

$$\mathrm{G}(f) = \{\, (a, f(a)) : a \in A\}\,;$$

c'est donc une partie de $A \times B$.

Proposition 1.5.1.1 *Une partie G de $A \times B$ est le graphe d'une application de A dans B si et seulement si les deux conditions suivantes sont vérifiées*
(a) $\forall a \in A, \exists b \in B$ *tel que* $(a, b) \in G$,
(b) $((a, b_1), (a, b_2) \in G) \Rightarrow b_1 = b_2$.

Preuve. La condition est évidemment nécessaire. Pour établir sa suffisance, il suffit de vérifier que l'application f de A dans B qui, à tout $a \in A$, associe l'élément unique $b \in B$ tel que $(a, b) \in G$, a G pour graphe.∎

Remarque. En fait, on peut définir le concept d'application de A dans B comme étant une partie G de $A\times B$ qui vérifie les conditions (a) et (b) de la proposition précédente. Cela évite de devoir définir les applications en recourant au mot "loi" qui lui n'a pas été défini. Cependant, dans une première étude des applications, il semble plus adéquat de recourir aux "lois".

Définitions. Soit f une application de A dans B.

Si A' est une partie non vide de A, l'ensemble $\{ f(a) : a \in A'\}$ est noté $f(A')$ et est appelé *image de A' par f* ou plus précisément *image directe de A' par f*. En particulier, pour tout $a \in A$, $\{f(a)\}$ est l'image de $\{a\}$, $f(a)$ étant appelé *la valeur de f en a*. De plus, nous posons $f(\emptyset) = \emptyset$.

Définition. Si B' est une partie de B, l'ensemble $\{ a \in A : f(a) \in B'\}$ est noté

$$\overset{-1}{f}(B') \quad \text{ou} \quad f^{-1}(B')$$

et est appelé *image inverse de B' par f*; en particulier, pour tout $b \in B$, l'image inverse de $\{b\}$, à savoir

$$f^{-1}(\{b\}) = \{ a \in A : f(a) = b\},$$

est un ensemble qui peut contenir plus d'un élément ou être vide. On écrit bien souvent $f^{-1}(b)$ à la place de $f^{-1}(\{b\})$. Il est clair que $f^{-1}(\emptyset) = \emptyset$.

Proposition 1.5.1.2 *Si f est une application de A dans B,*
a) $A' \subset A'' \subset A \Rightarrow f(A') \subset f(A'')$,
b) $B' \subset B'' \subset B \Rightarrow f^{-1}(B') \subset f^{-1}(B'')$,
c) *la loi f qui, à toute partie A' de A, associe la partie $f(A')$ de B est une application de $\wp(A)$ dans $\wp(B)$,*
d) *la loi f^{-1} qui, à toute partie B' de B, associe la partie $f^{-1}(B')$ de A est une application de $\wp(B)$ dans $\wp(A)$,*

e) $A_j \subset A$ *pour tout* $j \in J$ *implique*

$$f\left(\bigcup_{j\in J} A_j\right) = \bigcup_{j\in J} f(A_j) \quad et \quad f\left(\bigcap_{j\in J} A_j\right) \subset \bigcap_{j\in J} f(A_j),$$

f) $B_j \subset B$ *pour tout* $j \in J$ *implique*

$$f^{-1}\left(\bigcup_{j\in J} B_j\right) = \bigcup_{j\in J} f^{-1}(B_j) \quad et \quad f^{-1}\left(\bigcap_{j\in J} B_j\right) = \bigcap_{j\in J} f^{-1}(B_j),$$

g) *pour tout* $A' \subset A$ *et tout* $B' \subset B$, *on a*

$$f(A \setminus A') \ ? \ B \setminus f(A') \quad et \quad f^{-1}(B \setminus B') = A \setminus f^{-1}(B').\blacksquare$$

Remarque. On a l'habitude de retenir les points e), f) et g) ci-dessus en disant que "pour les images inverses, les symboles $\cup$, $\cap$ et $\setminus$ se comportent bien tandis que pour les images directes tout va mal sauf pour $\cup$".

Remarque. Voici un exemple d'application f pour laquelle l'image $f(\bigcap_{j\in J} A_j)$ diffère de $\bigcap_{j\in J} f(A_j)$. Il suffit de prendre $A = \{0,1\}$, $B = \{0\}$ et de définir f par $f(0) = f(1) = 0$. De fait, il vient $f(\{0\} \cap \{1\}) = \emptyset$ et $f(\{0\}) \cap f(\{1\}) = B$.

En voici un autre exemple qui ne fait pas intervenir $\emptyset$. Il suffit de poser $A = \{1,2,3\}$, $B = \{1,2\}$, $f(1) = 1$, $f(2) = 2$ et $f(3) = 1$. En effet, il vient alors $f(\{1,2\} \cap \{2,3\}) = \{2\}$ et $f(\{1,2\}) \cap f(\{2,3\}) = B$.

Définition. Etant donné les aplications $f\colon A \to B$ et $g\colon B \to C$, il est clair que la loi

$$g \circ f\colon A \to C \quad a \mapsto g(f(a))$$

est une application de A dans C. Elle est appelée *composition de* f *et de* g et est souvent notée $g(f)$.

Proposition 1.5.1.3 *Etant donné* $f\colon A \to B$ *et* $g\colon B \to C$, *et une partie* C' *de* C, *on a*

$$(g \circ f)^{-1}(C') = \left(f^{-1} \circ g^{-1}\right)(C').\blacksquare$$

1.5.2 Injections, surjections et bijections

Définitions. Une application f de A dans B est une

a) *injection* si l'image inverse de tout élément de B contient un élément au plus; il revient au même de dire que, si deux éléments de A ont même image, alors ils sont égaux, ou que les images de deux éléments distincts de A sont différentes,

b) *surjection*, on dit aussi une *application de A sur B*, si on a $f(A) = B$, c'est-à-dire que, pour tout $b \in B$, il existe $a \in A$ tel que $f(a) = b$,

c) *bijection*, on dit aussi une *correspondance biunivoque*, si f est à la fois une injection et une surjection. Remarquons immédiatement qu'alors la loi qui, à tout élément de B, associe l'élément unique de A dont il est l'image est également une bijection de B sur A: on la note f^{-1} et cela ne cause pas d'ambiguïté.

Proposition 1.5.2.1 *Une application $f: A \to B$ est*

a) *injective si et seulement s'il existe une application $g: B \to A$ telle que $g \circ f = \mathrm{id}_A$,*

b) *surjective s'il existe une application $g: B \to A$ telle que $f \circ g = \mathrm{id}_B$,*
$* \to$ *la réciproque ayant lieu si on admet l'axiome du choix* $\leftarrow *$,

c) *bijective si et seulement s'il existe des applications $g: B \to A$ et $h: B \to A$ telles que $g \circ f = \mathrm{id}_A$ et $f \circ h = \mathrm{id}_B$. De plus on a alors $g = h$.*

Preuve. a) La condition est nécessaire. Choisissons un point a_0 de A. On vérifie de suite que $g: B \to A$ défini par $g(f(a)) = a$ pour tout $a \in A$ et $g(b) = a_0$ pour tout $b \in B \setminus f(A)$ convient. La condition est suffisante car, si les points a_1, a_2 de A ont même image, il vient $a_1 = g(f(a_1)) = g(f(a_2)) = a_2$.

b) De fait, pour tout $b \in B$, on a $f(g(b)) = b$.

$* \to$ Inversement, l'axiome du choix affirme l'existence d'une application $g: B \to A$ telle que $g(b) \in f^{-1}(\{b\})$ pour tout $b \in B$. $\leftarrow *$

c) La condition est nécessaire. Il suffit de définir g et h par $g(f(a)) = h(f(a)) = a$ pour tout $a \in A$. La suffisance de la condition résulte aussitôt de a) et b).∎

Proposition 1.5.2.2 *Si les applications $f: A \to B$ et $g: B \to C$ sont injectives* (resp. *surjectives; bijectives*)*, il en est de même pour $g \circ f$.*∎

$* \to$

Définitions. L'ensemble A *est en bijection* ou *en correspondance biunivoque avec* l'ensemble B, ce qu'on note $A \simeq B$, s'il existe une bijection de A sur B. Il est clair que A est en bijection avec B si et seulement si B est en bijection avec A. Ceci permet d'introduire la locution "*A et B sont en bijection* ou *en correspondance biunivoque*" pour traduire ce fait.

Remarque. Le fait que deux ensembles soient en bijection ne signifie pas qu'ils ont le "même nombre d'éléments", ce qui serait du reste une notion à définir dans le cas des ensembles non finis. Ainsi l'ensemble $\mathbb{N}_0$ est en bijection avec l'ensemble $\{2,4,6,\ldots\}$ des entiers strictement positifs et pairs, comme l'établit l'application $f\colon n \mapsto 2n$. De même, l'intervalle $]0,1[$ est en bijection avec l'intervalle $]a,b[$ quels que soient $a,b \in \mathbb{R}$ tels que $a < b$, comme le montre l'application $f\colon x \mapsto a + (b-a)x$. Cependant on peut établir qu'un ensemble fini A est en bijection avec un ensemble B si et seulement si B est fini et a le même nombre d'éléments que A.

Proposition 1.5.2.3 a) *Pour tout ensemble A, $A \simeq A$.*

b) *Etant donné deux ensembles A et B, on a $(A \simeq B) \Leftrightarrow (B \simeq A)$.*

c) *Etant donné des ensembles A, B et C,*

$$(A \simeq B,\, B \simeq C) \Rightarrow (A \simeq C).\blacksquare$$

Exemple. *Si les ensembles A et B sont en bijection, alors les ensembles $\wp(A)$ et $\wp(B)$ sont en bijection.*

Proposition 1.5.2.4 *Pour tout ensemble A, il n'existe pas de bijection de A sur $\wp(A)$.*

Preuve. Supposons qu'il existe une bijection f de A sur $\wp(A)$. Posons $A' = \{\, a \in A : a \notin f(a)\}$. Bien sûr, A' est une partie de A. Dès lors, il existe un élément a' de A tel que $f(a') = A'$. Cet élément a' doit alors appartenir à A' ou à $A \setminus A'$. Mais, d'une part, $a' \in A'$ ne peut avoir lieu car cela signifie que a' n'appartient pas à $f(a') = A'$ et, d'autre part, $a' \in A \setminus A'$ ne peut avoir lieu car alors a' appartient à $f(a') = A'$. D'où une contradiction.$\blacksquare$

Lemme 1.5.2.5 (sandwich) *Si les ensembles A, B, C sont tels que $A \subset C \subset B$ et si A et B sont en bijection, alors A et C sont en bijection de même que C et B.*

Preuve. Soit f une bijection de A sur B. Posons $A_0 = f^{-1}(C \setminus A)$ et, pour tout entier $m \geq 1$, $A_m = f^{-1}(A_{m-1})$. Cela étant, considérons l'ensemble $A' = \bigcup_{m=0}^{\infty} A_m$: ses éléments sont ceux de A auxquels correspondent successivement des éléments de A puis un élément de $C \setminus A$. Etablissons que f est une bijection de A' sur $A' \cup (C \setminus A)$. C'est une application de A' dans $A' \cup (C \setminus A)$ car, pour tout $a' \in A'$, il existe un entier $n \in \mathbb{N}$ tel que $a' \in A_n$ ce qui implique $f(a') \in C \setminus A$ si n est égal à 0 et $f(a') \in A_{n-1}$ sinon. Il est alors clair que f est une injection de A' dans $A' \cup (C \setminus A)$. C'est aussi une surjection car, d'une part, pour tout $c \in A'$, il existe $a \in A$ tel que $f(a) = c$ et ainsi a appartient à A' et d'autre part, pour tout élément de $C \setminus A$, il existe bien sûr un élément a de A tel que $f(a) = c$, c'est-à-dire un élément de A_0.

De là, la loi g définie de A dans C par

$$\begin{array}{rcll} g(a) & = & f(a), & \forall a \in A' \\ g(a) & = & a, & \forall a \in A \setminus A', \end{array}$$

est une bijection de $A = A' \cup (A \setminus A')$ sur $C = (A' \cup (C \setminus A)) \cup (A \setminus A')$.

Enfin, comme on a $A \simeq B$ et $A \simeq C$, on a aussi $C \simeq B$.∎

Théorème 1.5.2.6 (Schroeder-Bernstein) *S'il existe une injection de l'ensemble A dans l'ensemble B et une injection de B dans A, alors A et B sont en bijection.*

Preuve. De fait, si f est une injection de A dans B et g une injection de B dans A, f est une bijection entre A et son image B', et g est une bijection entre B et son image A'. Dès lors, $g \circ f$ est une bijection entre A et une partie de A, à savoir $g \circ f(A)$. D'où la conclusion par le lemme du sandwich.∎
$\leftarrow *$

1.5.3 Exemples d'ensembles en bijection

$* \rightarrow$

Proposition 1.5.3.1 *Toute partie A de $\mathbb{R}^n$ d'intérieur non vide est en bijection avec $\mathbb{R}^n$.*

Preuve. D'une part, A est en bijection avec une partie de $\mathbb{R}^n$, par exemple avec A lui-même au moyen de l'application identité. D'autre part, si x_0 est un point intérieur à A, il existe $r > 0$ tel que $b = \{\, x : |x - x_0| < r\}$ soit inclus dans A. Il existe alors une injection de $\mathbb{R}^n$ dans la partie b de A, à savoir par exemple

$$f : x \mapsto x_0 + rx/(1 + |x|),$$

comme on le vérifie aisément. D'où la conclusion par le théorème de Schroeder-Bernstein.∎

Rappel. Rappelons que, mis sous forme décimale, un nombre réel n'admet qu'un seul développement sauf dans les cas où cette forme décimale se termine soit par une suite illimitée de 9, soit par une suite illimitée de 0 auquel cas on a, par exemple, $a{,}a_1 \ldots a_n 99 \ldots = a{,}a_1 \ldots (a_n + 1)$ si $a_n \neq 9$.

Proposition 1.5.3.2 *Quels que soient n, $n' \in \mathbb{N}_0$, les espaces $\mathbb{R}^n$ et $\mathbb{R}^{n'}$ sont en bijection.*

Preuve. Vu la proposition précédente, il suffit d'établir que, pour tout entier $n \geq 1$, l'intervalle ouvert $I =]0,1[\times \cdots \times]0,1[$ de $\mathbb{R}^n$ est en bijection avec l'intervalle ouvert $J =]0,1[$ de $\mathbb{R}$. Cela résulte d'une application du théorème de Schroeder-Bernstein. D'une part, l'application

$$f \colon x \mapsto (x, 1/2, \ldots, 1/2)$$

est évidemment une injection de J dans I. D'autre part, au point $x = (x_1, \ldots, x_n) \in I$ dont les composantes s'écrivent

$$x_j = 0{,}j_1 j_2 j_3 \ldots$$

où, pour éviter toute ambiguïté, on convient d'éviter les suites illimitées de 9 dans le développement décimal, associons le nombre

$$0{,}1_1 2_1 \ldots n_1 1_2 2_2 \ldots n_2 \ldots$$

On voit aisément que cette loi est une bijection entre I et une partie de J.∎

Remarque. Cette dernière bijection est bien une bijection entre I et une partie de J distincte de J et ne donne pas une solution au problème posé car, par exemple, le nombre $0{,}0191919\ldots \in J$ n'est l'image d'aucun élément de $I \subset \mathbb{R}^2$. En effet, il devrait provenir du point $(0{,}0999\ldots; 0{,}111\ldots) = (0{,}1; 0{,}111\ldots)$ dont l'image est $(0{,}1101010\ldots)$.

Proposition 1.5.3.3 *Pour tout $n \in \mathbb{N}_0$, on a $\mathbb{R}^n \simeq \wp(\mathbb{N}_0)$.*

Preuve. Vu les propositions précédentes, nous avons $\mathbb{R}^n \simeq]0,1[$. On conclut alors par le théorème de Schroeder-Bernstein. D'une part,

$$f \colon 0{,}x_1 x_2 x_3 \ldots \mapsto \{1x_1, 1x_1x_2, 1x_1x_2x_3, \ldots\}$$

est une bijection entre $]0,1[$ et une partie de $\wp(\mathbb{N}_0)$, comme on le vérifie aisément. D'autre part,

$$g \colon A \mapsto 0{,}a_1 a_2 a_3 \ldots,$$

où $a_j = 0$ si $j \notin A$ et $a_j = 1$ si $j \in A$ est une bijection entre $\wp(\mathbb{N}_0)$ et une partie de $]0,1[$, comme on le vérife de suite.∎

Proposition 1.5.3.4 *Pour tout $n \in \mathbb{N}_0$, l'ensemble des intervalles de $\mathbb{R}^n$ est en bijection avec $\mathbb{R}^n$.*

Preuve. Vu les propositions précédentes, il suffit d'établir que l'ensemble des intervalles de $\mathbb{R}^n$ est en bijection avec $]0,1[$. Cela résulte du théorème de Schroeder-Bernstein. D'une part, tout intervalle de $\mathbb{R}^n$ peut être caractérisé par $2n$ nombres ordonnés et $2n$ symboles ordonnés appartenant à $\{],[\}$, donc un point de $\mathbb{R}^{4n}$ si on convient, par exemple de remplacer $]$ par -1 et $[$ par 1. Dès lors, par les propositions précédentes, l'ensemble des intervalles de $\mathbb{R}^n$ est en bijection avec une partie de $]0,1[$. D'autre part,

$$f: x \mapsto\ ]0,x[\ \times \cdots \times\ ]0,x[$$

est visiblement une bijection entre $]0,1[$ et une partie de l'ensemble des intervalles de $\mathbb{R}^n$.∎

Proposition 1.5.3.5 *Pour tout $n \in \mathbb{N}_0$, l'ensemble des suites de $\mathbb{R}^n$ est en bijection avec $\mathbb{R}^n$.*

Preuve. Vu les deux premières propositions de ce paragraphe, on est amené à établir que l'ensemble des suites de $]0,1[$ est en bijection avec $]0,1[$. D'une part, $x \mapsto x, x, x, \ldots$ est une injection de $]0,1[$ dans l'ensemble des suites de $]0,1[$. D'autre part, à la suite $(x_m)_{m\in\mathbb{N}_0}$ de $]0,1[$ dont les éléments s'écrivent

$$\begin{array}{lcl} x_1 & = & 0{,}a_1a_2a_3\ldots \\ x_2 & = & 0{,}b_1b_2b_3\ldots \\ x_3 & = & 0{,}c_1c_2c_3\ldots \\ \ldots & & \end{array}$$

où, pour éviter toute ambiguïté, nous convenons d'éliminer les suites illimitées de 9, associons le nombre

$$0{,}a_1a_2b_1a_3b_2c_1\ldots$$

On voit aisément que cette loi est une bijection entre l'ensemble des suites de $]0,1[$ et une partie de $]0,1[$. D'où la conclusion par le théorème de Schroeder-Bernstein.∎

Proposition 1.5.3.6 *Pour tout $n \in \mathbb{N}_0$, l'ensemble $\mathrm{F}(\mathbb{R}^n)$ des fonctions définies sur $\mathbb{R}^n$ est en bijection avec $\wp(\mathbb{R}^n)$.*

Preuve. D'une part, toute fonction définie sur $\mathbb{R}^n$ est caractérisée par son graphe, qui est une partie de $\mathbb{R}^{n+2}$, or $\wp(\mathbb{R}^{n+2})$ est en bijection avec $\wp(\mathbb{R}^n)$. D'autre part, toute partie de $\mathbb{R}^n$ est caractérisée par sa fonction caractéristique, c'est-à-dire par une fonction sur $\mathbb{R}^n$.∎
$\leftarrow *$

1.5.4 Ensembles dénombrables

$\to *$

Définitions. Un ensemble A est *dénombrable* s'il est en bijection avec une partie de $\mathbb{N}_0$, sinon il est dit *non dénombrable.*

Numéroter les éléments d'un ensemble A, c'est donner une bijection d'un ensemble du type $\{ m \in \mathbb{N}_0 : m \leq n\}$ avec $n \in \mathbb{N}$ ou de $\mathbb{N}_0$ sur A; on ne peut donc numéroter les éléments d'un ensemble que si et seulement si celui-ci est dénombrable.

Pour établir qu'un ensemble est dénombrable, on fournit d'habitude un procédé de numérotation de ses éléments. Pour établir qu'un ensemble n'est pas dénombrable, on procède généralement par l'absurde.

L'existence d'ensembles non dénombrables semble a priori paradoxale car on imagine assez aisément qu'il est possible de "numéroter successivement tous ses éléments".

Cependant $\wp(\mathbb{N})$ n'est pas dénombrable car $\mathbb{N}$ n'est pas en bijection avec $\wp(\mathbb{N})$.

Voici certainement l'exemple le plus fameux d'ensemble non dénombrable.

Proposition 1.5.4.1 *L'intervalle* $]0,1[$ *n'est pas dénombrable.*

Preuve. Procédons par l'absurde. On sait que $]0,1[$ ne peut être en bijection avec une partie finie de $\mathbb{N}_0$. Supposons que f soit une bijection de $\mathbb{N}_0$ sur $]0,1[$. Le nombre $0{,}x_1x_2x_3\ldots$ où x_m diffère de la m-ème décimale de $f(m)$ et de 9 (si on convient de représenter tout élément de $]0,1[$ au moyen de son développement décimal où on rejette les formes qui contiennent une suite illimitée de 9), appartient à $]0,1[$ et diffère cependant de $f(m)$ quel que soit m. D'où une contradiction avec la surjectivité de f.

En voici une démonstration transcendante. Il suffit d'établir que $\mathbb{N}_0$ n'est pas en bijection avec $]0,1[$. Or nous savons que $]0,1[$ et $\mathbb{R}$ sont en bijection, de même que $\mathbb{R}$ et $\wp(\mathbb{N}_0)$. Les ensembles $]0,1[$ et $\wp(\mathbb{N}_0)$ sont donc en bijection. La conclusion provient alors de ce que A et $\wp(A)$ ne sont jamais en bijection, quel que soit l'ensemble A. ∎

Voici un critère très important de dénombrabilité.

Critère 1.5.4.2 *Si les éléments d'un ensemble A peuvent être caractérisés par un nombre fini d'indices* (mais ce nombre d'indices peut varier d'un élément à l'autre) *prenant leurs valeurs dans des ensembles dénombrables, alors A est dénombrable.*

Preuve. Soit $a(n_1, \ldots, n_N)$ la représentation annoncée des éléments de A où nous pouvons supposer que chaque indice n_j varie dans une partie de $\mathbb{N}_0$.

Pour tout n, soit A_n l'ensemble des éléments de A dont la somme des indices dans la représentation est égale à n. Bien sûr, chaque A_n est fini et peut donc être numéroté.

On peut alors numéroter les éléments de A de la façon suivante: on numérote les éléments de A_1, puis ceux de A_2 et ainsi de suite.∎

Proposition 1.5.4.3 a) *Toute partie d'un ensemble dénombrable est dénombrable.*

b) *Deux ensembles en bijection sont simultanément dénombrables ou non dénombrables.*

c) *Toute union dénombrable d'ensembles dénombrables est dénombrable.*

Preuve. a) et b) sont immédiats.

c) Si, pour tout $m \in \mathbb{N}_0$, l'ensemble A_m est dénombrable, on peut numéroter ses éléments et obtenir

$$A_m = \left\{ x_i^{(m)} : i \in I_m \right\},$$

avec $I_m \subset \mathbb{N}_0$. Mais alors, $\cup_{m=1}^{\infty} A_m$ est en bijection avec une partie de l'ensemble $\{ (m, i) : m, i \in \mathbb{N}_0 \}$, qui est dénombrable par application du critère. D'où la conclusion.∎

Voici également un important critère de non dénombrabilité; on l'obtient de suite en effectuant un raisonnement analogue à celui de la preuve du fait que $]0, 1[$ est non dénombrable.

Critère 1.5.4.4 *Si les éléments d'un ensemble A peuvent être caractérisés par une infinité d'indices prenant au moins deux valeurs* (un indice qui ne peut prendre qu'une seule valeur peut évidemment être ignoré: il ne joue aucun rôle)*, alors A est non dénombrable.*∎

Proposition 1.5.4.5 *L'ensemble des points rationnels de $\mathbb{R}^n$ est dénombrable quel que soit $n \in \mathbb{N}_0$.*

Preuve. Rappelons qu'un point de $\mathbb{R}^n$ est rationnel si chacune de ses composantes est un nombre rationnel. Vu le critère de dénombrabilité, il suffit donc d'établir que l'ensemble des points rationnels de $\mathbb{R}$ est dénombrable. Or un nombre réel est rationnel si et seulement s'il peut s'écrire sous la forme $(-1)^i \, r/s$ avec $i \in \{1, 2\}$, $r \in \mathbb{N}$ et $s \in \mathbb{N}_0$.∎

Proposition 1.5.4.6 a) *L'ensemble des polynômes à coefficients dans* $\mathbb{Z}$ *est dénombrable. De là, l'ensemble des nombres algébriques* (c'est-à-dire des zéros de ces polynômes) *est dénombrable.*

b) *L'ensemble des nombres transcendants est non dénombrable.*

Preuve. a) est une application immédiate du critère de dénombrabilité.

b) Un nombre transcendant étant un nombre réel qui n'est pas algébrique, la conclusion provient de ce que $\mathbb{R}$ n'est pas dénombrable.∎

$\leftarrow *$

Chapitre 2

L'espace euclidien $\mathbb{R}^n$

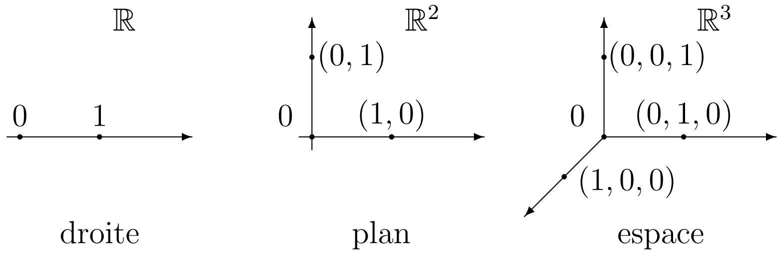
2.1 Introduction

Pour $n = 1$, 2 ou 3, il existe des représentations géométriques bien connues et commodes de l'espace euclidien $\mathbb{R}^n$, à savoir successivement

a) la droite orientée munie d'une origine et d'une unité, qui est une représentation géométrique de $\mathbb{R}$, l'ensemble des nombres réels,

b) le plan muni d'un système d'axes perpendiculaires avec unité, qui est une représentation géométrique de $\mathbb{R}^2$ mais aussi de $\mathbb{C}$, l'ensemble des nombres complexes,

c) l'espace physique rapporté à un système d'axes dextrorsum avec unité, qui est une représentation géométrique de $\mathbb{R}^3$.



Cependant les besoins de l'analyse mathématique et de ses applications à la physique, la mécanique, la chimie, ... montrent qu'il est indispensable de sortir du cadre de ces trois exemples. Ainsi la position d'un point matériel au repos dans l'espace physique nécessite trois coordonnées en mécanique classique et dès lors $\mathbb{R}^3$ convient; mais s'il est en mouvement, on introduit une quatrième coordonnée: le temps. D'autres exemples, très nombreux, seront introduits dans les autres cours.

Dans ce chapitre, notre but est de définir et étudier un espace $\mathbb{R}^n$ quel que soit l'entier $n \in \mathbb{N}_0$. Bien sûr, pour $n > 3$, il n'y a plus de représentation graphique (mais déjà la représentation graphique de $\mathbb{R}^3$ sur un plan présente quelques difficultés). Dès lors, comment allons-nous procéder? La réponse est la suivante: *nous allons définir* $\mathbb{R}^n$ *et étudier cet espace par analogie avec ce qui se passe pour* $\mathbb{R}$, $\mathbb{R}^2$ *et* $\mathbb{R}^3$. Il s'ensuit que la définition de $\mathbb{R}^n$ et l'introduction des différentes notions que nous allons étudier dans $\mathbb{R}^n$ doivent se faire de manière purement analytique, tout en étant une généralisation naturelle de ce qui se passe dans $\mathbb{R}$, $\mathbb{R}^2$ et $\mathbb{R}^3$.

Insistons immédiatement sur le fait suivant. Les représentations graphiques de $\mathbb{R}$, $\mathbb{R}^2$ et $\mathbb{R}^3$ sont très utiles non seulement pour visualiser ce qui se passe dans ces espaces, mais encore pour suggérer comment procéder dans $\mathbb{R}^n$. Cependant un tel dessin ne constitue jamais une preuve.

2.2 Espace $\mathbb{R}^n$

2.2.1 Définition

Définitions. Etant donné $n \in \mathbb{N}_0$, on appelle *point de* $\mathbb{R}^n$ tout ensemble ordonné de n nombres réels. Si $r_1, \ldots, r_n$ sont ces n nombres réels, on désigne le point correspondant par

$$(r_1, \ldots, r_n)$$

ou même par un symbole unique tel que x si aucune ambiguïté sur les r_j n'est possible, r_j étant appelé la *j-ème composante de x* et noté

$$[x]_j$$

ou même x_j si aucune ambiguïté n'est possible.

Quelques point particuliers de $\mathbb{R}^n$ reçoivent une dénomination et une notation spéciales: d'une part

$$0 = (0, \ldots, 0)$$

est appelé *origine de* $\mathbb{R}^n$ et, d'autre part, pour tout $j \in \{1, \ldots, n\}$,

$$e_j = (\underbrace{0, \ldots, 0, 1}_{j}, 0, \ldots, 0)$$

est appelé *j-ème point unité de* $\mathbb{R}^n$. Il y a donc n points unité dans $\mathbb{R}^n$.

Définitions. *L'espace euclidien* $\mathbb{R}^n$ de *dimension* n est l'ensemble des points de $\mathbb{R}^n$ muni des notions d'égalité, de combinaison linéaire et de produit scalaire que nous allons introduire. En particulier, l'espace euclidien de dimension 1 est noté $\mathbb{R}$ et non $\mathbb{R}^1$, notation qui va se justifier d'elle-même tout aussitôt.

a) Egalité de deux points de $\mathbb{R}^n$. Deux points x, y de $\mathbb{R}^n$ sont *égaux* si, pour tout $j \in \{1, \ldots, n\}$, on a $x_j = y_j$.

Par exemple, les points unité e_j et e_k de $\mathbb{R}^n$ sont égaux si et seulement si j est égal à k.

Remarque. Pour définir l'égalité entre points de $\mathbb{R}^n$, nous nous sommes ramenés à n égalités dans $\mathbb{R}$. Insistons sur le fait que $\mathbb{R}$ a une structure beaucoup plus riche car on y introduit aussi les notions de comparaison suivantes:
$\leq$ (que nous lisons "*inférieur ou égal à*"),
$<$ (que nous lisons "*strictement inférieur à*"),
$\geq$ (que nous lisons "*supérieur ou égal à*"),
$>$ (que nous lisons "*strictement supérieur à*").

En ce qui concerne ces notions de comparaison, l'espace $\mathbb{R}$ devra toujours être considéré séparément.

b) Combinaison linéaire dans $\mathbb{R}^n$. Etant donné x, $y \in \mathbb{R}^n$ et $r \in \mathbb{R}$, on introduit les points rx et $x + y$ de $\mathbb{R}^n$ par

$$[rx]_j = r[x]_j \quad \text{et} \quad [x + y]_j = [x]_j + [y]_j, \quad \forall j \in \{1, \ldots, n\}.$$

Plus généralement, si K est un nombre entier supérieur ou égal à 1, si x_1, …, x_K sont des points de $\mathbb{R}^n$ et si r_1, … r_K sont des nombres réels, la *combinaison linéaire* correspondante est le point de $\mathbb{R}^n$ dont la j-ème composante est donnée par

$$r_1 \left[x_1\right]_j + \cdots + r_K \left[x_K\right]_j = \sum_{k=1}^{K} r_k \left[x_k\right]_j ,$$

c'est-à-dire par la combinaison linéaire correspondante des j-èmes composantes des points. Elle est notée

$$r_1 x_1 + \cdots + r_K x_K \quad \text{ou} \quad \sum_{k=1}^{K} r_k x_k.$$

Les nombres r_1, …, r_K sont appelés les *coefficients* de cette combinaison linéaire.

Remarque. Pour définir la notion de combinaison linéaire dans $\mathbb{R}^n$, nous nous sommes ramenés à n combinaisons linéaires dans $\mathbb{R}$. De la sorte, si nous désirons identifier $\mathbb{C}$ et $\mathbb{R}^2$, remarquons bien que nous n'avons introduit jusqu'à présent que les combinaisons linéaires à coefficients réels dans $\mathbb{C}$. Rappelons que si

$$z = (\Re z, \Im z) = \Re z + i \Im z \quad \text{et} \quad z' = (\Re z', \Im z') = \Re z' + i \Im z'$$

sont des nombres complexes, leur *produit* est le nombre complexe zz' défini par

$$\begin{aligned} zz' &= (\Re z \Re z' - \Im z \Im z', \Re z \Im z' + \Im z \Re z') \\ &= \Re z \Re z' - \Im z \Im z' + i(\Re z \Im z' + \Im z \Re z'). \end{aligned}$$

Cela étant, si $z_1, \ldots z_K$ et $c_1, \ldots c_K$ sont des nombres complexes, la *combinaison linéaire* (à coefficients complexes) correspondante est bien sûr le nombre complexe

$$c_1 z_1 + \cdots + c_K z_K = \sum_{k=1}^{K} c_k z_k.$$

Il s'agit là d'une opération spécifique à $\mathbb{C}$, qui doit être étudiée séparément.

On peut aussi introduire la notion de combinaison linéaire de parties de $\mathbb{R}^n$: si K est un nombre entier supérieur ou égal à 1, si $A_1, \ldots, A_K$ sont des parties non vides de $\mathbb{R}^n$ et si $r_1, \ldots, r_K$ sont des nombres réels, la *combinaison linéaire* correspondante est l'ensemble

$$\sum_{k=1}^{K} r_k A_k = \left\{ \sum_{k=1}^{K} r_k x_k : x_k \in A_k, \forall k \in \{1, \ldots, K\} \right\},$$

noté également $r_1 A_1 + \cdots + r_K A_K$.

En particulier, si A est une partie non vide de $\mathbb{R}^n$,

a) un *translaté de* A est une partie B de $\mathbb{R}^n$ pour laquelle il existe un point x de $\mathbb{R}^n$ tel que $B = \{x\} + A$, ce qu'on note plutôt $B = x + A$.

b) un *homothétique de* A est une partie B de $\mathbb{R}^n$ pour laquelle il existe $r \in \mathbb{R}$ tel que $B = rA$.

c) Produit scalaire dans $\mathbb{R}^n$. *Le produit scalaire des points x, y de* $\mathbb{R}^n$ est le nombre réel

$$\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n = \sum_{j=1}^{n} x_j y_j.$$

Par exemple, pour tout $x \in \mathbb{R}^n$, on a

$$[x]_j = \langle x, e_j \rangle, \quad \forall j \in \{1, \ldots, n\},$$

et, par conséquent,

$$x = \sum_{j=1}^{n} \langle x, e_j \rangle \, e_j.$$

Remarque. Pour définir le produit scalaire dans $\mathbb{R}^n$, nous nous sommes ramenés à une somme de produits dans $\mathbb{R}$.

Dans $\mathbb{C}$, cette notion correspond à une notion peu utilisée.

Dans $\mathbb{R}$ et dans $\mathbb{C}$, on introduit aussi la *division* par tout nombre non nul. Il s'agit là d'une opération spécifique à $\mathbb{R}$ et $\mathbb{C}$, qu'il faudra étudier séparément.

Voici quelques propriétés fondamentales du produit scalaire.

Proposition 2.2.1.1 *Pour tous x, y, $z \in \mathbb{R}^n$ et tout $r \in \mathbb{R}$, on a*

a) $\langle x, x\rangle \geq 0$. *De plus,* $\langle x, x\rangle = 0$ *a lieu si et seulement si* $x = 0$.

b) $\langle x, y\rangle = \langle y, x\rangle$, (c'est-à-dire que le produit scalaire est commutatif),

c) $\langle rx, y\rangle = r\,\langle x, y\rangle$,

d) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$.∎

Remarque. Pour toutes combinaisons linéaires $\sum_{j=1}^{J} r_j x_j$ et $\sum_{k=1}^{K} s_k y_k$ de points de $\mathbb{R}^n$, on déduit aussitôt de c) et d) la formule suivante

$$\left\langle \sum_{j=1}^{J} r_j x_j, \sum_{k=1}^{K} s_k y_k \right\rangle = \sum_{j=1}^{J}\sum_{k=1}^{K} r_j s_k \langle x_j, y_k\rangle$$

dans laquelle le choix du premier indice sommatoire est arbitraire, celui du second étant libre également à condition de le prendre différent du premier.

2.2.2 Module d'un point

Définition. *Le module d'un point x de $\mathbb{R}^n$* est le nombre positif ou nul

$$|x| = \sqrt{\sum_{j=1}^{n} x_j^2} = \sqrt{\langle x, x\rangle}.$$

On a donc notamment $|0| = 0$ et $|e_j| = 1$ pour tout $j \in \{1, \ldots, n\}$.

Remarque. Pour $n = 1$ et $n = 2$, on réobtient bien les notions de module introduites dans $\mathbb{R}$ et dans $\mathbb{C}$ respectivement, à savoir

$$|x| = \left\{ \begin{array}{rcl} x & \text{si} & x \geq 0 \\ -x & \text{si} & x < 0 \end{array} \right\}, \quad \forall x \in \mathbb{R},$$

et

$$|z| = \sqrt{(\Re z)^2 + (\Im z)^2}, \quad \forall z \in \mathbb{C}.$$

Passons aux propriétés du module.

Théorème 2.2.2.1 *Pour tous points x, $y \in \mathbb{R}^n$, on a*

a) $|x| = 0 \Leftrightarrow x = 0$,

b) $|rx| = |r|\,|x|\,, \quad \forall r \in \mathbb{R}$,

c) *l'***inégalité de Cauchy-Schwarz:** $|\langle x, y\rangle| \leq |x|\,|y|$.

De plus, l'égalité $|\langle x, y\rangle| = |x|\,|y|$ a lieu si et seulement si on a $x = 0$ ou $y = sx$ avec $s \in \mathbb{R}$.

Preuve. a) et b) sont triviaux.

c) Si x est égal à 0, c'est trivial. Sinon, pour tout $r \in \mathbb{R}$, il vient

$$\begin{aligned} 0 \leq |rx + y|^2 &= \langle rx + y, rx + y\rangle \\ &= \langle x, x\rangle\, r^2 + \langle x, y\rangle\, r + \langle y, x\rangle\, r + \langle y, y\rangle \\ &= |x|^2\, r^2 + 2\,\langle x, y\rangle\, r + |y|^2\,. \end{aligned}$$

Comme $P(r) = |x|^2\, r^2 + 2\,\langle x, y\rangle\, r + |y|^2$ est un trinôme du second degré en la variable r dont le signe est constant, son réalisant est négatif: on a donc $\langle x, y\rangle^2 - |x|^2\,|y|^2 \leq 0$. La conclusion résulte alors aussitôt d'un passage aux racines carrées dans l'inéquation $\langle x, y\rangle^2 \leq |x|^2\,|y|^2$.

Passons à l'égalité. Si on a $x = 0$ ou $y = sx$ avec $s \in \mathbb{R}$, on vérifie de suite que l'égalité a lieu. Inversement, si l'égalité a lieu et si x diffère de 0, $P(r)$ est un trinôme du second degré qui admet un zéro double r_0 car son réalisant est nul. On a alors $|r_0 x + y| = 0$. D'où la conclusion.∎

Remarque. a) Pour tous x, $y \in \mathbb{R}$, on a $|\langle x, y\rangle| = |xy| = |x|\,|y|$.

b) Pour tous z, $z' \in \mathbb{C}$, on a $|zz'| = |z|\,|z'|$.

Voici également d'autres propriétés fort importantes du module.

Proposition 2.2.2.2 *Pour tous x, $y \in \mathbb{R}^n$, on a*

a) $|x_j| \leq |x| \leq \sum_{k=1}^{n} |x_k|\,, \quad \forall j \leq n$,

b) *l'***inégalité de Minkowski:** $|x + y| \leq |x| + |y|$.

Plus généralement, si $x^{(1)}$, ..., $x^{(J)}$ sont des points de $\mathbb{R}^n$ et r_1, ..., r_J des nombres réels en nombre fini, on a

$$\left|\sum_{j=1}^{J} r_j x^{(j)}\right| \leq \sum_{j=1}^{J} |r_j|\,\left|x^{(j)}\right|.$$

De plus, l'égalité $|x + y| = |x| + |y|$ a lieu si et seulement si on a $x = 0$ ou $y = rx$ avec $r \geq 0$.

c) $||x| - |y|| \leq |x - y|$. *De plus l'égalité $||x| - |y|| = |x - y|$ a lieu si et seulement si on a $x = 0$ ou $y = rx$ avec $r \geq 0$.*

Preuve. a) est immédiat.

b) L'inégalité $|x+y| \leq |x| + |y|$ s'obtient directement par passage aux racines carrées dans

$$\begin{aligned} |x+y|^2 &= \langle x+y, x+y \rangle = |x|^2 + 2\langle x, y\rangle + |y|^2 \\ &\leq |x|^2 + 2|\langle x, y\rangle| + |y|^2 \\ &\leq |x|^2 + 2|x|\,|y| + |y|^2 = (|x| + |y|)^2. \end{aligned}$$

Le cas général s'obtient aisément par récurrence.

De plus, l'égalité $|x+y| = |x| + |y|$ a lieu si et seulement si $\langle x, y\rangle$ est ≥ 0 et égal à $|x|\,|y|$, donc si et seulement si les deux conditions suivantes sont vérifiées
(i) $\langle x, y\rangle \geq 0$,
(ii) $x = 0$ ou $y = rx$ avec $r \in \mathbb{R}$.
La conclusion est alors immédiate.

c) De fait, de $x = x - y + y$, on tire de suite $|x| \leq |x-y| + |y|$ donc $|x| - |y| \leq |x-y|$. En permutant les rôles de x et de y, on obtient de la même manière $|y| - |x| \leq |y-x| = |x-y|$. D'où la conclusion en ce qui concerne l'inégalité.

Passons à l'égalité. Elle a lieu si et seulement si l'une des égalités

$$|x| - |y| = |x-y| \quad \text{ou} \quad |y| - |x| = |x-y|$$

a lieu. La première de ces égalités a lieu si et seulement si

$$|x| = |(x-y) + y| = |x-y| + |y|,$$

c'est-à-dire si et seulement si on a $x = y$ ou $y = r(x-y)$ avec $r \geq 0$, ce qui revient à la condition $y = rx$ avec $0 \leq r \leq 1$. Dès lors, la seconde égalité revient à la condition $x = 0$ ou $y = rx$ avec $1 \leq r$. D'où la conclusion.∎

2.2.3 Distance euclidienne

Définition. Une *distance* sur un ensemble non vide E est une loi d qui, à tout couple ordonné x, y de points de E associe un nombre réel $d(x, y)$, qui jouit des propriétés suivantes:
a) $d(x, y) \geq 0$,
b) $d(x, y) = 0$ si et seulement si $x = y$,
c) $d(x, y) = d(y, x)$,
d) $d(x, y) \leq d(x, z) + d(z, y)$, propriété connue sous le nom d'**inégalité triangulaire**.

Un ensemble non vide E muni d'une distance d est un *espace métrique;* il est noté (E, d).

Proposition 2.2.3.1 *La loi* d *qui, à tout couple ordonné* x, $y \in \mathbb{R}^n$, *associe le nombre* $\mathrm{d}(x, y) = |x - y|$ *est une distance sur* $\mathbb{R}^n$.

Preuve. Cela résulte aussitôt des propriétés du module.∎

Définition. Cette distance d sur $\mathbb{R}^n$ est appelée *distance euclidienne.*

Voici tout d'abord quelques propriétés générales vérifiées par tout espace métrique donc, en particulier, par $\mathbb{R}^n$ muni de la distance euclidienne.

Proposition 2.2.3.2 *Pour tous points* x, y, z, t *de l'espace métrique* (E, d), *on a*

$$|d(x, z) - d(z, y)| \leq d(x, y)$$

et

$$|d(x, y) - d(z, t)| \leq d(x, z) + d(y, t).$$

Preuve. La première inégalité résulte aussitôt des inégalités

$$d(x, z) - d(z, y) \leq d(x, y) \quad \text{et} \quad d(y, z) - d(z, x) \leq d(y, x)$$

qu'on obtient directement des relations c) et d) de la définition d'une distance. La seconde inégalité résulte aussitôt de

$$d(x, y) - d(z, t) \leq d(x, z) + d(z, t) + d(t, y) - d(z, t)$$

et

$$d(z, t) - d(x, y) \leq d(z, x) + d(x, y) + d(y, t) - d(x, y).$$∎

Voici ensuite quelques propriétés particulières relatives à la distance euclidienne.

Proposition 2.2.3.3 *Pour tous points* x, y, z *de* $\mathbb{R}^n$ *et tout nombre réel* r, *on a*

a) $\mathrm{d}(rx, ry) = |r| \, \mathrm{d}(x, y)$,
b) $\mathrm{d}(x + y, y + z) = \mathrm{d}(x, z)$,
c) $\mathrm{d}(|x|, |y|) \leq \mathrm{d}(x, y)$,
d) $\mathrm{d}(x_j, y_j) \leq \mathrm{d}(x, y)$ *pour tout* $j \in \{1, \ldots, n\}$.

Preuve. Cela résulte aussitôt des propriétés du module.∎

Dans $\mathbb{C}$, on a en outre la propriété suivante où $\overline{z} = \Re z - i\Im z$ désigne le nombre complexe conjugué de $z = \Re z + i\Im z \in \mathbb{C}$.

Proposition 2.2.3.4 *Pour tous z, $z' \in \mathbb{C}$, on a* $\mathrm{d}(\overline{z}, \overline{z'}) = \mathrm{d}(z, z')$.∎

Enfin, en guise de complément, signalons les deux résultats suivants.

Exercice. *Etant donné des points x, y, z de $\mathbb{R}^n$, établir que l'égalité*

$$d(x,y) = d(x,z) + d(z,y)$$

a lieu si et seulement s'il existe un nombre réel r tel que $0 \leq r \leq 1$ pour lequel $z = rx + (1-r)y$.

Suggestion. Pour $x' = x - z$ et $y' = z - y$, cette égalité s'écrit $|x' + y'| = |x'| + |y'|$ et a donc lieu si et seulement si on a $x = z$ ou $z - y = r(x - z)$ avec $r \geq 0$, ce qui permet de conclure aussitôt.∎

Exercice. *Etant donné des points x, y, z de $\mathbb{R}^n$, établir que l'égalité*

$$|\mathrm{d}(x,z) - \mathrm{d}(z,y)| = \mathrm{d}(x,y)$$

a lieu si et seulement si l'une des deux conditions suivantes est vérifiée
a) $x = y$,
b) *il existe un nombre réel $r \leq 0$ ou $r \geq 1$ tel que $z = rx + (1-r)y$.*

Suggestion. Cette égalité a lieu si et seulement si une des égalités suivantes a lieu $\mathrm{d}(x,z) = \mathrm{d}(x,y) + \mathrm{d}(y,z)$, $\mathrm{d}(z,y) = \mathrm{d}(z,x) + \mathrm{d}(x,y)$. Or la première a lieu si et seulement si y s'écrit $y = rx + (1-r)z$ avec $0 \leq r \leq 1$, c'est-à-dire si et seulement si l'une des deux conditions suivantes est vérifiée: (a) $x = y$, (b) $z = rx + (1-r)y$ avec $r \leq 0$. De la même manière, la deuxième égalité a lieu si et seulement si l'une des deux égalités suivantes est vérifiée: (a') $x = y$, (b') $z = rx + (1-r)y$ avec $r \geq 1$. D'où la conclusion.

2.2.4 Intervalles et boules

Introduisons tout d'abord quelques parties remarquables de $\mathbb{R}$.

Notations. Etant donné des nombres réels a, b tels que $a < b$, on introduit les notations suivantes:

$$\begin{array}{rcl}
]a,b[& = & \{x \in \mathbb{R} : a < x < b\}\,, \quad]a,+\infty[\;=\; \{x \in \mathbb{R} : x > a\}\,, \\
[a,b] & = & \{x \in \mathbb{R} : a \leq x \leq b\}\,, \quad [a,+\infty[\;=\; \{x \in \mathbb{R} : x \geq a\}\,, \\
]a,b] & = & \{x \in \mathbb{R} : a < x \leq b\}\,, \quad]-\infty,b[\;=\; \{x \in \mathbb{R} : x < b\}\,, \\
[a,b[& = & \{x \in \mathbb{R} : a \leq x < b\}\,, \quad]-\infty,b] \;=\; \{x \in \mathbb{R} : x \leq b\}\,, \\
]-\infty,+\infty[& = & \mathbb{R}
\end{array}$$

Définitions. Les *intervalles de* $\mathbb{R}$ sont les parties de $\mathbb{R}$ qui s'écrivent selon une des notations que nous venons d'introduire. Si l'intervalle se présente sous l'une des formes

$$]a,b[\,,[a,b]\,,]a,b]\,,[a,b[\,,]a,+\infty[\,,[a,+\infty[$$
$$(\text{resp. }]a,b[\,,[a,b]\,,]a,b]\,,[a,b[\,,]-\infty,b[\,,]-\infty,b]),$$

on dit que a est son *origine* (resp. que b est son *extrémité*).

Si un intervalle de $\mathbb{R}$ a une origine a et une extrémité b—il s'écrit donc $]a,b[$, $[a,b]$, $]a,b]$ ou $[a,b[$—, sa *longueur* est égale à $b-a$.

Il y a deux manières de généraliser ces concepts à $\mathbb{R}^n$.

Définitions. Un *intervalle de* $\mathbb{R}^n$ est un ensemble qui s'écrit

$$I = I_1 \times \cdots \times I_n = \prod_{j=1}^{n} I_j,$$

où I_1, ..., I_n sont des intervalles de $\mathbb{R}$, I_j étant appelé le *j-ème intervalle constitutif de I*.

Un *cube de* $\mathbb{R}^n$—on dit *carré de* $\mathbb{R}^2$—est un intervalle dont tous les intervalles constitutifs admettent une origine et une extrémité, et ont même longueur, ce nombre étant appelé *côté* du cube.

Définitions. Soient a un point de $\mathbb{R}^n$ et r un nombre réel strictement positif. On qualifie de

a) *boule* de *centre* a et de *rayon* r, les ensembles

$$\{\, x \in \mathbb{R}^n : |x-a| < r \} \quad \text{et} \quad \{\, x \in \mathbb{R}^n : |x-a| \leq r \}\,.$$

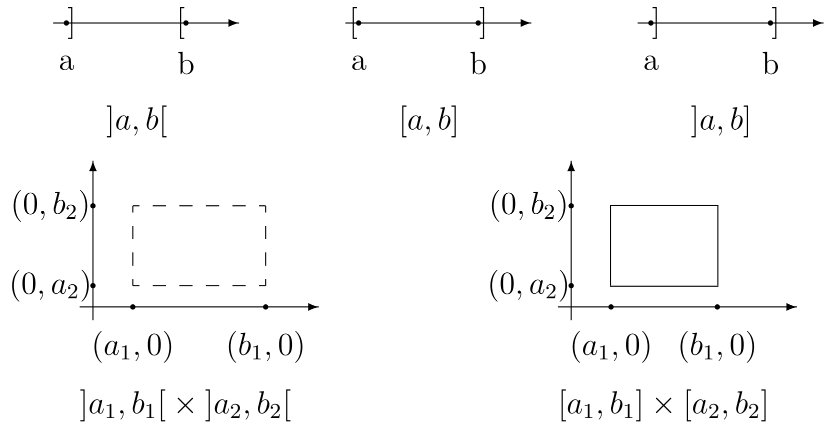
(Bientôt nous distinguerons ces deux ensembles en désignant le premier par l'expression de *boule ouverte* et le second par *boule fermée*.)

b) *sphère* de *centre* a et de *rayon* r, l'ensemble $\{\, x \in \mathbb{R}^n : |x-a| = r \}$.

Dans $\mathbb{R}^2$, on utilise plutôt les mots *cercle* ou *disque* en lieu et place de boule et le mot *circonférence* à la place de sphère.

Si n est égal à 1, on retrouve les intervalles $\{\, x \in \mathbb{R} : |x-a| < r \} = \,]a-r, a+r[$ et $\{\, x \in \mathbb{R} : |x-a| \leq r \} = [a-r, a+r]$, ainsi que les ensembles réduits à deux points distincts car on a $\{\, x \in \mathbb{R} : |x-a| = r \} = \{a-r, a+r\}$.

Signalons quelques conventions adoptées dans la représentation graphique de ces ensembles.



2.2.5 Réseaux finis et quadrillages décimaux

Définitions. Un *semi-intervalle de* $\mathbb{R}$ est un intervalle de $\mathbb{R}$ qui s'écrit sous l'une des formes

$$]a, b]\,,]-\infty, b]\,,]a, +\infty[\,,]-\infty, +\infty[\,.$$

Un *semi-intervalle de* $\mathbb{R}^n$ est un intervalle de $\mathbb{R}^n$ dont les intervalles constitutifs sont chacun des semi-intervalles de $\mathbb{R}$.

Soient $J(1), \ldots, J(n)$ des éléments de $\mathbb{N}_0$ et, pour tous $j \in \{1, \ldots, n\}$ et $k \in \{1, \ldots, J(j) - 1\}$, soit $r_{j,k}$ un nombre réel tel que

$$\left\{ \begin{array}{c} -\infty < r_{1,1} < \ldots < r_{1,J(1)-1} < +\infty \\ \ldots \\ -\infty < r_{n,1} < \ldots < r_{n,J(n)-1} < +\infty \end{array} \right. \qquad (*)$$

Posons $r_{1,0} = -\infty$, $r_{1,J(1)} = +\infty$, $\ldots$, $r_{n,0} = -\infty$ et $r_{n,J(n)} = +\infty$. On vérifie alors de suite que les semi-intervalles $I = I_1 \times \cdots \times I_n$ de $\mathbb{R}^n$ où, pour tout $j \in \{1, \ldots, n\}$, l'origine et l'extrémité de I_j sont deux symboles consécutifs pris parmi $r_{j,0}, \ldots, r_{j,J(j)}$, constituent une partition de $\mathbb{R}^n$. Une telle partition est appelée *réseau fini de* $\mathbb{R}^n$ et ses éléments *mailles* de ce réseau.

Insistons sur le fait que nous n'avons pas exclu le cas $-\infty < +\infty$ dans les inégalités reprises en (*).

Définitions. Soit m un élément de $\mathbb{N}$. Le *quadrillage décimal d'équidistance* 10^{-m} *de* $\mathbb{R}^n$ est l'ensemble des semi-intervalles $I = I_1 \times \cdots \times I_n$ de $\mathbb{R}^n$ tels que chacun des semi-intervalles constitutifs s'écrive sous la forme $]k.10^{-m}, (k+1).10^{-m}]$ avec $k \in \mathbb{Z}$.

On vérifie de suite qu'il s'agit d'une partition de $\mathbb{R}^n$, dont les éléments sont des semi-cubes de côté 10^{-m}, appelés *mailles* du quadrillage décimal d'équidistance 10^{-m}.

2.2.6 Compléments sur les nombres réels

Définitions. Pour tout nombre réel $r \neq 0$, on appelle *signature de* r le nombre

$$\operatorname{sign}(r) = \begin{cases} 1 & \text{si} \quad r > 0, \\ -1 & \text{si} \quad r < 0. \end{cases}$$

Si r est un nombre réel, on appelle

a) *partie positive de* r le nombre

$$r_+ = \begin{cases} r & \text{si} \quad r \geq 0, \\ 0 & \text{si} \quad r < 0, \end{cases}$$

b) *partie négative de* r le nombre

$$r_- = \begin{cases} 0 & \text{si} \quad r \geq 0, \\ -r & \text{si} \quad r < 0, \end{cases}$$

Ces nombres sont intimement liés à r et $|r|$.

Proposition 2.2.6.1 *Pour tous r, $r' \in \mathbb{R}$, on a*

a) $r_+ = (-r)_-$ *et* $r_- = (-r)_+$,

b) $r = r_+ - r_-$ *et* $|r| = r_+ + r_-$,

c) $r_+ = \frac{1}{2}(|r| + r)$ *et* $r_- = \frac{1}{2}(|r| - r)$,

d) $\left|r_+ - r'_+\right| \leq |r - r'|$ *et* $\left|r_- - r'_-\right| \leq |r - r'|$.

Preuve. a), b) et c) sont immédiats.

d) De fait, on a successivement

$$\left|r_\pm - r'_\pm\right| \leq \tfrac{1}{2}\left||r| \pm r - (|r'| \pm r')\right| \leq \tfrac{1}{2}\left||r| - |r'|\right| + \tfrac{1}{2}|r - r'| \leq |r - r'| . \blacksquare$$

Ces notions de parties positive et négative de nombres réels peuvent être en quelque sorte généralisée au cas d'un nombre fini de nombres réels.

Définitions. A tout nombre fini de nombres réels $r_1, \ldots r_J$, on associe

a) la *borne supérieure de* $r_1, \ldots, r_J$: c'est le plus grand des nombres $r_1, \ldots, r_J$; il est noté

$$\sup\{r_1, \ldots, r_J\} \quad \text{ou} \quad \sup_{1 \leq j \leq J} r_j,$$

b) la *borne inférieure de* $r_1, \ldots, r_J$: c'est le plus petit des nombres $r_1, \ldots, r_J$; il est noté

$$\inf\{r_1, \ldots, r_J\} \quad \text{ou} \quad \inf_{1 \leq j \leq J} r_j.$$

Remarquons de suite que la borne supérieure et la borne inférieure d'un nombre fini de nombres réels
a) ont un sens; on dit qu'elles *existent*,
b) sont *réalisées*, c'est-à-dire qu'il s'agit de nombres qui appartiennent à l'ensemble $\{r_1, \ldots, r_J\}$.

Voici quelques propriétés relatives à ces notions.

Proposition 2.2.6.2 *Pour tout $J \in \mathbb{N}_0$ et tous $r_1, \ldots r_J \in \mathbb{R}$,*
a) *on a*

$$\begin{array}{rcl} r_+ &=& \sup\{r, 0\}, \\ r_- &=& \sup\{-r, 0\}, \\ |r| &=& \sup\{r, -r\} \; = \; \sup\{r_+, r_-\}, \\ 0 &=& \inf\{r_+, r_-\}, \end{array}$$

b) *on a*

$$\begin{array}{rcl} \sup\{r_1, \ldots, r_J\} &=& -\inf\{-r_1, \ldots, -r_J\}, \\ \inf\{r_1, \ldots, r_J\} &=& -\sup\{-r_1, \ldots, -r_J\}, \end{array}$$

c) *on obtient* $\sup\{r_1, \ldots, r_J\}$ *et* $\inf\{r_1, \ldots, r_J\}$ *au moyen d'un nombre fini de combinaisons linéaires et de passage aux parties positives, aux parties négatives ou aux modules.*

Preuve. a) et b) sont immédiats.
c) Pour deux éléments, cela résulte aussitôt des égalités

$$\begin{array}{rcl} \sup\{r_1, r_2\} = r_1 + (r_2 - r_1)_+ &=& r_1 + (r_1 - r_2)_- = \dfrac{1}{2}(r_1 + r_2 + |r_1 - r_2|), \\ \inf\{r_1, r_2\} = r_1 - (r_1 - r_2)_+ &=& r_1 - (r_2 - r_1)_- = \dfrac{1}{2}(r_1 + r_2 - |r_1 - r_2|). \end{array}$$

Pour plus de deux nombres, on procède par récurrence en notant que

$$\begin{array}{rcl} \sup\{r_1, \ldots, r_J\} &=& \sup\{\ldots, \sup\{\sup\{r_1, r_2\}, r_3\}, \ldots, r_J\} \\ \inf\{r_1, \ldots, r_J\} &=& \inf\{\ldots, \inf\{\inf\{r_1, r_2\}, r_3\}, \ldots, r_J\}. \end{array}$$

2.2.7 Parties majorées ou minorées de $\mathbb{R}$

Définitions. Une partie A de $\mathbb{R}$ est
a) *majorée* s'il existe $M \in \mathbb{R}$ tel que $A \subset \,]-\infty, M]$, c'est-à-dire tel que $x \leq M$ pour tout $x \in A$. Un tel nombre M est appelé un *majorant* ou une *majorante de* A.
b) *minorée* s'il existe $m \in \mathbb{R}$ tel que $A \subset [m, +\infty[$, c'est-à-dire tel que $x \geq m$ pour tout $x \in A$. Un tel nombre m est appelé un *minorant* ou une *minorante de* A.

Remarquons immédiatement que seuls les intervalles de $\mathbb{R}$ d'un des types

$$]a,b[\,,[a,b]\,,]a,b]\,,[a,b[\,,]-\infty,b[\quad \text{et} \quad]-\infty,b]$$
$$(\text{resp. }]a,b[\,,[a,b]\,,]a,b]\,,[a,b[\,,]a,+\infty[\quad \text{et} \quad [a,+\infty[)$$

sont majorés (resp. minorés). De la sorte, toutes les possibilités existent indépendamment les unes des autres, ainsi que le prouvent les ensembles $]-\infty,-1[$, $[-1,1]$, $]1,+\infty[$ et $\mathbb{R}$.

Si M (resp. m) est une majorante (resp. minorante) de $A \subset \mathbb{R}$, il en est bien sûr de même pour tout nombre réel $M' \geq M$ (resp. $m' \leq m$). Mais il peut en être bien différemment pour les nombres réels M' (resp. m') tels que $M' < M$ (resp. $m' > m$). On en arrive ainsi au concept suivant.

Définitions. Un nombre réel M (resp. m) est une *borne supérieure* (resp. *borne inférieure*) *de* A si les deux conditions suivantes sont réalisées:

a) M (resp. m) est une majorante (resp. une minorante) de A,

b) pour toute majorante M' (resp. minorante m') de A, on a $M' \geq M$ (resp. $m' \leq m$).

De manière imagée, une borne supérieure (resp. inférieure) apparaît comme un majorant (resp. un minorant) qu'on ne peut pas améliorer.

Signalons sans démonstration le résultat suivant de la théorie des nombres réels. Il est fondamental: c'est un critère puissant permettant de déterminer si une partie de $\mathbb{R}$ admet une borne supérieure (resp. inférieure).

Théorème 2.2.7.1 (existence) a) *Toute partie majorée de* $\mathbb{R}$ *admet une borne supérieure.*

b) *Toute partie minorée de* $\mathbb{R}$ *admet une borne inférieure.*∎

Ce résultat essentiel pour l'analyse une fois admis, nous allons établir rigoureusement les autres propriétés.

Théorème 2.2.7.2 (unicité) a) *Toute partie majorée de* $\mathbb{R}$ *admet une borne supérieure unique.*

b) *Toute partie minorée de* $\mathbb{R}$ *admet une borne inférieure unique.*

Preuve. Cela résulte aussitôt de la définition car, par exemple, si M et M' sont des bornes supérieures de A, on doit avoir $M \leq M'$ et $M' \leq M$, donc $M = M'$.∎

Ce théorème permet donc de parler de *la* borne supérieure d'une partie majorée de $\mathbb{R}$ et de *la* borne inférieure d'une partie minorée de $\mathbb{R}$. De la sorte, nous pouvons introduire les notations suivantes.

Notations. a) La borne supérieure d'une partie majorée A de $\mathbb{R}$ est notée

$$\sup A \quad \text{ou} \quad \sup_P x$$

si P est une propriété définissant A.

b) La borne inférieure d'une partie minorée A de $\mathbb{R}$ est notée

$$\inf A \quad \text{ou} \quad \inf_P x$$

si P est une propriété définissant A.

Passons à présent aux propriétés des bornes supérieures et des bornes inférieures des parties de $\mathbb{R}$.

Critère 2.2.7.3 a) *Une majorante M d'une partie A de $\mathbb{R}$ est la borne supérieure de A si et seulement si, pour tout $\varepsilon > 0$, il existe $x \in A$ tel que $M - \varepsilon \leq x$.*

b) *Une minorante m d'une partie A de $\mathbb{R}$ est la borne inférieure de A si et seulement si, pour tout $\varepsilon > 0$, il existe $x \in A$ tel que $m + \varepsilon \geq x$.*

Preuve. a) La condition est nécessaire. De fait, étant strictement inférieur à M, $M - \varepsilon$ n'est pas une majorante de A.

La condition est suffisante. De fait, aucun nombre $M' < M$ ne peut alors être une majorante de A.

b) s'établit de même.∎

Corollaire 2.2.7.4 a) *Si M est une majorante de A et si M appartient à A, alors M est la borne supérieure de A.*

b) *Si m est une minorante de A et si m appartient à A, alors m est la borne inférieure de A.*∎

Vu ce qui précède, on conçoit que les deux possibilités suivantes peuvent se présenter.

Définitions. Soit M (resp. m) la borne supérieure (resp. inférieure) de la partie majorée (resp. minorée) A de $\mathbb{R}$.

a) D'une part, M (resp. m) peut appartenir à A. On dit alors que la borne supérieure (resp. inférieure) est *réalisée* ou *atteinte* et que M (resp. m) est un *maximum* (resp. *minimum*) *de* A. C'est toujours le cas si A est fini, comme nous l'avons vu au paragraphe précédent.

b) D'autre part, M (resp. m) peut ne pas appartenir à A; on dit alors que la borne supérieure (resp. inférieure) n'est pas réalisée.

En fait, ces deux possibilités existent comme le montre immédiatement la considération des ensembles $]-1, 1[$, $[-1, 1]$, $]-1, 1]$ et $[-1, 1[$.

2.2.8 Bornés

Définition. Une partie B de $\mathbb{R}^n$ est *bornée* si elle est incluse dans une boule de centre 0, c'est-à-dire s'il existe $C > 0$ tel que $|x| \leq C$ pour tout $x \in B$.

Proposition 2.2.8.1 *Une partie B de $\mathbb{R}$ est bornée si et seulement si elle est majorée et minorée.*

Preuve. La condition est nécessaire. De fait, si $C > 0$ est tel que $|x| \leq C$ pour tout $x \in B$, il vient $-C \leq x \leq C$ pour tout $x \in B$ et dès lors C est un majorant et $-C$ un minorant de B.

La condition est suffisante. De fait, si M et m sont respectivement une majorante et une minorante de B, on vérifie de suite que la majoration $|x| \leq \sup\{|m|, |M|\}$ a lieu pour tout $x \in B$.∎

Exemple. *Un intervalle de $\mathbb{R}$ est borné si et seulement s'il est d'un des types $]a, b[$, $[a, b]$, $]a, b]$ ou $[a, b[$.*

Exemple. *Dans $\mathbb{R}^n$, toute boule est bornée.* De fait, si b désigne une des boules de centre $a \in \mathbb{R}^n$ et de rayon $r > 0$, $r + |a|$ est un nombre strictement positif tel que

$$|x| = |x - a + a| \leq |x - a| + |a| \leq r + |a|, \quad \forall x \in b.$$

Le cas des intervalles de $\mathbb{R}^n$ est réglé par le résultat suivant.

Critère 2.2.8.2 *Une partie B de $\mathbb{R}^n$ est bornée si et seulement si, pour tout $j \in \{1, \ldots, n\}$, l'ensemble $\{\, x_j : x \in B\}$ est borné dans $\mathbb{R}$.*

Preuve. La condition est nécessaire. De fait, si $C > 0$ donne lieu à $|x| \leq C$ pour tout $x \in B$, alors on a $|x_j| \leq C$ pour tout $x \in B$ et tout $j \in \{1, \ldots, n\}$.

La condition est suffisante. De fait, si, pour tout $j \in \{1, \ldots, n\}$, l'ensemble $\{\, x_j : x \in B\}$ est inclus dans $[a_j, b_j]$, on obtient de suite

$$|x| \leq \sum_{j=1}^{n} |x_j| \leq \sum_{j=1}^{n} \sup\{|a_j|, |b_j|\}, \quad \forall x \in B.∎$$

Corollaire 2.2.8.3 *Un intervalle de $\mathbb{R}^n$ est borné si et seulement si chacun de ses intervalles constitutifs est borné dans $\mathbb{R}$.*∎

Proposition 2.2.8.4 a) *Toute partie d'un borné est bornée. En particulier, toute intersection de bornés est bornée.*

b) *Toute union finie de bornés est bornée. En particulier, toute partie finie est bornée.*

Preuve. a) est trivial.

b) Si $B_1, \ldots, B_J$ sont des bornés de $\mathbb{R}^n$ avec $J \in \mathbb{N}_0$, alors, pour tout $j \in \{1, \ldots, J\}$, il existe $C_j > 0$ tel que $|x| \leq C_j$ pour tout $x \in B_j$. Cela étant, $B = \bigcup_{j=1}^{J} B_j$ est borné car on a $|x| \leq \sup\{C_1, \ldots, C_J\}$ pour tout $x \in B$.∎

2.2.9 Diamètre d'une partie

Définition. Soit A une partie non vide de $\mathbb{R}^n$. Si $\{\,|x-y| : x, y \in A\}$ est une partie majorée de $\mathbb{R}$, on appelle *diamètre de A* le nombre

$$\operatorname{diam}(A) = \sup\{\,|x-y| : x, y \in A\} = \sup_{x,y \in A} |x-y|\,.$$

La description des parties non vides de $\mathbb{R}^n$ qui admettent un diamètre est donnée par la propriété suivante.

Proposition 2.2.9.1 *Une partie non vide de $\mathbb{R}^n$ admet un diamètre si et seulement si elle est bornée.*

Preuve. La condition est nécessaire. De fait, si A est une partie non vide de $\mathbb{R}^n$ qui admet un diamètre, choisissons un élément a de A. Il vient alors

$$|x| \leq |x-a+a| \leq |x-a| + |a| \leq \operatorname{diam}(A) + |a|\,, \quad \forall x \in A,$$

c'est-à-dire que A est borné.

La condition est suffisante. Si $B \subset \mathbb{R}^n$ est borné et non vide, il existe $C > 0$ tel que $|x| \leq C$ pour tout $x \in B$ donc tel que $|x-y| \leq |x| + |y| \leq 2C$ pour tous x, $y \in B$, ce qui suffit.∎

Exemple. *Toute boule de $\mathbb{R}^n$ admet un diamètre. Plus précisément, pour tout $a \in \mathbb{R}^n$ et tout $r > 0$, on a*

$$\operatorname{diam}(\{\,x : |x-a| < r\}) = \operatorname{diam}(\{\,x : |x-a| \leq r\}) = 2r.$$

Nous savons que toute boule de $\mathbb{R}^n$ est non vide et bornée, donc admet un diamètre. Cela étant, d'une part, $2r$ est une majorante de

$$\{\,|x-y| : |x-a| \leq r, |y-a| \leq r\}\,,$$

vu la formule

$$|x-y| = |(x-a)+(a-y)| \leq |x-a| + |y-a|\,, \quad \forall x, y \in \mathbb{R}^n.$$

D'autre part, pour tout $s \in \,]0, r[$, on vérifie aisément que les points $a + se_1$ et $a - se_1$ appartiennent à la boule considérée et sont tels que $|(a+se_1) - (a-se_1)| = 2s$. D'où la conclusion vu le critère 2.2.7.3 relatif aux bornes supérieures.

Exemple. *Tout intervalle borné admet un diamètre. Plus précisément, le diamètre de l'intervalle $I = I_1 \times \cdots \times I_n$ est égal à $|b-a|$ si, pour tout $j \in \{1, \ldots n\}$, a_j est l'origine de I_j et b_j son extrémité.* Nous savons que tout intervalle est non vide donc que tout intervalle borné admet un diamètre. Cela étant, d'une part, $|b-a|$ est une majorante de l'ensemble $\{\, |x-y| : x, y \in I\}$ car on a

$$|x-y| = \sqrt{\sum_{j=1}^{n}(x_j - y_j)^2} \leq \sqrt{\sum_{j=1}^{n}(b_j - a_j)^2} = |b-a|\,, \quad \forall x, y \in I.$$

D'autre part, pour tout $s \in \,]0, 1/2[$, on vérifie aisément que les points $a + s(b-a)$ et $a + (1-s)(b-a)$ appartiennent à I et vérifient

$$|a + s(b-a) - a - (1-s)(b-a)| = |1-2s|\,|b-a|\,.$$

La conclusion s'ensuit directement du critère 2.2.7.3 relatif aux bornes supérieure et inférieure. Si les points a et b appartiennent à I, on peut conclure beaucoup plus vite.▮

Remarques. a) Si le diamètre d'une partie B non vide de $\mathbb{R}^n$ existe, cela n'entraîne pas qu'il est réalisé: il n'existe pas nécessairement des points x, y de B tels que $|x-y| = \operatorname{diam}(B)$. Par exemple, on a bien sûr $\operatorname{diam}(]a, b[) = b - a$ et cependant il n'existe pas de points x, $y \in \,]a, b[$ tels que $|x - y| = b - a$.

b) Pourtant il existe des parties B non vides de $\mathbb{R}^n$ qui admettent un diamètre réalisé. $* \rightarrow$ C'est le cas des boules fermées: les points $a + re_1$ et $a - re_1$ appartiennent à la boule $\{\, x : |x-a| \leq r\}$ et leur distance est égale à $2r$. Plus généralement, nous établirons au paragraphe 2.4.3 que tout compact non vide de $\mathbb{R}^n$ admet un diamètre réalisé. $\leftarrow *$

c) Enfin le diamètre d'une partie non vide de $\mathbb{R}^n$ existe et est égal à 0 si et seulement si cette partie est un singleton.

2.2.10 Distance de deux parties

Etant donné deux parties non vides A et A' de $\mathbb{R}^n$, on appelle *distance de A à A'* le nombre

$$\mathrm{d}(A, A') = \inf\{\, |x-y| : x \in A, y \in A'\} = \inf_{x \in A,\, y \in A'} |x - y|\,.$$

Si l'ensemble A' est le singleton $\{x\}$, la distance de x à A est notée $\mathrm{d}(x, A)$ plutôt que $\mathrm{d}(\{x\}, A)$.

Proposition 2.2.10.1 *Si A, B et C sont des parties non vides de $\mathbb{R}^n$,*

a) $\mathrm{d}(A, B) \geq 0$,

b) $\mathrm{d}(A, B) = \mathrm{d}(B, A)$,

c) $\mathrm{d}(A, C) \leq \mathrm{d}(A, B) + \operatorname{diam}(B) + \mathrm{d}(B, C)$ *si B est borné.*

Preuve. a) et b) sont triviaux.

c) Pour tous $a \in A$, b, $b' \in B$ et $c \in C$, on a évidemment

$$\begin{array}{rcl} \mathrm{d}(A,C) & \leq & |a-c| \;=\; |(a-b)+(b-b')+(b'-c)| \\ & \leq & |a-b|+|b-b'|+|b'-c| \\ & \leq & |a-b|+\mathrm{diam}(B)+|b'-c|\,. \end{array}$$

Pour tous $a \in A$ et $b \in B$, le nombre $\mathrm{d}(A,C)-|a-b|-\mathrm{diam}(B)$ est donc une minorante de $\{\,|b'-c| : b' \in B, c \in C\}$. De la sorte, on obtient

$$\mathrm{d}(A,C)-|a-b|-\mathrm{diam}(B) \leq \mathrm{d}(B,C), \quad \forall a \in A, \forall b \in B.$$

Cela étant, le nombre $\mathrm{d}(A,C)-\mathrm{diam}(B)-\mathrm{d}(B,C)$ est une minorante de l'ensemble $\{\,|a-b| : a \in A, b \in B\}$; on a donc

$$\mathrm{d}(A,C)-\mathrm{diam}(B)-\mathrm{d}(B,C) \leq \mathrm{d}(A,B),$$

ce qui suffit.∎

Remarques. a) La distance de deux parties non vides de $\mathbb{R}^n$ peut valoir 0 sans que ces parties ne soient égales. Pour s'en convaincre, il suffit bien sûr de considérer deux parties non disjointes. Cependant la distance de deux parties non vides et disjointes de $\mathbb{R}^n$ peut aussi valoir 0: c'est par exemple le cas pour les ensembles $]-\infty,0[$ et $[0,+\infty[$.

b) Comme $\mathrm{d}(A,A')=0$ n'implique par $A=A'$, d n'est pas une distance sur l'ensemble des parties non vides de $\mathbb{R}^n$. Il s'agit d'un abus de langage consacré par l'usage.

c) Il n'est pas possible d'améliorer la formule c) de la proposition précédente. En effet, pour les ensembles $A=[0,1]$, $B=[2,3]$ et $C=[5,6]$, on vérifie de suite que l'égalité

$$\mathrm{d}(A,C)=\mathrm{d}(A,B)+\mathrm{diam}(B)+\mathrm{d}(B,C)$$

a lieu avec $\mathrm{diam}(B)>0$.

d) Insistons sur le fait suivant, déjà remarqué à la remarque a). La distance de deux parties non vides A, A' de $\mathbb{R}^n$ n'est pas nécessairement réalisée, c'est-à-dire qu'il n'existe pas nécessairement des points $a \in A$ et $a' \in A'$ pour lesquels l'égalité $\mathrm{d}(A,A')=\mathrm{d}(a,a')$ a lieu.

e) $* \rightarrow$ Au paragraphe 2.4.3, nous établirons que, dans $\mathbb{R}^n$, la distance d'un compact non vide à un fermé non vide est toujours réalisée. $\leftarrow *$

Proposition 2.2.10.2 *Pour tous x, $y \in \mathbb{R}^n$ et toute partie non vide A de $\mathbb{R}^n$, on a*

$$|\mathrm{d}(x,A)-\mathrm{d}(y,A)| \leq |x-y|\,.$$

Preuve. Cela résulte aussitôt des majorations

$$\begin{array}{rcl} \mathrm{d}(x,A) & \leq & \mathrm{d}(x,y)+\mathrm{diam}(\{y\})+\mathrm{d}(y,A) \\ \mathrm{d}(y,A) & \leq & \mathrm{d}(y,x)+\mathrm{diam}(\{x\})+\mathrm{d}(x,A) \end{array}$$

car on a $\mathrm{diam}(\{x\})=\mathrm{diam}(\{y\})=0$.∎

2.3 Suites dans $\mathbb{R}^n$

2.3.1 Définition

Il est préférable de formuler dès à présent la définition d'une suite en toute généralité et ne pas se limiter aux seules suites de $\mathbb{R}^n$.

Définitions. Une *suite* de l'ensemble non vide E est une application de $\mathbb{N}_0$ dans E; c'est donc une loi qui, à tout $m \in \mathbb{N}_0$, associe un élément x_m de E, appelé *m-ème élément* de la suite.

La suite elle-même est notée

$$(x_m)_{m\in\mathbb{N}_0} \quad \text{ou} \quad x_1, x_2, x_3, \ldots$$

ou même tout simplement on parle de la suite x_m si le contexte est clair.

En particulier si $(x(m))_{m\in\mathbb{N}_0}$ est une suite de E et si, à tout $x(m)$, on associe un élément $y_{x(m)}$ d'un ensemble E', on parle également de la suite $(y_{x(m)})_{m\in\mathbb{N}_0}$ qui, à tout $m \in \mathbb{N}_0$, associe l'élément $y_{x(m)}$ de E'. Ceci conduit notamment aux situations suivantes. Soit $(y_m)_{m\in\mathbb{N}_0}$ une suite de E. Comme $(x_m = 2m)_{m\in\mathbb{N}_0}$ est une suite de $\mathbb{N}_0$, on peut parler de la suite y_{2m} des éléments pairs de la suite de départ. De même, étant donné $M \in \mathbb{N}_0$, $(M+m-1)_{m\in\mathbb{N}_0}$ est une suite dans $\mathbb{N}_0$ et on peut donc parler de la suite y_{M+m-1} dont le m-ème élément est y_{M+m-1}.

Remarques. a) Deux suites x_m et y_m sont donc égales si et seulement si, pour tout $m \in \mathbb{N}_0$, on a $x_m = y_m$.

b) La suite $(x_m)_{m\in\mathbb{N}_0}$ et l'ensemble $\{\, x_m : m \in \mathbb{N}_0\}$ de ses éléments sont deux notions distinctes, à ne pas confondre.

Moyens d'obtention d'une suite. Une suite peut être obtenue au moyens des trois procédés suivants:

a) en donnant explicitement l'application qui, à tout $m \in \mathbb{N}_0$, associe l'élément x_m de E. Il en est ainsi de

$$x_m = m^2 + m + 1, \quad \forall m \in \mathbb{N}_0.$$

b) en donnant explicitement une relation de récurrence: on donne un entier $M \in \mathbb{N}_0$, des points $x_1, \ldots, x_M$ de E et une application qui, à tout entier $m > M$ et aux points $x_1, \ldots, x_{m-1}$, associe un point x_m de E. Il en est ainsi de

$$x_1 = 1,\ x_2 = 2 \ \text{ et } \ x_m = \sqrt{x_{m-1}x_{m-2}} \ \text{ pour tout } \ m \geq 3.$$

c) étant donné une suite $(E_m)_{m\in\mathbb{N}_0}$ d'ensembles non vides, il existe une suite $(x_m)_{m\in\mathbb{N}_0}$ telle que, pour tout $m \in \mathbb{N}_0$, x_m appartienne à E_m.

Remarque. Les moyens a) et b) d'obtention d'une suite permettent de déterminer explicitement chaque élément de la suite considérée. Il n'en est pas de même avec le procédé c) qui affirme tout simplement l'*existence* d'une suite.

Définition. Une *sous-suite* de la suite $(x_m)_{m\in\mathbb{N}_0}$ est une suite $(x_{k(m)})_{m\in\mathbb{N}_0}$ pour laquelle

$$k(1) \geq 1 \ \text{ et } \ k(m+1) > k(m) \ \text{ pour tout } \ m \in \mathbb{N}_0$$

(on a donc $k(m) \geq m$ pour tout $m \in \mathbb{N}_0$).

L'interprétation de cette notion est la suivante. Ecrivons les éléments de la suite

$$x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, \ldots$$

Supprimons les éléments x_j dont l'indice ne coïncide pas avec un des nombres $k(m)$: il reste par exemple

$$x_2, x_5, x_6, x_8, \ldots$$

si on a $k(1) = 2$, $k(2) = 5$, $k(3) = 6$, $k(4) = 8$, ... Cela étant, si, pour tout $m \in \mathbb{N}_0$, on baptise y_m le m-ème élément de ce qui “reste”, on obtient la suite $\left(y_m = x_{k(m)}\right)_{m\in\mathbb{N}_0}$.

Exemples. Etant donné une suite $(x_m)_{m\in\mathbb{N}_0}$, on peut considérer ses sous-suites

$$(x_{2m})_{m\in\mathbb{N}_0}, (x_{2m-1})_{m\in\mathbb{N}_0}, (x_{m^2})_{m\in\mathbb{N}_0}, \ldots$$

qui correspondent respectivement à

$$k(m) = 2m,\ k(m) = 2m-1,\ k(m) = m^2, \ldots \quad \forall m \in \mathbb{N}_0.$$

2.3.2 Suites convergentes

Le concept de “suite convergente” est fondamental en analyse. Dans un premier temps, il suffit de l'introduire dans $\mathbb{R}^n$, ce qui rend son étude plus concrète.

Définitions. Une suite $(x_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$ *converge vers* $a \in \mathbb{R}^n$ (on dit aussi *tend vers* $a \in \mathbb{R}^n$) si, pour tout $\varepsilon > 0$, il existe un nombre réel M tel que

$$m \geq M \Rightarrow |x_m - a| \leq \varepsilon;$$

le point a est alors appelé *limite de la suite* x_m et on écrit

$$x_m \to a \ \text{ ou } \ \lim_{m\to\infty} x_m = a.$$

Interprétation. Si a est un point de $\mathbb{R}^n$, appelons *voisinage de* a toute partie de $\mathbb{R}^n$ contenant une boule centrée en a. Cela étant, la suite $(x_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$ converge vers a si et seulement si, pour tout voisinage V de a, il existe un nombre réel M tel que, pour tout $m \geq M$, on ait $x_m \in V$. On exprime bien souvent ce fait sous la forme imagée que voici: la suite x_m converge vers a si et seulement si elle finit par "entrer et rester" dans tout voisinage de a.

Définitions. Une suite $(x_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$ *converge* ou est *convergente* s'il existe un point a de $\mathbb{R}^n$ vers lequel elle converge; dans la cas contraire, on dit que la suite *diverge* ou est *divergente*.

Exemples. a) *Soit a un point de $\mathbb{R}^n$. La suite "constante" $(x_m = a)_{m\in\mathbb{N}_0}$ converge vers a.*

b) *La suite $(x_m = 1/m)_{m\in\mathbb{N}_0}$ converge vers 0 dans $\mathbb{R}$.* De fait, pour tout $\varepsilon > 0$, $1/\varepsilon$ est un nombre réel tel que, pour tout $m \geq 1/\varepsilon$, on a $|1/m - 0| \leq \varepsilon$.

c) *Si la partie non vide A de $\mathbb{R}$ est majorée* (resp. *minorée*)*, il existe une suite de A qui converge vers la borne supérieure* (resp. *inférieure*) *de A.* C'est une conséquence directe du critère 2.2.7.3 relatif aux bornes supérieure et inférieure. Considérons par exemple le cas où la partie A est majorée et désignons par a sa borne supérieure. Vu ce critère, pour tout $m \in \mathbb{N}_0$, l'ensemble

$$A_m = A \bigcap \{\, x : |x - a| \leq 1/m\}$$

n'est pas vide. Cela étant, le troisième moyen d'obtention d'une suite appliqué à la suite A_m procure aussitôt une suite x_m telle que $x_m \in A_m$ pour tout $m \in \mathbb{N}_0$, c'est-à-dire une suite de A telle que $|x_m - a| \leq 1/m$ pour tout $m \in \mathbb{N}_0$, ce qui suffit. $* \rightarrow$ Cet exemple sera amélioré au paragraphe 2.3.4. $\leftarrow *$

Remarque. Si la suite $(x_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$ converge vers a, a peut être égal à chacun des éléments, à une infinité des éléments, à un nombre fini des éléments ou à aucun des éléments de la suite ainsi que le confirment les suites suivantes

$$\begin{aligned}
&x_m = 0, \quad \forall m \in \mathbb{N}_0,\\
&x_{2m} = 0 \ \text{ et } \ x_{2m-1} = 1/m, \quad \forall m \in \mathbb{N}_0,\\
&x_1 = \ldots = x_9 = 0 \ \text{ et } \ x_m = 1/m, \quad \forall m \geq 10,\\
&x_m = 1/m, \quad \forall m \in \mathbb{N}_0.
\end{aligned}$$

Voici trois propriétés fondamentales des suites convergentes qui ne font intervenir que le module de $\mathbb{R}^n$—plus loin nous étudierons les propriétés des suites qui font intervenir davantage que le module.

Théorème 2.3.2.1 (unicité de la limite) *Si une suite converge, sa limite est unique.*

Preuve. Supposons que la suite x_m converge à la fois vers a et vers b. Si a diffère de b, $|a-b|$ est un nombre strictement positif et, par conséquent, il existe des nombres réels M et N tels que

$$\begin{aligned} m \geq M &\Rightarrow |x_m - a| \leq \tfrac{1}{3}\,|a-b|\,, \\ m \geq N &\Rightarrow |x_m - b| \leq \tfrac{1}{3}\,|a-b|\,. \end{aligned}$$

Au total, on obtient

$$m \geq \sup\{M, N\} \Rightarrow |a-b| \leq |a-x_m| + |x_m - b| \leq \tfrac{2}{3}\,|a-b|\,,$$

ce qui est contradictoire.∎

Théorème 2.3.2.2 *Toute sous-suite d'une suite convergente converge vers la limite de la suite.*

Preuve. Soit x_m une suite qui converge vers a et soit $x_{k(m)}$ une sous-suite de cette suite. Pour tout $\varepsilon > 0$, il existe un nombre réel M tel que

$$m \geq M \Rightarrow |x_m - a| \leq \varepsilon$$

donc tel que

$$m \geq M \Rightarrow \left|x_{k(m)} - a\right| \leq \varepsilon$$

car on a $k(m) \geq m$ pour tout $m \in \mathbb{N}_0$. D'où la conclusion.∎

Théorème 2.3.2.3 *L'ensemble des points d'une suite convergente est borné.*

Preuve. De fait, si la suite x_m converge vers a, il existe un nombre réel M tel que

$$m \geq M \Rightarrow |x_m - a| \leq 1$$

donc tel que

$$|x_m| \leq \sup\{|x_1|\,, \ldots, |x_{M-1}|\,, |a| + 1\}, \quad \forall m \in \mathbb{N}_0.$$∎

Enfin, introduisons une manière particulière de converger dans $\mathbb{R}^n$.

Définition. Soient A une partie et a un point de $\mathbb{R}^n$. La suite $(x_m)_{m\in\mathbb{N}_0}$ *converge dans A vers a* si elle converge vers a et s'il existe un nombre réel M tel que $x_m \in A$ pour tout $m \geq M$. On exprime bien souvent la deuxième partie de cette condition sous une des formes suivantes: "et si x_m appartient à A pour m suffisamment grand" ou "et si la suite finit par entrer et rester dans A". Il convient de bien remarquer que nous n'avons pas exigé que a appartienne à A.

2.3.3 Suites divergentes, suites convergentes vers ∞

a) Suites divergentes

La suite $(x_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$ ne converge pas vers $a \in \mathbb{R}^n$ si et seulement si

$$\exists \varepsilon > 0 \text{ tel que, } \forall M \in \mathbb{R}, \exists m \geq M \text{ tel que } |x_m - a| > \varepsilon.$$

Cela étant, exprimer qu'une suite ne converge pas devient fastidieux car cela équivaut à “pour tout $a \in \mathbb{R}^n$, la suite x_m ne converge pas vers a”.

Cependant, en niant une des propriétés des suites convergentes, on obtient automatiquement un critère assez aisé de divergence.

Ainsi, le résultat “toute sous-suite d'une suite convergente dans $\mathbb{R}^n$ converge vers la même limite” donne un puissant critère de divergence. Il suffit, en effet, pour affirmer que la suite $(x_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$ ne converge pas, de produire deux sous-suites qui convergent vers des limites différentes. Ainsi, si a et b sont deux points distincts de $\mathbb{R}^n$, la suite définie par

$$x_{2m} = a \text{ et } x_{2m-1} = b, \quad \forall m \in \mathbb{N}_0,$$

ne converge pas.

Il en est de même du résultat “l'ensemble des points d'une suite convergente de $\mathbb{R}^n$ est borné”. Ainsi, la suite de $\mathbb{R}$ définie par $x_m = m$ pour tout $m \in \mathbb{N}_0$ ne converge pas.

b) Suites convergentes vers ∞

Introduisons à présent une manière très particulière mais fort importante de diverger dans $\mathbb{R}^n$.

Si la suite $(x_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$ jouit de la propriété suivante “pour tout $N > 0$, il existe un nombre réel M tel que, pour tout $m \geq M$, on ait $|x_m| \geq N$”, il est certain que l'ensemble de ses points n'est pas borné donc que cette suite ne converge pas. Cependant, à cause d'un abus consacré par l'usage, on introduit les définitions suivantes.

Définitions. La suite $(x_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$ *converge vers l'infini* ou *tend vers l'infini* et on écrit

$$x_m \to \infty \text{ ou } \lim_{m\to\infty} x_m = \infty$$

si

$$\forall N > 0, \exists M\colon m \geq M \Rightarrow |x_m| \geq N,$$

et on dit que l'infini est la *limite* de la suite.

Interprétation. Appelons *voisinage de l'infini dans* $\mathbb{R}^n$ toute partie de $\mathbb{R}^n$ contenant le complémentaire d'une boule centrée à l'origine. La suite $(x_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$ converge vers l'infini si et seulement si, pour tout voisinage V de l'infini, il existe un nombre réel M tel que, pour tout $m \geq M$, on ait $x_m \in V$. On exprime bien souvent ce fait au moyen de l'expression imagée suivante: la suite $(x_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$ converge vers l'infini si et seulement si elle "finit par entrer et rester dans tout voisinage de l'infini".

Insistons encore une fois: "la convergence vers l'infini dans $\mathbb{R}^n$" caractérise en fait une manière particulière de diverger! Si on a adopté cette formule qui a priori étonne, c'est parce que la formulation des propriétés des suites qui convergent vers l'infini est souvent fort proche de celle des suites convergentes.

Cela apparaît déjà clairement dans l'interprétation précédente. Un autre exemple, très éloquent, est donné par le résultat suivant (et sa démonstration).

Proposition 2.3.3.1 *Toute sous-suite d'une suite qui converge vers l'infini, converge également vers l'infini.*

Preuve. Soit $(x_m)_{m\in\mathbb{N}_0}$ une suite de $\mathbb{R}^n$ qui converge vers l'infini et soit $x_{k(m)}$ une sous-suite de cette suite. Pour tout $N > 0$, il existe un nombre réel M tel que

$$m \geq M \Rightarrow |x_m| \geq N$$

donc tel que

$$m \geq M \Rightarrow \left|x_{k(m)}\right| \geq N$$

car on a $k(m) \geq m$ pour tout $m \in \mathbb{N}_0$. D'où la conclusion.∎

Remarques. Cependant il existe des résultats propres aux suites de $\mathbb{R}^n$ qui convergent vers un point de $\mathbb{R}^n$. Il en est ainsi par exemple de "l'ensemble des éléments d'une suite convergente de $\mathbb{R}^n$ est borné" alors que, si une suite de $\mathbb{R}^n$ converge vers l'infini, l'ensemble de ses éléments n'est évidemment pas borné.

Si on désire insister sur le fait que la suite converge vers un point de $\mathbb{R}^n$, on parle de suite *convergente vers une limite finie.* Ainsi, on peut énoncer "si une suite converge vers une limite finie, alors l'ensemble de ses éléments est borné". Cependant, il ne peut y avoir de confusion puisque, d'une part, on parle de suite convergente ou convergente vers a et, d'autre part, de suite convergente vers ∞ alors que ∞ n'est pas un élément de $\mathbb{R}^n$.

2.3.4 Suites numériques

Définitions. On appelle *suite numérique* toute suite de $\mathbb{C}$ et, plus particulièrement, *suite numérique réelle* toute suite de $\mathbb{R}$.

En recourant aux signes d'inégalité dans $\mathbb{R}$, on introduit les définitions suivantes pour les suites numériques réelles.

Une suite numérique réelle $(x_m)_{m\in\mathbb{N}_0}$ est

a) *croissante* si on a $x_m \le x_{m+1}$ pour tout $m \in \mathbb{N}_0$; on écrit $x_m \uparrow$,

b) *strictement croissante* si on a $x_m < x_{m+1}$ pour tout $m \in \mathbb{N}_0$,

c) *décroissante* si on a $x_m \ge x_{m+1}$ pour tout $m \in \mathbb{N}_0$; on écrit $x_m \downarrow$,

d) *strictement décroissante* si on a $x_m > x_{m+1}$ pour tout $m \in \mathbb{N}_0$,

e) *monotone* si elle est croissante ou décroissante,

f) *strictement monotone* si elle est strictement croissante ou strictement décroissante.

De même, dans $\mathbb{R}$, les signes d'inégalité permettent de préciser la manière de converger de certaines suites.

Définitions. Soit $(x_m)_{m\in\mathbb{N}_0}$ une suite numérique réelle qui converge vers $a \in \mathbb{R}$. On dit qu'elle converge

a) *à droite vers a* s'il existe un nombre réel M tel que $m \ge M$ entraîne $x_m \ge a$; il revient donc au même de dire qu'elle converge vers a dans l'ensemble $[a, +\infty[$. On écrit $x_m \to a^+$,

b) *à gauche vers a* s'il existe un nombre réel M tel que $m \ge M$ entraîne $x_m \le a$; il revient donc au même de dire qu'elle converge vers a dans l'ensemble $]-\infty, a]$. On écrit $x_m \to a^-$,

c) *en croissant vers a* ou qu'elle *croît vers a* si elle est croissante. On écrit $x_m \uparrow a$,

d) *en croissant strictement vers a* ou qu'elle *croît strictement vers a* si elle est strictement croissante,

e) *en décroissant vers a* ou qu'elle *décroît vers a* si elle est décroissante. On écrit $x_m \downarrow a$,

f) *en décroissant strictement vers a* ou qu'elle *décroît strictement vers a* si elle est strictement décroissante.

Exercice. Si $(r_m)_{m\in\mathbb{N}}$ est une suite de nombres strictement positifs tels que $r_0 = 1$ et $r_m^2 \le r_{m-1}r_{m+1}$ pour tout $m \in \mathbb{N}_0$, établir que la suite $(\sqrt[m]{r_m})_{m\in\mathbb{N}_0}$ est croissante.

Suggestion. Il suffit de procéder par récurrence. De $r_1^2 \le r_2$, on tire $r_1 \le \sqrt{r_2}$. Cela étant, si on a $r_1 \le \ldots \le \sqrt[m]{r_m}$, il vient

$$r_m^2 \le r_{m-1}r_{m+1} \le r_m^{(m-1)/m} r_{m+1} \text{ donc } \sqrt[m]{r_m} \le \sqrt[m+1]{r_{m+1}}.\square$$

Exemple. *Si la partie A de $\mathbb{R}$ est majorée* (resp. *minorée*)*, il existe une suite monotone de A qui converge vers la borne supérieure* (resp. *inférieure*) *de A. Si la borne supérieure* (resp. *inférieure*) *de A n'est pas réalisée, on peut en outre exiger que la suite soit strictement monotone.* Considérons tout d'abord le cas où $A \subset \mathbb{R}$

est majoré. D'une part, si la borne supérieure de A est réalisée, c'est-à-dire si A admet un maximum M, c'est trivial: il suffit de prendre la suite constante $x_m = M$ pour tout $m \in \mathbb{N}_0$. D'autre part, si la borne supérieure M de A n'est pas atteinte, pour tout $m \in \mathbb{N}_0$, l'ensemble $]M - 1/m, M[$ contient au moins un point de A. On en déduit aisément qu'il existe une suite $k(m)$ strictement croissante de $\mathbb{N}_0$ telle que, pour tout $m \in \mathbb{N}_0$, l'ensemble

$$A_m = A \bigcap]M - 1/k(m), M - 1/k(m+1)[$$

diffère de $\emptyset$. Le moyen c) d'obtention d'une suite affirme alors qu'il existe une suite x_m telle que $x_m \in A_m$ pour tout $m \in \mathbb{N}_0$. Cela étant, on vérifie aisément que cette suite x_m convient. Considérons à présent le cas où A est minoré: on peut soit adapter la démonstration précente, soit considérer l'ensemble majoré $A' = \{ -x : x \in A\}$: on vérifie alors que la suite $-x_m$ convient si la suite x_m convient pour la borne supérieure de A'.

De même, dans $\mathbb{R}$, on peut qualifier certaines suites qui convergent vers l'infini.

Définitions. La suite $(x_m)_{m \in \mathbb{N}_0}$ de $\mathbb{R}$

a) *converge vers* $+\infty$ si, pour tout $N > 0$, il existe un nombre réel M tel que $m \geq M$ implique $x_m \geq N$. Il s'agit donc des suites de $\mathbb{R}$ qui convergent vers l'infini dans $]0, +\infty[$. On écrit $x_m \to +\infty$.

b) *converge vers* $-\infty$ si, pour tout $N > 0$, il existe un nombre réel M tel que $m \geq M$ implique $x_m \leq -N$. Il s'agit donc des suites de $\mathbb{R}$ qui convergent vers l'infini dans $]-\infty, 0[$. On écrit $x_m \to -\infty$.

c) *converge en croissant vers* $+\infty$ ou *croît vers* $+\infty$ si elle est croissante et converge vers $+\infty$. On écrit $x_m \uparrow +\infty$.

d) *converge en croissant strictement vers* $+\infty$ ou *croît strictement vers* $+\infty$ si elle est strictement croissante et converge vers $+\infty$.

e) *converge en décroissant vers* $-\infty$ ou *décroît vers* $-\infty$ si elle est décroissante et converge vers $-\infty$. On écrit $x_m \downarrow -\infty$.

f) *converge en décroissant strictement vers* $-\infty$ ou *décroît strictement vers* $-\infty$ si elle est strictement décroissante et converge vers $-\infty$.

2.3.5 Propriétés des suites convergentes

Etudions à présent les propriétés des suites convergentes de $\mathbb{R}^n$ et de leurs limites vis-à-vis des opérations que nous avons introduites entre les points de $\mathbb{R}^n$.

a) Vis-à-vis de la combinaison linéaire

Définitions. Si J appartient à $\mathbb{N}_0$, si $(x_m^{(1)})_{m\in\mathbb{N}_0}, \ldots, (x_m^{(J)})_{m\in\mathbb{N}_0}$ sont des suites de $\mathbb{R}^n$ (resp. de $\mathbb{C}$) et si $r_1, \ldots, r_J$ sont des nombres réels (resp. complexes), la *combinaison linéaire* (resp. la *combinaison linéaire complexe*) correspondante est la suite dont le m-ème élément est égal à $\sum_{j=1}^J r_j x_m^{(j)}$. Elle est notée

$$\left(\sum_{j=1}^J r_j x_m^{(j)}\right)_{m\in\mathbb{N}_0} \quad \text{ou même} \quad \sum_{j=1}^J r_j x_m^{(j)}$$

si aucune ambiguïté n'est possible.

Remarquons immédiatement que, dans cette définition, nous pouvons supposer que tous les coefficients diffèrent de 0.

Voici le résultat principal.

Théorème 2.3.5.1 *Soient $(x_m)_{m\in\mathbb{N}_0}$ et $(y_m)_{m\in\mathbb{N}_0}$ des suites de $\mathbb{R}^n$ et $(r_m)_{m\in\mathbb{N}_0}$ une suite de $\mathbb{R}$.*

a) i) *Si les suites x_m et y_m convergent respectivement vers x et y, alors la suite $x_m + y_m$ converge vers $x + y$.*

a) ii) *Si la suite x_m converge vers x et la suite r_m vers r, alors la suite $r_m x_m$ converge vers rx.*

Dès lors, toute combinaison linéaire de suites convergentes converge vers la combinaison linéaire correspondante des limites.

b) *Si la suite x_m converge vers l'infini et si la suite y_m est bornée, alors la suite $x_m + y_m$ converge vers l'infini.*

c) *Si la suite x_m converge vers l'infini et s'il existe $r > 0$ tel que $|r_m| \geq r$ pour tout $m \in \mathbb{N}_0$, alors la suite $r_m x_m$ converge vers l'infini.*

Preuve. a) i) De fait, pour tout $\varepsilon > 0$, il existe des nombres réels M_1 et M_2 tels que

$$\begin{aligned} m \geq M_1 &\Rightarrow |x_m - x| \leq \varepsilon/2 \\ m \geq M_2 &\Rightarrow |y_m - y| \leq \varepsilon/2 \end{aligned}$$

donc tels que

$$m \geq \sup\{M_1, M_2\} \Rightarrow |x_m + y_m - x - y| \leq \varepsilon.$$

a) ii) Comme la suite r_m converge, elle est bornée: il existe donc $C > 0$ tel que $|r_m| \leq C$ pour tout $m \in \mathbb{N}_0$. Cela étant, pour tout $\varepsilon > 0$, il existe des nombres réels M_1 et M_2 tels que

$$\begin{aligned} m \geq M_1 &\Rightarrow |x_m - x| \leq \varepsilon/(2C) \\ m \geq M_2 &\Rightarrow |r_m - r| \leq \varepsilon/(2\,|x| + 2) \end{aligned}$$

donc tels que, pour tout $m \geq \sup\{M_1, M_2\}$, on a

$$\begin{aligned} |r_m x_m - rx| &= |r_m(x_m - x) + (r_m - r)x| \\ &\leq |r_m|\,|x_m - x| + |r_m - r|\,|x| \;\leq\; \varepsilon. \end{aligned}$$

b) Comme la suite y_m est bornée, il existe $C > 0$ tel que $|y_m| \leq C$ pour tout $m \in \mathbb{N}_0$. Cela étant, pour tout $N > 0$, il existe M tel que

$$m \geq M \Rightarrow |x_m| \geq N + C$$

donc tel que

$$m \geq M \Rightarrow |x_m + y_m| \geq |x_m| - |y_m| \geq N.$$

c) De fait, pour tout $N > 0$, il existe M tel que

$$m \geq M \Rightarrow |x_m| \geq N/r$$

donc tel que

$$m \geq M \Rightarrow |r_m x_m| \geq r\,|x_m| \geq N.\blacksquare$$

Dans $\mathbb{C}$, ce résultat peut être précisé de la manière suivante.

Proposition 2.3.5.2 *L'énoncé précédent s'applique aux suites numériques si on y remplace "$\mathbb{R}^n$" et "$\mathbb{R}$" par "$\mathbb{C}$" ainsi que "combinaison linéaire" par "combinaison linéaire à coefficients complexes".*

Preuve. La même démonstration convient car on a $|z + z'| \leq |z| + |z'|$ ainsi que $|zz'| = |z|\,|z'|$ pour tous z, $z' \in \mathbb{C}$.$\blacksquare$

Remarque. Dans le cas des suites numériques réelles, on en déduit aussitôt de nombreux résultats relatifs à la convergence par la droite ou par la gauche ainsi qu'à la convergence vers $+\infty$ ou vers $-\infty$.

b) Vis-à-vis du produit scalaire

Définitions. Dans $\mathbb{R}^n$ (resp. $\mathbb{C}$), le *produit scalaire* (resp. *produit*) des suites $(x_m)_{m\in\mathbb{N}_0}$ et $(y_m)_{m\in\mathbb{N}_0}$ est la suite numérique réelle (resp. complexe) dont le m-ème élément est égal au produit scalaire $\langle x_m, y_m\rangle$ (resp. au produit $x_m y_m$). Elle est notée

$$(\langle x_m, y_m\rangle)_{m\in\mathbb{N}_0} \quad \text{(resp. } (x_m y_m)_{m\in\mathbb{N}_0})$$

ou même $\langle x_m, y_m\rangle$ (resp. $x_m y_m$) si aucune ambiguïté n'est possible.

Théorème 2.3.5.3 a) *Le produit scalaire de deux suites convergentes de $\mathbb{R}^n$ converge vers le produit scalaire de leurs limites.*

b) *Si les suites $(x_m)_{m\in\mathbb{N}_0}$ et $(y_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$ sont respectivement convergente vers 0 et bornée, alors la suite $\langle x_m, y_m\rangle$ converge vers 0.*

Preuve. a) Soient x_m et y_m des suites convergentes de $\mathbb{R}^n$, de limites respectives x et y. Toute suite convergente étant bornée, il existe $C > 0$ tel que $|y_m| \leq C$ pour tout $m \in \mathbb{N}_0$. Cela étant, pour tout $\varepsilon > 0$, il existe des nombres réels M_1 et M_2 tels que

$$\begin{aligned} m \geq M_1 \quad &\Rightarrow \quad |x_m - x| \leq \varepsilon/(2C) \\ m \geq M_2 \quad &\Rightarrow \quad |y_m - y| \leq \varepsilon/(2\,|x| + 2). \end{aligned}$$

Au total, pour tout $m \geq \sup\{M_1, M_2\}$, on a

$$\begin{aligned} |\langle x_m, y_m\rangle - \langle x, y\rangle| \;&=\; |\langle x_m - x, y_m\rangle + \langle x, y_m - y\rangle| \\ &\leq\; |x_m - x|\,|y_m| + |x|\,|y_m - y| \;\leq\; \varepsilon. \end{aligned}$$

b) Comme la suite y_m est bornée, il existe $C > 0$ tel que $|y_m| \leq C$ pour tout $m \in \mathbb{N}_0$. Cela étant, pour tout $\varepsilon > 0$, il existe M tel que

$$m \geq M \Rightarrow |x_m| \leq \varepsilon/C,$$

donc tel que

$$m \geq M \Rightarrow |\langle x_m, y_m\rangle| \leq |x_m|\,|y_m| \leq \varepsilon,$$

ce qui suffit.∎

Remarque. Dans $\mathbb{R}^n$ avec $n \geq 2$, il n'existe pas de propriété relative à la suite $\langle x_m, y_m\rangle$ si la suite x_m converge vers l'infini. Ainsi, dans $\mathbb{R}^2$, on vérifie directement les résultats suivants:

$$\begin{array}{l|l} a_m = (m,0) \to \infty & e_m = ((-1)^m/m, 0) \to 0 \\ b_m = (m^{-1/2}, 0) \to 0 & f_m = (0,m) \to \infty \\ c_m = (1/m, 0) \to 0 & g_m = ((-1)^m/m, m) \to \infty \\ d_m = (m^{-2}, 0) \to 0 & h_m = (1,m) \to \infty \end{array}$$

alors que

$$\begin{array}{l|l} \langle a_m, b_m\rangle = \sqrt{m} \uparrow +\infty & \langle a_m, f_m\rangle = 0 \to 0 \\ \langle a_m, c_m\rangle = 1 \to 1 & \langle a_m, g_m\rangle = (-1)^m \not\to \\ \langle a_m, d_m\rangle = 1/m \to 0 & \langle a_m, h_m\rangle = m \uparrow +\infty \\ \langle a_m, e_m\rangle = (-1)^m \not\to & \end{array}$$

Passons aux suites numériques.

Proposition 2.3.5.4 a) *Le théorème précédent s'applique aux suites numériques si on y remplace "produit scalaire" par "produit", "$\mathbb{R}^n$" par "$\mathbb{C}$" et "$\langle x_m, y_m \rangle$" par "$x_m y_m$".*

b) *Si J appartient à $\mathbb{N}_0$ et si les suites numériques $z_m^{(1)}$, ..., $z_m^{(J)}$ convergent respectivement vers $z^{(1)}$, ..., $z^{(J)}$, alors la suite numérique de m-ème élément égal à $z_m^{(1)} \cdots z_m^{(J)}$ converge vers $z^{(1)} \cdots z^{(J)}$.*

c) *Si la suite numérique z_m converge vers z_0 et si on a $z_m \neq 0$ pour tout $m \in \mathbb{N}$, alors la suite $1/z_m$ converge vers $1/z_0$.*

d) *Si la suite numérique z_m n'a aucun élément nul, alors elle converge vers 0 si et seulement si la suite $1/z_m$ converge vers l'infini.*

Preuve. a) La même démonstration convient car on a $|zz'| = |z|\,|z'|$ pour tous z, $z' \in \mathbb{C}$.

b) Pour $J = 2$, c'est un cas particulier de a). Un raisonnement par récurrence permet de conclure aussitôt.

c) Pour tout $\varepsilon > 0$, il existe M tel que

$$m \geq M \Rightarrow |z_m - z_0| \leq \inf \left\{ |z_0|/2, \varepsilon\, |z_0|^2/2 \right\}$$

donc tel que

$$m \geq M \Rightarrow \left| \frac{1}{z_m} - \frac{1}{z_0} \right| = \frac{|z_0 - z_m|}{|z_0|\,|z_m|} \leq \frac{\varepsilon\, |z_0|^2}{2} \cdot \frac{1}{|z_0|} \cdot \frac{1}{||z_0| - |z_0 - z_m||} \leq \varepsilon.$$

d) est immédiat.∎

Remarque. Dans le cas des suites numériques réelles, on en déduit aussitôt de nombreux résultats relatifs à la convergence par la droite ou par la gauche, ainsi qu'à la convergence vers $+\infty$ ou vers $-\infty$.

c) Vis-à-vis des grandeurs attachées aux points

Considérons d'abord le recours aux composantes.

Proposition 2.3.5.5 *Dans $\mathbb{R}^n$, la suite x_m converge vers x si et seulement si, pour tout $j \in \{1, \ldots, n\}$, la suite $([x_m]_j)_{m \in \mathbb{N}_0}$ converge vers la j-ème composante de x.*

Preuve. Cela résulte immédiatement des majorations

$$|[x]_j| \leq |x| \leq \sum_{k=1}^{n} |[x]_k|, \quad \forall x \in \mathbb{R}^n, \forall j \in \{1, \ldots, n\}.\blacksquare$$

Remarques. a) En particulier, la suite numérique z_m converge vers z dans $\mathbb{C}$ si et seulement si les suites $\Re z_m$ et $\Im z_m$ convergent respectivement vers $\Re z$ et $\Im z$.

b) Il n'existe pas de propriété générale de ce type pour une suite de $\mathbb{R}^n$ qui converge vers l'infini. Pour s'en convaincre, il suffit de considérer les suites x_m et y_m de $\mathbb{R}^2$ définies par

$$x_m = (m, 0), \; y_{2m} = (2m, 0) \text{ et } y_{2m-1} = (0, 2m-1), \quad \forall m \in \mathbb{N}_0.$$

c) La proposition ramène en fait l'étude de la convergence des suites dans $\mathbb{R}^n$ à celle de n suites dans $\mathbb{R}$.

Passons au module.

Proposition 2.3.5.6 *Soit $(x_m)_{m\in\mathbb{N}_0}$ une suite de $\mathbb{R}^n$.*

a) *Si la suite x_m converge vers x, alors la suite $|x_m|$ converge vers $|x|$.*

b) *La suite x_m converge vers 0 si et seulement si la suite $|x_m|$ converge vers 0.*

c) *La suite x_m converge vers l'infini si et seulement si la suite $|x_m|$ converge vers l'infini.*

Preuve. a) résulte aussitôt de l'inégalité $||x| - |y|| \leq |x - y|$, valable pour tous x, $y \in \mathbb{R}^n$.

b) et c) sont triviaux.∎

Remarque. Si la suite $|x_m|$ converge vers a avec $a \neq 0$, on ne peut rien dire au sujet de la suite x_m. Cela est vrai même dans $\mathbb{R}$, ainsi que l'établit la considération de la suite $(-1)^m a$.

Passons aux suites numériques.

Proposition 2.3.5.7 *La suite numérique z_m converge vers z si et seulement si la suite $\overline{z_m}$ converge vers $\overline{z}$.*∎

Enfin certaines considérations sont particulières aux suites numériques réelles.

Proposition 2.3.5.8 *La suite numérique réelle x_m converge vers x si et seulement si les suites x_{m+} et x_{m-} convergent respectivement vers x_+ et x_-.*

Par suite, si J est un entier supérieur ou égal à 2 et si les suites numériques réelles $x_{1,m}$, ..., $x_{J,m}$ convergent respectivement vers x_1, ..., x_J, alors les suites numériques réelles

$$\sup\{x_{1,m}, \ldots, x_{J,m}\} \quad \textit{et} \quad \inf\{x_{1,m}, \ldots, x_{J,m}\}$$

convergent respectivement vers $\sup\{x_1, \ldots, x_J\}$ *et* $\inf\{x_1, \ldots, x_J\}$.∎

Remarque. Si la suite numérique réelle x_m converge vers l'infini, on ne peut rien affirmer au sujet des suites numériques réelles x_{m+} et x_{m-}. Pour s'en convaincre, il suffit de considérer les suites numériques réelles x_m, y_m et z_m suivantes

$$x_m = m, \; y_m = -m \text{ et } z_m = (-1)^m m, \quad \forall m \in \mathbb{N}_0.$$

d) Vis-à-vis des signes d'égalité et d'inégalité

Proposition 2.3.5.9 *Soient x_m et y_m des suites numériques réelles qui convergent respectivement vers x et y.*

a) *Si on a $x \neq y$, alors il existe M tel que $x_m \neq y_m$ pour tout $m \geq M$.*
Dès lors, si x diffère de a, il existe M tel que $x_m \neq a$ pour tout $m \geq M$.

b) *Si on a $x < y$, alors il existe M tel que $x_m < y_m$ pour tout $m \geq M$.*
En particulier, si on a $x < a$ (resp. $x > a$)*, il existe M tel que $x_m < a$* (resp. $x_m > a$) *pour tout $m \geq M$.*

Preuve. a) De fait, il existe M tel que

$$m \geq M \Rightarrow |x_m - x| \leq \tfrac{1}{3}|x - y| \quad \text{et} \quad |y_m - y| \leq \tfrac{1}{3}|x - y|,$$

donc tel que

$$m \geq M \Rightarrow |x_m - y_m| \geq |x - y| - |x - x_m| - |y - y_m| \geq \tfrac{1}{3}|x - y|.$$

Le cas particulier résulte aussitôt de la considération de la suite constante $y_m = a$ pour tout $m \in \mathbb{N}_0$.

b) Il existe en effet M tel que

$$m \geq M \Rightarrow |x_m - x| \leq \tfrac{1}{3}|x - y| \quad \text{et} \quad |y_m - y| \leq \tfrac{1}{3}|x - y|,$$

donc tel que, par exemple dans le cas $x < y$, $m \geq M$ implique $x_m < y_m$.

Le cas particulier résulte aussitôt de la considération de la suite constante $y_m = a$ pour tout $m \in \mathbb{N}_0$.∎

Proposition 2.3.5.10 a) *Si la suite numérique réelle x_m converge vers x et si, pour tout $m \in \mathbb{N}_0$, on a $x_m <$* (resp. $\leq, >, \geq$) *a, alors on a $x \leq$* (resp. $\leq, \geq, \geq$) a.

b) *Si la suite numérique réelle x_m converge vers $+\infty$* (resp. $-\infty$) *et si la suite numérique réelle y_m est telle que $x_m \leq$* (resp. $\geq$) *y_m pour tout $m \in \mathbb{N}_0$, alors on a $y_m \to +\infty$* (resp. $-\infty$)*.*

Preuve. a) est direct par contradiction, vu la partie b) de la proposition précédente.

b) est trivial.∎

Remarque. On recourt bien souvent à la partie a) de cette dernière proposition selon la formule "*lors d'un passage à la limite dans $\mathbb{R}$, les signes d'égalité se maintiennent avec établissement éventuel du signe d'égalité*".

Théorème 2.3.5.11 (étau) *Si les suites numériques réelles x_m et y_m convergent vers a et si la suite numérique réelle z_m vérifie $x_m \leq z_m \leq y_m$ pour tout $m \in \mathbb{N}_0$, alors la suite z_m converge vers a.*

Preuve. Cela résulte aussitôt des majorations évidentes

$$|z_m - a| \leq \sup\{|x_m - a|, |y_m - a|\}, \quad \forall m \in \mathbb{N}_0. \blacksquare$$

2.3.6 Suites de Cauchy

Position du problème. Soit $(x_m)_{m\in\mathbb{N}_0}$ une suite de $\mathbb{R}^n$. Comment faire pour déterminer si cette suite converge? Recourir à la définition de la convergence d'une suite est utopique car, ne connaissant pas la limite éventuelle a priori, on devrait tester tous les points de $\mathbb{R}^n$ pour vérifier s'il y en a un vers lequel la suite converge! Il convient plutôt d'appliquer des critères de convergence. L'objet de ce paragraphe est d'établir de tels critères.

Une première réponse est donnée par le résultat suivant.

Théorème 2.3.6.1 *Toute suite numérique réelle, croissante et majorée converge vers la borne supérieure de l'ensemble de ses éléments.*

De même, toute suite numérique réelle, décroissante et minorée converge vers la borne inférieure de l'ensemble de ses éléments.

Preuve. Considérons d'abord le cas d'une suite x_m numérique réelle, croissante et majorée. Désignons par x la borne supérieure de l'ensemble de ses éléments. Tout revient à démontrer que $x_m \to x$. Or, pour tout $\varepsilon > 0$, vu le critère 2.2.7.3 relatif aux bornes supérieure et inférieure, il existe un élément de $\{x_m : m \in \mathbb{N}_0\}$, c'est-à-dire un élément x_M de la suite, tel que $|x_M - x| \leq \varepsilon$. On en déduit aussitôt que

$$m \geq M \Rightarrow |x_m - x| = x - x_m \leq x - x_M \leq \varepsilon,$$

ce qui suffit.

Le cas d'une suite numérique réelle, décroissante et minorée x_m se traite de même. On peut cependant aussi procéder comme suit: la suite numérique réelle $-x_m$ est croissante et majorée donc converge vers la borne supérieure de $\{-x_m : m \in \mathbb{N}_0\}$ qui est l'opposé de la borne inférieure de l'ensemble $\{x_m : m \in \mathbb{N}_0\}$. D'où la conclusion. $\blacksquare$

Voici également un lemme théorique relatif aux suites numériques réelles.

Lemme 2.3.6.2 *Si une sous-suite d'une suite numérique réelle croissante* (resp. *décroissante*) *converge, alors la suite converge.*

Preuve. Soit x_m une suite numérique réelle croissante dont la sous-suite $x_{k(m)}$ converge. Comme toute suite convergente est bornée, il existe C tel que $x_{k(m)} \leq C$ pour tout $m \in \mathbb{N}_0$. Cela étant, la suite x_m est majorée car on a alors $x_m \leq x_{k(m)} \leq C$ pour tout $m \in \mathbb{N}_0$. D'où la conclusion par la proposition précédente.

Dans le cas d'une suite numérique réelle décroissante, on peut soit procéder de manière semblable, soit remarquer que la suite $-x_m$ est croissante.∎

Une deuxième réponse, fondamentale, repose sur la notion de suite de Cauchy.

Définition. Une suite $(x_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$ *est de Cauchy* si, pour tout $\varepsilon > 0$, il existe un nombre réel M tel que

$$p, q \geq M \Rightarrow |x_p - x_q| \leq \varepsilon.$$

Le résultat suivant est à la base du critère de Cauchy.

Proposition 2.3.6.3 *Si une suite de Cauchy admet une sous-suite convergente, elle converge.*

Preuve. Soit $x_{k(m)}$ une sous-suite convergente d'une suite de Cauchy x_m; désignons par x sa limite. Pour tout $\varepsilon > 0$, il existe donc des nombres M_1 et M_2 tels que

$$\begin{array}{lcl} m \geq M_1 & \Rightarrow & \left|x_{k(m)} - x\right| \leq \varepsilon/2 \\ p, q \geq M_2 & \Rightarrow & |x_p - x_q| \leq \varepsilon/2. \end{array}$$

D'où la conclusion car on a alors

$$m \geq \sup\{M_1, M_2\} \Rightarrow |x_m - x| \leq \left|x_m - x_{k(m)}\right| + \left|x_{k(m)} - x\right| \leq \varepsilon.$$ ∎

Critère 2.3.6.4 (Cauchy) *Une suite de $\mathbb{R}^n$ converge si et seulement si elle est de Cauchy.*

Preuve. La condition est nécessaire. De fait, si la suite x_m de $\mathbb{R}^n$ converge vers x, pour tout $\varepsilon > 0$, il existe M tel que

$$m \geq M \Rightarrow |x_m - x| \leq \varepsilon/2,$$

donc tel que

$$p, q \geq M \Rightarrow |x_p - x_q| \leq |x_p - x| + |x - x_q| \leq \varepsilon.$$

La condition est suffisante. Vu la proposition précédente, il suffit, étant donné une suite x_m de Cauchy dans $\mathbb{R}^n$, d'en extraire une sous-suite convergente.

Donnons d'abord une construction d'une telle sous-suite. Pour $k(1)$, on prend le premier entier $k \in \mathbb{N}_0$ tel que

$$p, q \geq k \Rightarrow |x_p - x_q| \leq 10^{-1}.$$

Ensuite, on construit les $k(m)$ au moyen de la récurrence suivante: si $k(1), \ldots, k(m-1)$ sont déterminés, on prend pour $k(m)$ le premier entier k strictement supérieur à $k(m-1)$ tel que

$$p, q \geq k \Rightarrow |x_p - x_q| \leq 10^{-m}.$$

Il est bien certain que la suite $x_{k(m)}$ ainsi construite est une sous-suite de la suite x_m.

Prouvons à présent que cette sous-suite converge. Il suffit pour cela d'établir que, pour tout $j \in \{1, \ldots, n\}$, la suite $[x_{k(m)}]_j$ des j-èmes composantes converge. Or, pour tout $m \in \mathbb{N}_0$, on a

$$[x_{k(m)}]_j = [x_{k(m)} - x_{k(m-1)}]_j + \ldots + [x_{k(2)} - x_{k(1)}]_j + [x_{k(1)}]_j.$$

De plus, la suite $S_{m,j,+}$ définie par

$$S_{m,j,+} = [x_{k(m)} - x_{k(m-1)}]_{j+} + \ldots + [x_{k(2)} - x_{k(1)}]_{j+} + [x_{k(1)}]_{j+}$$

pour tout $m \in \mathbb{N}_0$ est croissante et majorée par $\left|x_{k(1)}\right| + 1$, comme on le vérifie aisément, elle est donc convergente. De même, la suite $S_{m,j,-}$ définie par

$$S_{m,j,-} = [x_{k(m)} - x_{k(m-1)}]_{j-} + \ldots + [x_{k(2)} - x_{k(1)}]_{j-} + [x_{k(1)}]_{j-}$$

pour tout $m \in \mathbb{N}_0$ est croissante et majorée par $\left|x_{k(1)}\right| + 1$ donc converge. Au total, la suite

$$S_{m,j,+} - S_{m,j,-} = [x_{k(m)}]_j,$$

converge. D'où la conclusion.∎

Remarques. a) Dans $\mathbb{R}^n$, il y a donc identification entre les suites convergentes et les suites de Cauchy. C'est la raison pour laquelle nous n'avons pas cherché à établir les propriétés des suites de Cauchy.

b) Il convient de remarquer qu'en général, $S_{m,j,+}$ (resp. $S_{m,j,-}$) n'est pas la partie positive (resp. négative) de $[x_{k(m)}]$.

Exercice. Etablir le *principe du choix de Bolzano-Weierstrass:*, à savoir la propriété suivante: *toute suite de* $\mathbb{R}$ *contient une sous-suite monotone.* En particulier, *de toute suite bornée de* $\mathbb{R}$*, on peut extraire une sous-suite convergente.*

En déduire le critère de Cauchy dans $\mathbb{R}$.

Suggestion. Soit r_m une suite de $\mathbb{R}$. Appelons *pivot* de cette suite tout entier $m \in \mathbb{N}_0$ tel que $r_n < r_m$ pour tout $n > m$. S'il y a une suite de pivots, nous pouvons les noter $x_{k(m)}$ avec $k(m) < k(m+1)$ pour tout $m \in \mathbb{N}_0$ et la suite $r_{k(m)}$ ainsi obtenue est strictement monotone. S'il n'y a qu'un nombre fini de pivots, il existe $k(1) \in \mathbb{N}$ strictement plus grand que chacun des pivots. Dès lors, $k(1)$ n'est pas un pivot et il existe $k(2) > k(1)$ tel que $r_{k(2)} \geq r_{k(1)}$. Comme $k(2)$ n'est pas un pivot, il existe $k(3) > k(2)$ tel que $r_{k(3)} \geq r_{k(2)}$. Et ainsi de suite. Le cas particulier résulte aussitôt du théorème 2.3.6.1.

Le critère de Cauchy résulte aussitôt de la proposition 2.3.6.3.∎

2.3.7 Exemples d'étude de la convergence de suites

Exercice. Etudier la convergence de la suite numérique réelle de m-ème élément égal à

$$r_m = \frac{1}{m+1} + \cdots + \frac{1}{2m} = \sum_{j=1}^{m} \frac{1}{m+j}.$$

Suggestion. D'une part, cette suite est strictement croissante car, pour tout $m \in \mathbb{N}_0$,

$$\begin{aligned} r_{m+1} - r_m &= \frac{1}{2m+1} + \frac{1}{2m+2} - \frac{1}{m+1} \\ &= \frac{(2m+2)+(2m+1)-(4m+2)}{2(m+1)(2m+1)} > 0. \end{aligned}$$

D'autre part, elle est majorée par 1 car on a

$$r_m = \frac{1}{m+1} + \cdots + \frac{1}{m+m} \leq m\,\frac{1}{m} = 1.$$

Vu le théorème 2.3.6.1, cette suite converge.

Exemple. (Série harmonique) *Etudier la convergence de la suite numérique réelle de m-ème élément égal à*

$$r_m = 1 + \frac{1}{2} + \cdots + \frac{1}{m} = \sum_{j=1}^{m} \frac{1}{j}.$$

Comme on a

$$r_{2m} - r_m = \frac{1}{m+1} + \cdots + \frac{1}{2m} \geq m \cdot \frac{1}{2m} = \frac{1}{2},$$

pour $\varepsilon = 1/3$, il n'existe pas d'entier M tel que $|r_p - r_q| \leq \varepsilon$ pour tous $p, q \geq M$. La suite n'est donc pas de Cauchy et, par conséquent, elle ne converge pas. On peut aussi remarquer que la suite considérée n'est pas bornée car on a

$$r_{2^m} = \sum_{j=1}^{m} (r_{2^j} - r_{2^{j-1}}) + r_1 \geq r_1 + \frac{m}{2}, \quad \forall m \in \mathbb{N}_0.$$

Théorème 2.3.7.1 (Cesaro) *Si la suite x_m converge vers x, la suite X_m définie par*

$$X_m = \frac{1}{m}(x_1 + \cdots + x_m) = \frac{1}{m}\sum_{j=1}^{m} x_j$$

converge également vers x.

Preuve. Pour tout $\varepsilon > 0$, il existe M tel que $|x_m - x| \leq \varepsilon/2$ pour tout $m \geq M$ donc tel que

$$m \geq M \Rightarrow \frac{1}{m}\sum_{j=1}^{M} |x_j - x| + \frac{1}{m}\sum_{j=M+1}^{m} |x_j - x| \leq \frac{1}{m}\sum_{j=1}^{M} |x_j - x| + \frac{\varepsilon}{2}.$$

On vérifie alors aisément qu'il existe $M' \geq M$ tel que

$$m \geq M' \Rightarrow \frac{1}{m}\sum_{j=1}^{M} |x_j - x| \leq \frac{\varepsilon}{2}.$$

Au total, pour tout $\varepsilon > 0$, il existe M' tel que $|X_m - x| \leq \varepsilon$ pour tout $m \geq M'$, ce qui suffit.∎

Exercice. *Etablir que, pour tout $a > 0$, la suite r_m définie par récurrence par*

$$r_0 > 0 \quad \textit{et} \quad r_{m+1} = \tfrac{1}{2}(r_m + a/r_m), \quad \forall m \in \mathbb{N},$$

converge vers $\sqrt{a}$.

Suggestion. Il est clair qu'on a $r_m > 0$ pour tout $m \in \mathbb{N}$. De plus, de l'inégalité $\sqrt{rs} \leq (r+s)/2$ valable pour tous r, $s > 0$, on tire

$$\sqrt{a} = \sqrt{r_m \cdot \frac{a}{r_m}} \leq \tfrac{1}{2}(r_m + a/r_m) = r_{m+1}, \quad \forall m \in \mathbb{N}.$$

Dès lors, pour tout $m \in \mathbb{N}_0$, on a $r_m^2 \geq a$ et par conséquent $a/r_m \leq r_m$. Au total, la suite r_m est définie, décroissante à partir de r_1 et minorée par $\sqrt{a}$, donc converge vu le théorème 2.3.6.1. Comme sa limite r doit vérifier $r \geq 0$ et $r = \frac{1}{2}(r + a/r)$, on a $r = \sqrt{a}$.

Exercice. *Si la suite a_m de $[0, +\infty[$ vérifie $a_{m+n} \leq a_m a_n$ pour tous m, $n \in \mathbb{N}_0$, établir que*

$$(a_m)^{1/m} \to \inf\left\{\, (a_k)^{1/k} : k \in \mathbb{N}_0 \right\}.$$

Suggestion. Posons $a = \inf\left\{\, (a_k)^{1/k} : k \in \mathbb{N}_0 \right\}$. Pour tout $\varepsilon > 0$, il existe bien sûr $m \in \mathbb{N}_0$ tel que $a_m \leq (a + \varepsilon)^m$. Posons $A = \sup\{a_1, \ldots, a_m\}$. Comme, pour tout $n \in \mathbb{N}_0$, il existe p_n, $q_n \in \mathbb{N}$ tels que $n = mp_n + q_n$ avec $0 \leq q_n < m$, il vient successivement

$$a_n = a_{mp_n+q_n} \leq a_{mp_n} a_{q_n} \leq (a + \varepsilon)^{mp_n} a_{q_n} \leq (a + \varepsilon)^{mp_n} A$$

donc

$$a \leq (a_n)^{1/n} \leq (a + \varepsilon)^{mp_n/n} A^{1/n} = (a + \varepsilon) A^{1/n} (a + \varepsilon)^{-q_n/n}.$$

Or la suite $A^{1/n}(a+\varepsilon)^{-q_n/n}$ converge vers 1 (on a en effet $r^{1/n} \to 1$ pour tout $r > 0$). Par conséquent, il vient $a \leq (a_n)^{1/n} \leq a + 2\varepsilon$ pour n suffisamment grand, ce qui permet de conclure.

$* \to$ Dès lors, pour tout élément a d'une algèbre normée $(A, \|\cdot\|)$, la suite $\|a^m\|^{1/m}$ converge vers $\inf\left\{\, \|a^k\|^{1/k} : k \in \mathbb{N}_0 \right\}$. $\leftarrow *$

2.4 Topologie de $\mathbb{R}^n$

2.4.1 Ouverts et fermés

Définition. Une partie Ω de $\mathbb{R}^n$ est *ouverte* (on dit aussi "*est un ouvert*") si tout point de Ω est le centre d'une boule incluse dans Ω ou encore si Ω est voisinage de chacun de ses points.

Exemples. a) Bien sûr, *$\mathbb{R}^n$ et $\emptyset$ sont des ouverts de $\mathbb{R}^n$.*

b) *Les seuls intervalles ouverts de $\mathbb{R}$ sont les intervalles des types suivants: $]a, b[$, $]-\infty, b[$, $]a, +\infty[$ et $]-\infty, +\infty[$.* On vérifie aisément que ces intervalles sont ouverts. De plus, aucun autre intervalle de $\mathbb{R}$ n'est ouvert: ainsi, par exemple, l'intervalle $[a, b]$ n'est pas ouvert car a y appartient et n'est le centre d'aucune boule incluse dans cet intervalle.

c) *Un intervalle $I = I_1 \times \cdots \times I_n$ de $\mathbb{R}^n$ est ouvert si et seulement si chacun de ses intervalles constitutifs $I_1, \ldots, I_n$ est ouvert dans $\mathbb{R}$.* La condition est nécessaire. Pour tout $j \in \{1, \ldots, n\}$ et tout $a \in I_j$, il existe un point $x \in I$ dont la j-ème composante est égale à a. Il existe ensuite $r > 0$ tel que $\{\, y : |x - y| \leq r\} \subset I$, donc tel que

$$\{\, b \in \mathbb{R} : |a - b| \leq r\} = \{\, [y]_j : |x - y| \leq r\} \subset I_j.$$

La condition est suffisante. Soit x un point de I. Pour tout $j \in \{1, \ldots, n\}$, on a $x_j \in I_j$ et il existe donc $r_j > 0$ tel que $\{ a \in \mathbb{R} : |x_j - a| \leq r_j \} \subset I_j$. Cela étant, on obtient de suite

$$\left\{ y : |x - y| \leq \inf_{1 \leq j \leq n} r_j \right\} \subset I.$$

d) *Pour tout $x \in \mathbb{R}^n$ et tout $r > 0$, la boule $b = \{ y : |x - y| < r \}$ est ouverte.* De fait, si y appartient à cette boule, on vérifie de suite que

$$\{ z : |y - z| < r - |x - y| \} \subset b.$$

Par contre, *la boule $b' = \{ y : |x - y| \leq r \}$ n'est pas ouverte.* Il suffit de remarquer que le point $x + re_1$ appartient assurément à b' et n'est le centre d'aucune boule incluse dans b': pour tout $r' > 0$, le point $x + re_1 + r'e_1$ appartient à $\{ y : |x + re_1 - y| \leq r' \}$ et n'appartient pas à b'.

e) *Pour toute partie non vide A de $\mathbb{R}^n$ et tout $r > 0$, l'ensemble*

$$A_r = \{ x : \mathrm{d}(x, A) < r \}$$

est un ouvert contenant A. Bien sûr, cet ensemble A_r contient A. De plus, il s'agit d'un ouvert car, pour tout $x \in A_r$, on a $\mathrm{d}(x, A) = r_x$ avec $r_x < r$ donc $\{ y : |x - y| < r - r_x \} \subset A_r$, comme on le vérifie de suite au moyen de l'inégalité triangulaire relative à la distance entre parties de $\mathbb{R}^n$.

Définition. Une partie de $\mathbb{R}^n$ est *fermée* (on dit aussi "*est un fermé*") si son complémentaire dans $\mathbb{R}^n$ est ouvert.

Dans $\mathbb{R}^n$, on peut donner une définition directe (c'est-à-dire qui ne recourt pas au complémentaire) des parties fermées grâce à la caractérisation suivante.

Théorème 2.4.1.1 *Une partie F de $\mathbb{R}^n$ est fermée si et seulement si elle contient la limite de toutes ses suites convergentes.*

Preuve. La condition est nécessaire. Soit x_m une suite de F qui converge vers x. Si x n'appartient pas au fermé F, il appartient à l'ouvert $\mathbb{R}^n \setminus F$. Il existe alors $r > 0$ tel que $\{ y : |x - y| \leq r \} \subset \mathbb{R}^n \setminus F$. Comme la suite x_m converge vers x, il existe M tel que $|x_m - x| \leq r$ donc tel que $x_m \in \mathbb{R}^n \setminus F$ pour tout $m \geq M$. D'où une contradiction.

La condition est suffisante. Si $\mathbb{R}^n \setminus F$ n'est pas ouvert, il existe $x \in \mathbb{R}^n \setminus F$ tel qu'aucune boule centrée en x ne soit incluse dans $\mathbb{R}^n \setminus F$. Cela étant, pour tout $m \in \mathbb{N}_0$, l'ensemble $A_m = F \cap \{ y : |x - y| \leq 1/m \}$ n'est pas vide. Vu le troisième moyen d'obtention d'une suite, il existe alors une suite x_m telle que $x_m \in A_m$ pour tout $m \in \mathbb{N}_0$. D'où une contradiction car il s'agit d'une suite de F qui converge vers x.∎

Exemples. a) Bien sûr, *$\mathbb{R}^n$ et $\emptyset$ sont des fermés de $\mathbb{R}^n$.*

b) *Les seuls intervalles fermés de $\mathbb{R}$ sont les intervalles des types suivants: $[a, b]$, $]-\infty, b]$, $[a, +\infty[$ et $]-\infty, +\infty[$.* On vérifie directement que ces intervalles sont fermés, au moyen des propriétés des suites numériques réelles convergentes vis-à-vis des signes d'inégalité. De plus, aucun autre intervalle n'est fermé: ainsi, par exemple, l'intervalle $I =]a, b[$ n'est pas fermé car $(x_m = a + (b - a)/(m + 1))_{m \in \mathbb{N}_0}$ est une suite de I qui converge vers a.

c) *Un intervalle $I = I_1 \times \cdots \times I_n$ de $\mathbb{R}^n$ est fermé si et seulement si chacun de ses intervalles constitutifs est fermé dans $\mathbb{R}$.* La condition est nécessaire. Soient $j \in \{1, \ldots, n\}$ et a_m une suite de I_j qui converge vers a. Pour tout $k \in \{1, \ldots, n\}$ différent de j, choisissons un point b_k de I_k et définissons la suite x_m de $\mathbb{R}^n$ par

$$x_m = (b_1, \ldots, b_{j-1}, a_m, b_{j+1}, \ldots, b_n).$$

Il s'agit évidemment d'une suite de I qui converge vers

$$x = (b_1, \ldots, b_{j-1}, a, b_{j+1}, \ldots, b_n).$$

De là, a appartient à I_j. La condition est suffisante. Soit x_m une suite de I qui converge vers x. Pour tout $j \in \{1, \ldots, n\}$, on sait alors que la suite $[x_m]_j$ est une suite I_j qui converge vers $[x]_j$. Si chacun des I_j est fermé, on a $[x]_j \in I_j$ pour tout j et par conséquent $x \in I$.

d) *Pour tout $x \in \mathbb{R}^n$ et tout $r > 0$, la boule $b = \{\, y : |x - y| \leq r\}$ est fermée.* De fait, si la suite y_m de b converge vers y, la suite $y_m - x$ converge vers $y - x$. On a donc $|y_m - x| \to |y - x|$ avec $|y_m - x| \leq r$ pour tout $m \in \mathbb{N}_0$, ce qui suffit.

Par contre, *la boule $b' = \{\, y : |y - x| < r\}$ n'est pas fermée.* De fait, on vérifie directement que $(y_m = x + (1 - 1/m)re_1)_{m \in \mathbb{N}_0}$ est une suite de b' qui converge vers le point $x + re_1$ qui n'appartient pas à b'.

e) *Toute partie finie de $\mathbb{R}^n$ est fermée.*

f) *Pour toute partie A de $\mathbb{R}^n$ différente de $\mathbb{R}^n$ et tout $r > 0$, l'ensemble*

$$A_{-r} = \{\, x : \mathrm{d}(x, \mathbb{R}^n \setminus A) \geq r\}$$

est un fermé inclus dans A. Il suffit de remarquer que $A_{-r} = \mathbb{R}^n \setminus (\mathbb{R}^n \setminus A)_r$.

Passons à présent aux propriétés des ouverts et des fermés.

Proposition 2.4.1.2 *Si le point x n'appartient pas au fermé F de $\mathbb{R}^n$, il existe des ouverts disjoints Ω_x et Ω_F contenant respectivement x et F.* (On dit que "$\mathbb{R}^n$ est un espace *régulier*".)

En particulier, si x et y sont deux points distincts de $\mathbb{R}^n$, il existe des ouverts disjoints Ω_x et Ω_y contenant respectivement x et y. (On dit que "$\mathbb{R}^n$ est un espace *séparé*".)

Preuve. Comme x appartient à l'ouvert $\mathbb{R}^n \setminus F$, il existe $r > 0$ tel que la boule $\{ y : |x-y| \leq r \}$ soit disjointe de F. Cela étant, on peut prendre

$$\Omega_x = \{ y : |x-y| < r/2 \} \quad \text{et} \quad \Omega_F = \{ y : \mathrm{d}(y, F) < r/2 \}.$$

De fait, nous savons que Ω_x est un ouvert contenant x et que Ω_F est un ouvert contenant F. De plus, Ω_x et Ω_F sont disjoints: l'existence d'un point y appartenant à l'intersection de ces deux ensembles conduirait en effet à la contradiction

$$\mathrm{d}(x, F) \leq |x-y| + \mathrm{diam}(\{y\}) + \mathrm{d}(y, F) < r.$$

Le cas particulier résulte aussitôt de ce que $\{y\}$ est fermé. On peut aussi vérifier directement que les boules ouvertes

$$\Omega_x = \left\{ z : |x-z| < \tfrac{1}{2} |x-y| \right\} \quad \text{et} \quad \Omega_y = \left\{ z : |y-z| < \tfrac{1}{2} |x-y| \right\}$$

conviennent. ∎

Remarque. Au paragraphe 4.2.1, nous établirons même que si F_1 et F_2 sont des fermés disjoints de $\mathbb{R}^n$, il existe des ouverts disjoints Ω_{F_1} et Ω_{F_2} contenant respectivement F_1 et F_2. (On dit que "$\mathbb{R}^n$ est un espace *normal*".)

Théorème 2.4.1.3 a) *Toute union d'ouverts est ouverte.*
b) *Toute intersection de fermés est fermée.*

Preuve. a) De fait, si un point appartient à une union d'ouverts, il appartient à l'un d'entre eux et est donc le centre d'une boule incluse dans cet ouvert donc dans l'union des ouverts.
b) s'obtient directement par passage aux complémentaires dans a), vu la formule

$$\mathbb{R}^n \setminus \left(\bigcap_{j \in J} A_j \right) = \bigcup_{j \in J} (\mathbb{R}^n \setminus A_j). \blacksquare$$

Théorème 2.4.1.4 a) *Toute intersection finie d'ouverts est ouverte.*
b) *Toute union finie de fermés est fermée.*

Preuve. a) Soient $\Omega_1, \ldots, \Omega_J$ des ouverts en nombre fini. Si un point x appartient à leur intersection, il appartient à chacun d'entre eux. Dès lors, pour tout $j \in \{1, \ldots, J\}$, il existe $r_j > 0$ tel que la boule $\{ y : |x-y| \leq r_j \}$ soit incluse dans Ω_j. Au total, $r = \inf\{r_1, \ldots, r_J\}$ est un nombre strictement positif tel que la boule $\{ y : |x-y| \leq r \}$ soit incluse dans chacun des Ω_j donc dans leur intersection.
b) s'obtient directement par passage aux complémentaires dans a). ∎

2.4.2 Intérieur, adhérence, frontière

Remarquons de suite qu'il existe des parties de $\mathbb{R}^n$ qui ne sont ni ouvertes, ni fermées. Il en est ainsi, par exemple, des intervalles des types $]a,b]$ et $[a,b[$ dans $\mathbb{R}$.

Dans ce paragraphe, à toute partie A de $\mathbb{R}^n$, nous allons associer un ouvert et un fermé particuliers qui permettent de "cerner" A.

Définitions. Un point x de $\mathbb{R}^n$ est *intérieur à* la partie A de $\mathbb{R}^n$ s'il est le centre d'une boule incluse dans A.

L'*intérieur de* A est l'ensemble des points intérieurs à A; il est noté A° ou $(A)^\circ$.

Proposition 2.4.2.1 *L'intérieur de A est un ouvert inclus dans A et contient tout ouvert inclus dans A.* (On dit que A° est "*le plus grand ouvert*" inclus dans A.)

En particulier, on a $A = A^\circ$ si et seulement si A est ouvert.

Preuve. Il est clair que A° est inclus dans A et il est immédiat que A° contient tout ouvert inclus dans A. Pour conclure, il suffit alors de prouver que A° est un ouvert. Or, pour tout $x \in A^\circ$, il existe $r > 0$ tel que $\{ y : |x-y| \leq r \} \subset A$ donc tel que l'ouvert $\{ y : |x-y| < r \}$ soit inclus dans A donc dans A°, ce qui suffit.∎

Exemples. On vérifie aisément les égalités

$$[a,b]^\circ =]a,b[\quad \text{et} \quad \{ y : |x-y| \leq r \}^\circ = \{ y : |x-y| < r \} .$$

Définitions. Un point x de $\mathbb{R}^n$ est *adhérent à* la partie A de $\mathbb{R}^n$ si toute boule de centre x est d'intersection non vide avec A.

L'*adhérence de* A est l'ensemble des points adhérents à A; il est noté $\overline{A}$ ou A^- ou $(A)^-$.

Voici une caractérisation interne de l'adhérence de A.

Proposition 2.4.2.2 *Dans $\mathbb{R}^n$, un point x appartient à l'adhérence de A si et seulement s'il existe une suite de A qui converge vers x.*

Preuve. La condition est nécessaire. De fait, si x est adhérent à A, alors, pour tout $m \in \mathbb{N}_0$, l'ensemble $A_m = A \cap \{ y : |x-y| \leq 1/m \}$ n'est pas vide. Il existe donc une suite x_m telle que $x_m \in A_m$ pour tout $m \in \mathbb{N}_0$. D'où la conclusion car il s'agit d'une suite de A qui converge vers x.

La condition est évidemment suffisante.∎

Proposition 2.4.2.3 *L'adhérence de A est un fermé contenant A et inclus dans tout fermé contenant A.* (On dit que A^- est "*le plus petit fermé*" contenant A.)

En particulier, on a $A = A^-$ si et seulement si A est fermé.

Preuve. Il est clair que A est inclus dans A^-. De plus, vu la proposition précédente, il est immédiat que A^- est inclus dans tout fermé contenant A. Pour conclure, il suffit donc de prouver que A^- est fermé. Cela résulte aussitôt de la formule

$$\mathbb{R}^n \setminus \overline{A} = (\mathbb{R}^n \setminus A)^\circ$$

qui exprime tout simplement que $x \in \mathbb{R}^n$ n'appartient pas à A^- si et seulement s'il existe $r > 0$ tel que la boule $\{\, y : |x - y| \leq r\}$ soit incluse dans $\mathbb{R}^n \setminus A$.∎

Exemples. On vérifie aisément les égalités

$$]a, b[^- = [a, b] \quad \text{et} \quad \{\, y : |x - y| < r\}^- = \{\, y : |x - y| \leq r\}\,.$$

Proposition 2.4.2.4 *Pour toute partie A et tout ouvert Ω de $\mathbb{R}^n$, on a $(\Omega \cap A)^- = (\Omega \cap \overline{A})^-$.*

Preuve. L'inclusion $(\Omega \cap A)^- \subset (\Omega \cap \overline{A})^-$ résulte aussitôt de ce que $(\Omega \cap \overline{A})^-$ est un fermé contenant $\Omega \cap A$.

Inversement, soit x un élément de $(\Omega \cap \overline{A})^-$. Il existe donc une suite x_m de $\Omega \cap \overline{A}$ qui converge vers x. Pour tout $m \in \mathbb{N}_0$, il existe $r_m > 0$ tel que $\{\, y : |x_m - y| \leq r_m\}$ soit inclus dans Ω et nous pouvons supposer que la suite r_m converge vers 0. Cela étant, pour tout $m \in \mathbb{N}_0$, $A_m = A \cap \{\, y : |x_m - y| \leq r_m\}$ n'est pas vide et il existe donc une suite y_m telle que $y_m \in A_m$ pour tout $m \in \mathbb{N}_0$. D'où la conclusion car il s'agit d'une suite de $\Omega \cap A$ qui converge vers x.∎

Remarque. Pour toute partie A de $\mathbb{R}^n$, nous venons d'introduire le plus grand ouvert A° et le plus petit fermé A^- tels que $A^\circ \subset A \subset A^-$. Pour savoir à quel point ces inclusions sont fines, il suffit de considérer l'ensemble $A^- \setminus A^\circ$, c'est-à-dire l'ensemble des points frontière introduit ci-après.

Définitions. Un point x de $\mathbb{R}^n$ est *frontière de A* si toute boule de centre x est d'intersection non vide avec A et avec $\mathbb{R}^n \setminus A$.

La *frontière de A* est l'ensemble des points frontière de A; il est noté $A^\bullet$ ou $(A)^\bullet$. Il est clair que $A^\bullet = A^- \setminus A^\circ$.

Exemples. On vérifie aisément les égalités

$$]a, b[^\bullet = [a, b]^\bullet = \{a, b\}$$

et

$$\{\, y : |x - y| < r\}^\bullet = \{\, y : |x - y| \leq r\}^\bullet = \{\, y : |x - y| = r\}\,.$$

Proposition 2.4.2.5 *Pour toute partie A de $\mathbb{R}^n$, $A^\bullet = A^- \cap (\mathbb{R}^n \setminus A)^-$.*
En particulier, la frontière de A est fermée et $A^\bullet = (\mathbb{R}^n \setminus A)^\bullet$.∎

Proposition 2.4.2.6 *La frontière d'un ouvert* (resp. *d'un fermé*) *est un ensemble d'intérieur vide.*

Preuve. Comme on a $A^\bullet = (\mathbb{R}^n \setminus A)^\bullet$ pour toute partie A de $\mathbb{R}^n$, il suffit d'établir la proposition dans le cas d'un ouvert Ω. Procédons par l'absurde. Supposons que x soit un point intérieur à $\Omega^\bullet$. Il existe alors $r > 0$ tel que la boule $\{ y : |x - y| \leq r\}$ soit incluse dans $\Omega^\bullet$. De plus, il doit exister un point z de Ω tel que $|x - z| \leq r/2$. Or un tel point z ne peut appartenir à $\Omega^\bullet$ puisqu'il est le centre d'une boule incluse dans Ω. D'où la conclusion.∎

Remarques. Il existe de nombreuses relations entre les ensembles A, A°, A^- et $A^\bullet$. Nous n'allons en citer que quelques unes, laissées comme exercices.

a) Pour toute partie A de $\mathbb{R}^n$, établir que A° et $A^\bullet$ sont disjoints et d'union égale à A^-. En particulier, on a $A^\bullet \subset A$ si et seulement si A est fermé. De même, on a $A^\bullet \subset \mathbb{R}^n \setminus A$ si et seulement si A est ouvert.

b) Pour toute partie A de $\mathbb{R}^n$, on a les égalités

$$\mathbb{R}^n \setminus A^\circ = (\mathbb{R}^n \setminus A)^- \text{ et } \mathbb{R}^n \setminus A^- = (\mathbb{R}^n \setminus A)^\circ,$$

qui permettent de déduire les propriétés de l'adhérence (resp. de l'intérieur) de A à partir de celles de l'intérieur (resp. l'adhérence) de son complémentaire.

$* \rightarrow$ Le recours aux opérations d'intérieur, d'adhérence et de frontière sur une même partie A de $\mathbb{R}^n$ n'introduit qu'un nombre fini d'ensembles distincts.

Proposition 2.4.2.7 *Tout ensemble formé à partir de A et des opérations d'intérieur, d'adhérence et de frontière est nécessairement égal à un des ensembles suivants:*

$$\begin{gathered}\emptyset, A\\ A^-, A^\circ, A^\bullet,\\ A^{-\circ}, A^{-\bullet}, A^{\circ-}, A^{\circ\bullet}, A^{\bullet\circ}, A^{\bullet\bullet},\\ A^{-\circ-}, A^{-\circ\bullet}, A^{\circ-\circ}, A^{\circ-\bullet}, A^{\bullet\circ-}, A^{\bullet\circ\bullet}.\end{gathered}$$

Preuve. Bien sûr, on rencontre A, A°, A^- et $A^\bullet$.

Considérons à présent les ensembles obtenus à partir de A et deux des opérations envisagées. En plus de ceux mentionnés dans l'énoncé, on peut aussi rencontrer $A^{--} = A^-$, $A^{\circ\circ} = A^\circ$ et $A^{\bullet-} = A^\bullet$.

Considérons ensuite les ensembles obtenus à partir de A et trois des opérations envisagées. En plus de ceux mentionnés dans l'énoncé, on peut rencontrer (tout en négligeant bien sûr ceux qui commencent ou se terminent par un des assemblages ${}^{--}$, ${}^{\circ\circ}$ ou ${}^{\bullet-}$ déjà éliminés)

$A^{-\bullet\circ} = \emptyset$ car la frontière d'un fermé est d'intérieur vide,

$A^{-\bullet\bullet} = A^{-\bullet}$ car la frontière d'un fermé est d'intérieur vide et car un fermé est égal à l'union de son intérieur et de sa frontière,

$A^{\circ\bullet\circ} = \emptyset$ car la frontière d'un ouvert est d'intérieur vide,

$A^{\circ\bullet\bullet} = A^{\circ\bullet}$ car la frontière d'un ouvert est d'intérieur vide et car un fermé est égal à l'union de son intérieur et de sa frontière,

$A^{\bullet\bullet\circ} = \emptyset$ car la frontière d'un fermé est d'intérieur vide,

$A^{\bullet\bullet\bullet} = A^{\bullet\bullet}$ car la frontière d'un fermé est d'intérieur vide et car un fermé est égal à l'union de son intérieur et de sa frontière.

Considérons ensuite les ensembles obtenus à partir de A et quatre des opérations envisagées. En plus des assemblages qui commencent ou se terminent par un des assemblages déjà éliminés, on trouve

$A^{-\circ-\circ} = A^{-\circ}$ car on a $A^{-\circ-} \supset A^{-\circ} \Rightarrow A^{-\circ-\circ} \supset A^{-\circ}$ et $A^{-\circ} \subset A^{-} \Rightarrow A^{-\circ-} \subset A^{-} \Rightarrow A^{-\circ-\circ} \subset A^{-\circ}$,

$A^{-\circ-\bullet} = A^{-\circ\bullet}$ car on a
$$\begin{aligned} A^{-\circ-\bullet} &= A^{-\circ--}(\mathbb{R}^n \setminus A^{-\circ-\circ}) \\ &= A^{-\circ-} \cap (\mathbb{R}^n \setminus A^{-\circ\circ}) = A^{-\circ\bullet}, \end{aligned}$$

$A^{\circ-\circ-} = A^{\circ-}$ car on a $A^{\circ-\circ} \subset A^{\circ-} \Rightarrow A^{\circ-\circ-} \subset A^{\circ-}$ et $A^{\circ-} \supset A^{\circ} \Rightarrow A^{\circ-\circ} \supset A^{\circ} \Rightarrow A^{\circ-\circ-} \supset A^{\circ-}$,

$A^{\circ-\circ\bullet} = A^{\circ-\bullet}$ car on a $A^{\circ-\circ\bullet} = A^{\circ-\circ-\bullet} = A^{\circ-\bullet}$, la première égalité provenant de $E^{-\circ\bullet} = E^{-\circ-\bullet}$ pour toute partie E de $\mathbb{R}^n$,

$A^{\bullet\circ-\circ} = A^{\bullet-\circ-\circ} = A^{\bullet-\circ} = A^{\bullet\circ}$,

$A^{\bullet\circ-\bullet} = A^{\bullet\circ--} \cap (\mathbb{R}^n \setminus A^{\bullet\circ-\circ}) = A^{\bullet\circ-} \cap (\mathbb{R}^n \setminus A^{\bullet-\circ-\circ})$, ce dernier ensemble étant égal à $A^{\bullet\circ-} \cap (\mathbb{R}^n \setminus A^{\bullet-\circ}) = A^{\bullet\circ-} \cap (\mathbb{R}^n \setminus A^{\bullet\circ\circ}) = A^{\bullet\circ\bullet}$.

Au total, aucune autre possibilité que celles annoncées dans l'énoncé ne peut apparaître.

Pour conclure, nous devons encore prouver qu'aucune des possibilités de l'énoncé n'est superflue. Pour s'en assurer, il suffit de procéder comme suit: on choisit a, b, c, $d \in \mathbb{R}$ tels que $a < b < c < d$ puis on considère l'ensemble
$$A = \,]-\infty, a[\, \cup \,]a, b[\, \cup \{c\} \cup \{\, x : x \in \mathbb{Q}\,, x \in \,]d, +\infty[\,\}\,.$$
Il est non vide et tel que

A^{-}	$=$	$]-\infty, b] \cup \{c\} \cup [d, +\infty[\,$,	A°	$=$	$]-\infty, a[\, \cup \,]a, b[\,$,
$A^{\bullet}$	$=$	$\{a, b, c\} \cup [d, +\infty[\,$,	$A^{-\circ}$	$=$	$]-\infty, b[\, \cup \,]d, +\infty[\,$,
$A^{-\bullet}$	$=$	$\{b, c, d\}$,	$A^{\circ-}$	$=$	$]-\infty, b]\,$,
$A^{\circ\bullet}$	$=$	$\{a, b\}$,	$A^{\bullet\circ}$	$=$	$]d, +\infty[\,$,
$A^{\bullet\bullet}$	$=$	$\{a, b, c, d\}$,	$A^{-\circ-}$	$=$	$]-\infty, b] \cup [d, +\infty[\,$,
$A^{-\circ\bullet}$	$=$	$\{b, d\}$,	$A^{\circ-\circ}$	$=$	$]-\infty, b[\,$,
$A^{\circ-\bullet}$	$=$	$\{b\}$,	$A^{\bullet\circ-}$	$=$	$[d, +\infty[\,$,
$A^{\bullet\circ\bullet}$	$=$	$\{d\}$.			

Nous voyons que, pour cet ensemble A particulier, toutes les possibilités de l'énoncé sont distinctes.

D'où la conclusion.∎← ∗

2.4.3 Compacts

Dans une première étude de l'espace $\mathbb{R}^n$, les parties compactes peuvent être introduites en recourant à la caractérisation de Heine-Borel.

Définition. Une partie K de $\mathbb{R}^n$ est *compacte* (on dit aussi que K est un *compact*) si elle est bornée et fermée.

Exemples. a) *Tout intervalle borné et fermé est compact.*
b) *Toute boule fermée de $\mathbb{R}^n$ est compacte.*
c) *Toute partie finie est compacte.*
d) *Si la suite x_m converge vers x_0 dans $\mathbb{R}^n$, alors $K = \{ x_m : m \in \mathbb{N}\}$ est un compact.* D'une part K est borné car toute suite convergente est bornée. D'autre part, K est fermé. De fait, soit y_m une suite de K qui converge et soit y_0 sa limite. S'il en existe une sous-suite $y_{k(m)}$ constante, bien sûr la limite y_0 appartient à K. Si y_m n'admet pas de sous-suite constante, on vérifie aisément qu'il en existe une sous-suite qui est en même temps une sous-suite de x_m et ainsi on a $y_0 = x_0 \in K$. D'où la conclusion.

Voici une propriété remarquable des compacts de $\mathbb{R}^n$.

Théorème 2.4.3.1 (Bolzano-Weierstrass) *De toute suite bornée de $\mathbb{R}^n$, on peut extraire une sous-suite de Cauchy.*

Dès lors, de toute suite d'un compact K de $\mathbb{R}^n$, on peut extraire une sous-suite convergente vers un point de K.

Preuve. Soit $(x_m)_{m\in\mathbb{N}_0}$ une suite bornée de $\mathbb{R}^n$. Procédons par récurrence. Le nombre de mailles du réseau décimal d'équidistance 10^{-1} qui contiennent au moins un élément de la suite est fini. Par suite, il existe au moins une maille I de ce réseau telle que $\{ m \in \mathbb{N}_0 : x_m \in I\}$ ne soit pas fini, soit I_1 une maille du réseau qui jouit de cette propriété et soit $k(1)$ un entier tel que $x_{k(1)}$ appartienne à I_1. Si $I_1, \ldots, I_{m-1}$ et $x_{k(1)}, \ldots, x_{k(m-1)}$ sont obtenus, on détermine I_m et $x_{k(m)}$ de la manière suivante. Le nombre de mailles du réseau décimal d'équidistance 10^{-m} incluses dans I_{m-1} est fini. Il existe donc au moins une de ces mailles, soit I, telle que l'ensemble $\{ m : x_m \in I\}$ ne soit pas fini. Alors I_m est une telle maille et $k(m)$ est un entier strictement supérieur à $k(m-1)$ et tel que $x_{k(m)} \in I_m$. Cela étant, la suite $x_{k(m)}$ est bien sûr une sous-suite de la suite x_m. Pour conclure, il suffit alors de noter que cette suite $x_{k(m)}$ est de Cauchy car on a

$$p \leq q \Rightarrow x_{k(p)}, x_{k(q)} \in I_p \Rightarrow \left|x_{k(p)} - x_{k(q)}\right| \leq \mathrm{diam}(I_p) = 10^{-p} n^{1/2}.$$

D'où la conclusion.∎

Voici quelques applications de ce théorème d'extraction.

Proposition 2.4.3.2 *Le diamètre de tout compact non vide de $\mathbb{R}^n$ est réalisé.*

En d'autres termes, pour tout compact non vide K de $\mathbb{R}^n$, il existe $x, y \in K$ tels que $|x - y| = \operatorname{diam}(K)$.

Preuve. Comme K est borné et non vide, il admet un diamètre donné par

$$\operatorname{diam}(K) = \sup \{ |x - y| : x, y \in K \} .$$

On sait alors qu'il existe une suite de $\{ |x - y| : x, y \in K\}$ qui converge vers $\operatorname{diam}(K)$, c'est-à-dire qu'il existe des suites x_m et y_m de K telles que la suite $|x_m - y_m|$ converge vers $\operatorname{diam}(K)$.

Par le théorème d'extraction, de la suite x_m, on peut extraire une sous-suite $x_{k(m)}$ qui converge vers un point x de K. Cela étant, $y_{k(m)}$ est une suite de K dont, par le théorème d'extraction également, on peut extraire une sous-suite $y_{l(k(m))}$ qui converge vers un point y de K. Remarquons alors que $x_{l(k(m))}$ est une sous-suite de $x_{k(m)}$, donc est une suite qui converge vers x. Au total, il vient

$$\left| x_{l(k(m))} - y_{l(k(m))} \right| \to |x - y|$$

avec $x, y \in K$ alors que $\left| x_{l(k(m))} - y_{l(k(m))} \right|$ est une sous-suite de $|x_m - y_m|$, donc est une suite qui converge vers $\operatorname{diam}(K)$.

D'où la conclusion.∎

Proposition 2.4.3.3 *Dans $\mathbb{R}^n$, la distance d'un compact non vide à un fermé non vide est réalisée.*

En d'autres termes, pour tout compact non vide K et tout fermé non vide F de $\mathbb{R}^n$, il existe $x \in K$ et $y \in F$ tels que $|x - y| = \mathrm{d}(K, F)$.

Preuve. Comme on a

$$\mathrm{d}(K, F) = \inf \{ |x - y| : x \in K, y \in F \} ,$$

il existe des suites x_m de K et y_m de F telles que $|x_m - y_m| \to \mathrm{d}(K, F)$. De la suite x_m du compact K, on peut extraire une sous-suite convergente dans K: soit $x_{k(m)} \to x$ une telle sous-suite. Cela étant, considérons la sous-suite $y_{k(m)}$ de la suite y_m. Comme, pour tout $m \in \mathbb{N}_0$, on a $|y_m| \leq |x_m| + |x_m - y_m|$, on voit que la suite y_m est bornée. Par conséquent, de la suite $y_{k(m)}$, on peut extraire une sous-suite convergente: soit $y_{l(k(m))} \to y$ une telle sous-suite. Comme $y_{l(k(m))}$ appartient au fermé F pour tout $m \in \mathbb{N}_0$, on obtient que y appartient à F. La conclusion est alors directe: on a bien sûr $x_{l(k(m))} \to x$ puis

$$x_{l(k(m))} - y_{l(k(m))} \to x - y \quad \text{donc} \quad \left| x_{l(k(m))} - y_{l(k(m))} \right| \to |x - y| ,$$

ce qui entraine $|x - y| = \mathrm{d}(K, F)$.∎

Corollaire 2.4.3.4 *Si le compact non vide K et le fermé non vide F de $\mathbb{R}^n$ sont tels que* $\mathrm{d}(K, F) = 0$*, alors* $K \cap F \neq \emptyset$*.*∎

Corollaire 2.4.3.5 *Si le compact non vide K et l'ouvert non vide Ω de $\mathbb{R}^n$ sont tels que $\Omega \neq \mathbb{R}^n$ et $K \subset \Omega$, alors* $\mathrm{d}(K, \mathbb{R}^n \setminus \Omega) > 0$*.*∎

Théorème 2.4.3.6 (Cantor) *Si $(K_m)_{m\in\mathbb{N}_0}$ est une suite de compacts non vides emboîtés en décroissant* (c'est-à-dire que $K_{m+1} \subset K_m$ pour tout $m \in \mathbb{N}_0$)*, alors $K = \bigcap_{m=1}^{\infty} K_m$ est un compact non vide.*

Preuve. Comme toute intersection de bornés est bornée et que toute intersection de fermés est fermée, nous savons déjà que K est compact.

Pour conclure, prouvons que K est non vide. Comme K_m diffère de $\emptyset$ pour tout $m \in \mathbb{N}_0$, il existe une suite x_m telle que $x_m \in K_m$ pour tout $m \in \mathbb{N}_0$. En particulier, la suite x_m est une suite de K_1 et on peut en extraire une sous-suite $x_{k(m)}$ qui converge vers un élément x_0 de K_1. En fait, on a $x_0 \in K_M$ pour tout $M \in \mathbb{N}_0$ car la suite $(y_m = x_{k(M+m)})_{m\in\mathbb{N}_0}$ est évidemment une suite de K_M et une sous-suite de $x_{k(m)}$. On a donc $x_0 \in K$.∎

2.5 Séries dans $\mathbb{R}^n$

2.5.1 Définition

Nous allons étudier à présent des suites de $\mathbb{R}^n$ qui se présentent sous une forme particulière.

Définitions. On appelle *série associée à la suite $(x_m)_{m\in\mathbb{N}_0}$ de $\mathbb{R}^n$*, la nouvelle suite $(X_k)_{k\in\mathbb{N}_0}$ dont le k-ème élément est égal à $X_k = \sum_{m=1}^{k} x_m$; c'est donc la suite

$$X_1 = x_1,\ X_2 = x_1 + x_2,\ X_3 = x_1 + x_2 + x_3,\ \ldots;$$

elle est notée $\sum_{m=1}^{\infty} x_m$.

Les x_m sont appelés les *termes* de la série; plus particulièrement, x_m est le *m-ème terme* de la série et on parle de la série de *terme général* x_m. De plus, X_k est appelé *k-ème somme partielle de la série.*

Ainsi, *la série $\sum_{m=1}^{\infty} x_m$ associée à la suite x_m apparaît comme étant la suite X_k des sommes partielles successives de la série.*

Remarques. a) La considération de la série $\sum_{m=1}^{\infty} x_m$ revient donc à celle de la suite X_k de ses sommes partielles.

b) Inversement, la considération de la suite x_m peut se ramener à celle d'une série $\sum_{m=1}^{\infty} y_m$. Il suffit pour cela de définir les y_m par $y_1 = x_1$ et $y_m = x_m - x_{m-1}$ pour tout entier $m \geq 2$ car on a évidemment $x_k = \sum_{m=1}^{k} y_m = Y_k$ pour tout $k \in \mathbb{N}_0$.

Afin de parler des séries de manière autonome (c'est-à-dire sans faire appel aux suites), nous sommes amenés à introduire les définitions suivantes.

Définitions. La série $\sum_{m=1}^{\infty} x_m$ est *convergente* si la suite X_k de ses sommes partielles converge. Dans ce cas, la limite de la suite X_k est appelée *somme de la série* et est notée $\sum_{m=1}^{\infty} x_m$.

La série est *divergente* si elle ne converge pas.

Remarque. Si la série associée à la suite x_m converge, la notation $\sum_{m=1}^{\infty} x_m$ représente donc tout aussi bien la série (c'est-à-dire la suite X_k de ses sommes partielles) que sa limite (c'est-à-dire la limite de la suite X_k). Il s'agit là d'un abus d'écriture consacré par l'usage et qui ne provoque jamais de confusion si le contexte est clair.

Voici la traduction du critère de Cauchy dans la langage des séries.

Critère 2.5.1.1 *La série $\sum_{m=1}^{\infty} x_m$ converge si et seulement si, pour tout $\varepsilon > 0$, il existe M tel que*

$$q \geq p \geq M \Rightarrow \left| \sum_{m=p}^{q} x_m \right| \leq \varepsilon.$$

Preuve. De fait, la série $\sum_{m=1}^{\infty} x_m$ converge si et seulement si la suite X_k de ses sommes partielles converge, c'est-à-dire si et seulement si cette suite X_k est de Cauchy. Or la suite X_k est de Cauchy si et seulement si, pour tout $\varepsilon > 0$, il existe M' tel que

$$p, q \geq M' \Rightarrow |X_p - X_q| \leq \varepsilon.$$

Pour conclure, il suffit alors de poser $M = M' + 1$ et de remarquer que, pour tous p, $q \in \mathbb{N}_0$ tels que $M \leq p \leq q$, on a

$$|X_q - X_{p-1}| = \left| \sum_{m=p}^{q} x_m \right| \leq \varepsilon. \blacksquare$$

La convergence de la série $\sum_{m=1}^{\infty} x_m$ impose des conditions sur ses termes x_m. En voici une qui découle directement du critère de Cauchy.

Proposition 2.5.1.2 *Si la série $\sum_{m=1}^{\infty} x_m$ converge, la suite x_m converge vers 0.*

Preuve. De fait, pour tout $\varepsilon > 0$, il existe M tel que

$$q \geq p \geq M \Rightarrow \left| \sum_{m=p}^{q} x_m \right| \leq \varepsilon,$$

donc en particulier tel que

$$m \geq M \Rightarrow |x_m| = \left| \sum_{j=m}^{m} x_j \right| \leq \varepsilon. \blacksquare$$

Remarque. La réciproque de cette proposition est fausse. En effet, la suite numérique réelle $x_m = 1/m$ converge vers 0 et nous avons déjà établi que la suite $\sum_{m=1}^{M} 1/m$ ne converge pas (cf. page 55), c'est-à-dire que *la série harmonique* $\sum_{m=1}^{\infty} 1/m$ *ne converge pas.*

Notation. Soient $(x_m)_{m\in\mathbb{N}_0}$ une suite de $\mathbb{R}^n$ et M un entier strictement positif. On peut alors introduire la suite y_m de m-ème élément défini par $y_m = x_{M+m-1}$ et considérer la série $\sum_{m=1}^{\infty} y_m$. En fait, cette dernière série est plutôt notée $\sum_{m=M}^{\infty} x_m$, sa signification étant claire et sa notation ne faisant pas appel aux y_m.

Théorème 2.5.1.3 *Si la série $\sum_{m=1}^{\infty} x_m$ converge, alors, pour tout entier $M \in \mathbb{N}_0$, la série $\sum_{m=M}^{\infty} x_m$ converge et*

$$\sum_{m=1}^{\infty} x_m = \sum_{m=1}^{M-1} x_m + \sum_{m=M}^{\infty} x_m.$$

Inversement, s'il existe $M \in \mathbb{N}_0$ tel que la série $\sum_{m=M}^{\infty} x_m$ converge, alors la série $\sum_{m=1}^{\infty} x_m$ converge.

Preuve. C'est une conséquence directe de la propriété relative aux combinaisons linéaires de suites convergentes, qui exprime que la suite $(X_k)_{k\geq M}$ définie par

$$X_k = \sum_{m=1}^{M-1} x_m + \sum_{m=M}^{k} x_m, \quad \forall k \geq M,$$

converge si et seulement si la suite $(Y_k)_{k\geq M} = \sum_{m=M}^{k} x_m$ converge, et qui donne également l'égalité annoncée entre les limites. $\blacksquare$

Définition. Si la série $\sum_{m=1}^{\infty} x_m$ converge, alors, pour tout $M \in \mathbb{N}_0$, la somme de la série $\sum_{m=M}^{\infty} x_m$ est appelé *M-ème reste* de la série $\sum_{m=1}^{\infty} x_m$.

Définition. Une série $\sum_{m=1}^{\infty} x_m$ est *numérique* si tous les x_m sont des nombres complexes; elle est *numérique réelle* si tous les x_m sont des nombres réels.

Proposition 2.5.1.4 *Une série numérique réelle à termes positifs converge si et seulement si la suite de ses sommes partielles est majorée.*

Preuve. De fait, la suite des sommes partielles est alors une suite numérique réelle croissante: elle converge si et seulement si elle est majorée. $\blacksquare$

2.5.2 Séries géométriques dans $\mathbb{C}$

Définition. Soit z un nombre complexe non nul. La *série géométrique de raison z* est la série $\sum_{m=0}^{\infty} z^m$.

Théorème 2.5.2.1 *La série géométrique de raison $z \in \mathbb{C} \setminus \{0\}$ converge si et seulement si on a $|z| < 1$, auquel cas sa somme vaut $1/(1-z)$.*

Preuve. Rappelons la formule relative aux progressions géométriques de raison $z \in \mathbb{C} \setminus \{0,1\}$:

$$\sum_{m=0}^{k} z^m = 1 + z + \cdots + z^k = \frac{1 - z^{k+1}}{1-z}.$$

Si on a $0 < |z| < 1$, cette formule montre immédiatement que la série géométrique de raison z converge vers

$$\lim_{k\to\infty} \left(\frac{1}{1-z} - \frac{z^{k+1}}{1-z} \right) = \frac{1}{1-z}$$

car on a alors

$$\left| \frac{z^{k+1}}{1-z} \right| = \frac{|z|^{k+1}}{|1-z|} \to 0.$$

De plus, pour tout $z \in \mathbb{C}$ tel que $|z| \geq 1$, la suite z^m ne tend pas vers 0 et la série géométrique de raison z ne peut converger.∎

2.5.3 Critères d'Abel

Les critères d'Abel relatifs à la convergence des séries sont basés sur l'inégalité suivante.

Lemme 2.5.3.1 (inégalité d'Abel) *Etant donné $J \in \mathbb{N}_0$; $x_1, \ldots, x_J \in \mathbb{R}^n$ et $r_1, \ldots, r_J \in \mathbb{R}$, on a*

$$\left| \sum_{j=1}^{J} r_j x_j \right| \leq \sum_{j=1}^{J-1} |r_j - r_{j+1}| \cdot \sup_{1\leq k \leq J-1} \left| \sum_{j=1}^{k} x_j \right| + |r_J| \cdot \left| \sum_{j=1}^{J} x_j \right|.$$

Preuve. Cela résulte aussitôt de l'identité

$$\begin{aligned}\sum_{j=1}^{J} r_j x_j &= (r_1 - r_2)x_1 + (r_2 - r_3)(x_1 + x_2) + \cdots \\ &\quad + (r_{J-1} - r_J)(x_1 + \cdots + x_{J-1}) + r_J(x_1 + \cdots + x_J).\end{aligned}$$ ∎

Critère 2.5.3.2 (Abel) *Si la suite r_m de $\mathbb{R}$ converge vers 0 et est telle que la série $\sum_{m=1}^{\infty} |r_m - r_{m+1}|$ converge* (c'est le cas de toute suite numérique positive décroissant vers 0) *et s'il existe $C > 0$ tel que la suite x_m de $\mathbb{R}^n$ vérifie la condition*

$$p \leq q \Rightarrow \left| \sum_{m=p}^{q} x_m \right| \leq C,$$

alors la série $\sum_{m=1}^{\infty} r_m x_m$ converge et, pour tout $M \in \mathbb{N}_0$, on a la majoration suivante du reste

$$\left| \sum_{m=M}^{\infty} r_m x_m \right| \leq C \sum_{m=M}^{\infty} |r_m - r_{m+1}| \, .$$

Preuve. Pour tout $\varepsilon > 0$, il existe $M \in \mathbb{N}_0$ tel que $|r_m| \leq \varepsilon/(2C)$ pour tout $m \geq M$ et $\sum_{m=M}^{\infty} |r_m - r_{m+1}| \leq \varepsilon/(2C)$ donc tel que

$$M \leq p \leq q \Rightarrow \left| \sum_{m=p}^{q} r_m x_m \right| \leq C \sum_{m=p}^{q-1} |r_m - r_{m+1}| + C \, |r_q| \leq \varepsilon.$$

Dès lors, la série converge car elle est de Cauchy. La majoration du reste s'obtient directement en passant à la limite pour $q \to \infty$ dans la majoration que nous venons d'obtenir.∎

Définition. Une *série alternée* est une série numérique réelle qui peut se mettre sous la forme $\sum_{m=1}^{\infty} (-1)^m r_m$, les r_m étant tous des nombres supérieurs ou égaux à 0.

Critère 2.5.3.3 (séries alternées) *Si r_m est une suite numérique réelle décroissante vers 0, la série alternée $\sum_{m=1}^{\infty} (-1)^m r_m$ converge et, pour tout $M \in \mathbb{N}_0$, on a la majoration suivante du reste*

$$\left| \sum_{m=M}^{\infty} (-1)^m r_m \right| \leq r_M.$$

Preuve. C'est une conséquence immédiate du premier critère d'Abel car, pour tous p, q tels que $p \leq q$, on a évidemment $\left| \sum_{m=p}^{q} (-1)^m \right| \leq 1$.∎

Critère 2.5.3.4 (Abel) *Si la suite r_m de nombres réels est telle que la série $\sum_{m=1}^{\infty} |r_m - r_{m+1}|$ converge* (c'est le cas de toute suite numérique positive et décroissante) *et si la série $\sum_{m=1}^{\infty} x_m$ converge dans $\mathbb{R}^n$, alors la série $\sum_{m=1}^{\infty} r_m x_m$ converge et, pour tout $M \in \mathbb{N}_0$, on a la majoration suivante du reste*

$$\left| \sum_{m=M}^{\infty} r_m x_m \right| \leq \left(|r_M| + 2 \sum_{m=M}^{\infty} |r_m - r_{m+1}| \right) \cdot \sup_{l \geq M} \left| \sum_{k=M}^{l} x_k \right| .$$

Preuve. Posons

$$C = 1 + |r_1| + \sum_{m=1}^{\infty} |r_m - r_{m+1}|$$

et remarquons que, vu l'égalité $r_m = r_1 + \sum_{j=1}^{m-1}(r_{j+1} - r_j)$ valable pour tout $m \in \mathbb{N}_0$, on a bien sûr $|r_m| \leq C$ pour tout $m \in \mathbb{N}_0$. Cela étant, la série $\sum_{m=1}^{\infty} r_m x_m$ est de Cauchy donc converge car, pour tout $\varepsilon > 0$, il existe M tel que

$$M \leq p \leq q \Rightarrow \left| \sum_{m=p}^{q} x_m \right| \leq \frac{\varepsilon}{2C}$$

donc tel que, pour tous p, $q \in \mathbb{N}_0$ tels que $M \leq p \leq q$, on a

$$\left| \sum_{m=p}^{q} r_m x_m \right| \leq \sum_{m=p}^{q-1} |r_m - r_{m+1}| \cdot \sup_{p \leq l \leq q-1} \left| \sum_{k=p}^{l} x_k \right| + |r_q| \cdot \left| \sum_{k=p}^{q} x_k \right| \leq \varepsilon.$$

L'inégalité portant sur le reste de la série s'obtient en passant à la limite pour $q \to \infty$ dans cette dernière inégalité.∎

Le cas $\mathbb{R}^n = \mathbb{C}$ se traîte de manière analogue et procure aussitôt les résultats suivants.

Lemme 2.5.3.5 (inégalité d'Abel) *Etant donné $J \in \mathbb{N}_0$ et $c_1, \ldots, c_J, z_1, \ldots, z_J \in \mathbb{C}$, on a l'inégalité*

$$\left| \sum_{j=1}^{J} c_j z_j \right| \leq \sum_{j=1}^{J-1} |c_j - c_{j+1}| \cdot \sup_{1 \leq k \leq J-1} \left| \sum_{j=1}^{k} z_j \right| + |c_J| \cdot \left| \sum_{j=1}^{J} z_j \right| . ∎$$

Critère 2.5.3.6 (Abel) *Si la suite c_m de $\mathbb{C}$ converge vers 0 et est telle que la série $\sum_{m=1}^{\infty} |c_m - c_{m+1}|$ converge et s'il existe $C > 0$ tel que la suite z_m de $\mathbb{C}$ vérifie la condition*

$$p \leq q \Rightarrow \left| \sum_{m=p}^{q} z_m \right| \leq C,$$

alors la série $\sum_{m=1}^{\infty} c_m z_m$ converge et, pour tout $M \in \mathbb{N}_0$, on a la majoration suivante du reste

$$\left| \sum_{m=M}^{\infty} c_m z_m \right| \leq C \sum_{m=M}^{\infty} |c_m - c_{m+1}| . ∎$$

Critère 2.5.3.7 (d'Abel) *Si la suite c_m de nombres complexes est telle que la série $\sum_{m=1}^{\infty} |c_m - c_{m+1}|$ converge et si la série $\sum_{m=1}^{\infty} z_m$ converge dans $\mathbb{C}$, alors la série $\sum_{m=1}^{\infty} c_m z_m$ converge et, pour tout $M \in \mathbb{N}_0$, on a la majoration suivante du reste*

$$\left| \sum_{m=M}^{\infty} c_m z_m \right| \leq \left(|c_M| + 2 \sum_{m=M}^{\infty} |c_m - c_{m+1}| \right) \cdot \sup_{l \geq M} \left| \sum_{k=M}^{l} z_k \right| . \blacksquare$$

2.5.4 Séries absolument convergentes et séries semi-convergentes

Introduisons une distinction essentielle parmi les séries convergentes.

Définition. Une série $\sum_{m=1}^{\infty} x_m$ de $\mathbb{R}^n$ est *absolument convergente* si la *série des modules* $\sum_{m=1}^{\infty} |x_m|$ converge.

Théorème 2.5.4.1 *Toute série absolument convergente $\sum_{m=1}^{\infty} x_m$ de $\mathbb{R}^n$ converge et donne lieu à l'inégalité*

$$\left| \sum_{m=1}^{\infty} x_m \right| \leq \sum_{m=1}^{\infty} |x_m| .$$

Preuve. Cela résulte aussitôt du critère de Cauchy car, pour tous p, $q \in \mathbb{N}_0$ tels que $p \leq q$, on a évidemment

$$\left| \sum_{m=p}^{q} x_m \right| \leq \sum_{m=p}^{q} |x_m| . \blacksquare$$

Remarque. Insistons sur le fait que la réciproque de ce résultat est fausse. Ainsi, à partir du critère de convergence des séries alternées, on établit directement que la série $\sum_{m=1}^{\infty} (-1)^m/m$ converge alors que nous savons que la série harmonique diverge.

Définition. Une série $\sum_{m=1}^{\infty} x_m$ de $\mathbb{R}^n$ est *semi-convergente* si elle converge sans être absolument convergente.

$* \rightarrow$ Les séries absolument convergentes jouissent de propriétés remarquables vis-à-vis des groupements et permutations de leurs termes.

Théorème 2.5.4.2 *Si la série $\sum_{m=1}^{\infty} x_m$ de $\mathbb{R}^n$ est absolument convergente, on peut arbitrairement grouper et permuter ses termes sans altérer sa convergence ni sa limite.*

Preuve. La situation la plus délicate est évidemment celle où on considère la série

$$\sum_{k=1}^{\infty}\sum_{m=1}^{\infty} x_{k(m)},$$

où les ensembles $A_k = \{\, k(m) : m \in \mathbb{N}_0 \}$ pour $k \in \mathbb{N}_0$ constituent une partition de $\mathbb{N}_0$ et où, pour tout $k \in \mathbb{N}_0$, l'application $k(\cdot)\colon \mathbb{N}_0 \to A_k$ est une numérotation de A_k. Dans les autres cas, la preuve qui suit se simplifie.

D'une part, pour tout $k \in \mathbb{N}_0$, la série $\sum_{m=1}^{\infty} x_{k(m)}$ converge absolument car la suite $\sum_{m=1}^{j} \left|x_{k(m)}\right|$ est croissante et majorée par $\sum_{m=1}^{\infty} |x_m|$.

D'autre part, étant donné $\varepsilon > 0$, il existe $M \in \mathbb{N}_0$ tel que $\sum_{m=M}^{\infty} |x_m| \leq \varepsilon/2$. Cela étant, puisque $\{x_1, \ldots, x_{M-1}\}$ est un ensemble fini, il existe d'abord $K \in \mathbb{N}_0$ tel que

$$\{1, \ldots, M-1\} \subset A_1 \cup \ldots \cup A_K$$

et ensuite $M' \in \mathbb{N}_0$ tel que

$$\{x_1, \ldots, x_{M-1}\} \subset \bigcup_{k=1}^{K} \left\{\, x_{k(m)} : m \leq M' \right\}.$$

Au total, pour tout $K' \geq K$, il vient

$$\begin{aligned}
&\left|\sum_{k=1}^{K'}\sum_{m=1}^{\infty} x_{k(m)} - \sum_{m=1}^{\infty} x_m\right| \\
&\quad\leq \left|\sum_{k=1}^{K'}\left(\sum_{m=1}^{\infty} x_{k(m)} - \sum_{m=1}^{M'} x_{k(m)}\right)\right| + \left|\sum_{k=1}^{K'}\sum_{m=1}^{M'} x_{k(m)} - \sum_{m=1}^{\infty} x_m\right| \leq \varepsilon,
\end{aligned}$$

ce qui suffit.∎

Remarque. Si la série $\sum_{m=1}^{\infty} x_m$ n'est pas absolument convergente, on ne peut pas impunément grouper ses termes.

Ainsi, la série $\sum_{m=1}^{\infty} (-1)^m$ ne converge pas car son terme général ne tend pas vers 0. Cependant la série $\sum_{m=1}^{\infty} \left((-1)^{2m-1} + (-1)^{2m}\right)$ converge trivialement vers 0.

Même dans une série semi-convergente, on ne peut pas permuter les termes sans rencontrer des problèmes. A ce sujet, le lemme suivant est aussi très instructif.

Lemme 2.5.4.3 *Si la série numérique réelle $\sum_{m=1}^{\infty} x_m$ est semi-convergente, les séries $\sum_{m=1}^{\infty} x_{m+}$ et $\sum_{m=1}^{\infty} x_{m-}$ divergent.*

Preuve. De fait, par exemple, si la série $\sum_{m=1}^{\infty} x_{m+}$ converge, la série

$$\sum_{m=1}^{\infty} x_{m-} = \sum_{m=1}^{\infty} (x_{m+} - x_m)$$

converge également comme différence de deux séries convergentes. Cela étant, la série $\sum_{m=1}^{\infty} x_m$ converge absolument car on a évidemment

$$\sum_{m=1}^{M} |x_m| \leq \sum_{m=1}^{\infty} x_{m+} + \sum_{m=1}^{\infty} x_{m-}, \quad \forall M \in \mathbb{N}_0. \blacksquare$$

Remarque. Si la série $\sum_{m=1}^{\infty} x_m$ n'est pas absolument convergente, on ne peut pas permuter impunément ses termes; la proposition suivante est particulièrement claire à ce sujet.

Proposition 2.5.4.4 *Si la série numérique réelle $\sum_{m=1}^{\infty} x_m$ est semi-convergente, alors*
a) *pour tout $r \in \mathbb{R}$, il existe une permutation y_m de la suite x_m telle que la série $\sum_{m=1}^{\infty} y_m$ converge vers r,*
b) *il existe une permutation y_m de la suite x_m telle que la série $\sum_{m=1}^{\infty} y_m$ converge vers $+\infty$* (resp. $-\infty$).

Preuve. Supposons, par exemple, r strictement positif. Nous allons construire les y_m par récurrence. Il existe un premier entier $M(1)$ tel que $\sum_{m=1}^{M(1)} x_{m+} > r$. On baptise alors y_k le k-ème $x_m \geq 0$ parmi les $x_1, \ldots x_{M(1)}$; soient $y_1, \ldots, y_{K(1)}$ les y_m ainsi déterminés. Ensuite, il existe un premier entier $M'(1)$ tel que

$$\sum_{m=1}^{M(1)} x_{m+} - \sum_{m=1}^{M'(1)} x_{m,-} < r.$$

Baptisons alors $y_{K(1)+k}$ le k-ème $x_m < 0$ parmi les $x_1, \ldots, x_{M'(1)}$; soient $y_{K(1)+1}, \ldots, y_{K'(1)}$ les $y_{K(1)+k}$ ainsi retenus. Cela étant, si les $M(1), \ldots, M(m-1)$, les $M'(1), \ldots, M'(m-1)$ et les $y_1, \ldots, y_{K'(m-1)}$ sont obtenus, il existe un premier entier $M(m)$ strictement supérieur à $M(m-1)$ tel que

$$\sum_{k=1}^{K'(m-1)} y_k + \sum_{m=M(m-1)+1}^{M(m)} x_{m+} > r$$

et on baptise $y_{K'(m-1)+k}$ le k-ème $x_m \geq 0$ parmi les $x_{M(m-1)+1}, \ldots, x_{M(m)}$, d'où les $y_{K'(m-1)+1}, \ldots, y_{K(m)}$. Il existe ensuite un premier entier $M'(m) > M'(m-1)$ tel que

$$\sum_{k=1}^{K(m)} y_k - \sum_{m=M'(m-1)+1}^{M'(m)} x_{m-} < r$$

et on baptise $y_{K(m)+k}$ le k-ème $x_m < 0$ parmi les $x_{M'(m-1)+1}, \ldots, x_{M'(m)}$. Comme la suite x_m tend vers 0 et que les suites $M(m)$ et $M'(m)$ croissent évidemment vers $+\infty$, on vérifie alors aisément que la série $\sum_{m=1}^{\infty} y_m$ converge vers r.

Les autres cas s'établissent de même (cf. également la preuve de la proposition suivante pour les cas $+\infty$ et $-\infty$).∎

Remarque. Ainsi G. Dirichlet a établi les résultats suivants:

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \quad \to \quad \ln(2)$$
$$\frac{1}{1} + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \ldots \quad \to \quad \frac{3}{2}\ln(2).$$

Proposition 2.5.4.5 *Si, pour toute permutation y_m des éléments de la suite x_m de $\mathbb{R}^n$, la série $\sum_{m=1}^{\infty} y_m$ converge, alors la série $\sum_{m=1}^{\infty} x_m$ est absolument convergente.*

Preuve. On se ramène directement au cas où x_m est une suite numérique réelle, en passant aux différentes composantes des x_m.

On sait que la série $\sum_{m=1}^{\infty} x_m$ converge. Si elle est semi-convergente, le lemme précédent assure que les séries $\sum_{m=1}^{\infty} x_{m+}$ et $\sum_{m=1}^{\infty} x_{m-}$ divergent. En utilisant la méthode de démonstration de la proposition précédente, on construit une permutation y_m de la suite x_m et une suite $K(k)$ strictement croissante de $\mathbb{N}_0$ de la manière suivante: $y_1, \ldots, y_{K(1)}$ sont les premiers $x_m \geq 0$ avec $K(1)$ déterminé par $\sum_{k=1}^{K(1)} x_{m+} > 1$; $y_{K(1)+1}$ est ensuite le premier élément strictement négatif de la suite x_m; puis par récurrence, $y_{K(k-1)+2}, \ldots, y_{K(k)}$ sont les $K(k) - K(k-1) - 1$ premiers éléments ≥ 0 de la suite x_m qui n'ont pas encore été retenus, $K(k)$ étant déterminé par $K(k) > K(k-1)$ et $\sum_{k=1}^{K(k)} y_k > k$; et $y_{K(k)+1}$ est le k-ème élément strictement négatif de la suite x_m. On vérifie alors aisément que la suite des sommes partielles de la série $\sum_{m=1}^{\infty} y_m$ converge vers $+\infty$ car la suite x_m converge vers 0.∎

Exercice. *Si la série numérique réelle $\sum_{m=1}^{\infty} r_m$ converge et si on a $r_m \geq r_{m+1} \geq 0$ pour tout $m \in \mathbb{N}_0$, alors la suite mr_m converge vers 0.*

Suggestion. Si ce n'est pas le cas, on établit aisément qu'il existe une sous-suite $r_{k(m)}$ de la suite r_m et $\varepsilon > 0$ tels que $k(m)r_{k(m)} \geq \varepsilon$ pour tout $m \in \mathbb{N}_0$. De plus, il est clair qu'on peut supposer avoir $k(m+1) \geq 2k(m)$ pour tout $m \in \mathbb{N}_0$. Dans ces conditions, pour tout entier $M \geq 2$, il vient

$$\sum_{m=1}^{k(M)} r_m \geq k(1) r_{k(1)} + \sum_{m=2}^{M} (k(m) - k(m-1)) r_{k(m)} \geq (M-1)\frac{\varepsilon}{2}.$$

D'où une contradiction.□

Exercice. *Si la série numérique réelle* $\sum_{m=1}^{\infty} r_m$ *converge et si on a* $r_m \downarrow 0$*, alors il existe une série numérique réelle* $\sum_{m=1}^{\infty} s_m$ *convergente telle que* $s_m \downarrow 0$ *et* $s_m/r_m \to +\infty$.

Suggestion. Posons $M_0 = 1$ et, pour tout $k \in \mathbb{N}$, choisissons un entier M_k tel que $2M_k \leq r_{M_{k-1}}$ et $\sum_{m=M_k}^{\infty} r_m \leq 2^{-2k}$. Ensuite, pour tout $k \in \mathbb{N}$ et tout $m \in \{M_{k-1}, \ldots, M_k - 1\}$, posons $s_m = \inf\{2^{k-1} r_{M_{k-1}}, 2^k r_m\}$. Il est clair que $s_m \downarrow 0$ et que $s_m/r_m \to +\infty$. Enfin la série $\sum_{m=1}^{\infty} s_m$ converge car

$$\sum_{m=M_{k-1}}^{M_k - 1} s_m \leq 2^k \sum_{m=M_{k-1}}^{M_k - 1} r_m \leq 2^{-k+2}$$

a lieu pour tout $k \in \mathbb{N}_0$.□

Exercice. *Si* $(r_m)_{m \in \mathbb{N}}$ *est une suite de nombres strictement positifs tels que* $r_0 = 1$ *et* $r_m^2 \leq r_{m-1} r_{m+1}$ *pour tout* $m \in \mathbb{N}_0$ *et si la série* $\sum_{m=1}^{\infty} r_{m-1}/(m r_m)$ *converge, 'etablir que la suite* $(\sqrt[m]{r_m})_{m \in \mathbb{N}_0}$ *croît vers* $+\infty$.

Suggestion. Nous avons déjà établi (cf. p. 44) que, dans ces conditions, la suite $\sqrt[m]{r_m}$ est croissante. Si elle ne croît pas vers $+\infty$, il existe $C > 0$ tel que $\sqrt[m]{r_m} \uparrow C$. Cela étant, posons $s_m = r_m C^{-m}$ pour tout $m \in \mathbb{N}$. Il est alors clair que $s_m \uparrow 1$ et que $s_m^2 \leq s_{m-1} s_{m+1}$ pour tout $m \in \mathbb{N}_0$. D'autre part, on a également

$$\sum_{m=1}^{\infty} \frac{s_{m-1}}{m s_m} = C \sum_{m=1}^{\infty} \frac{r_{m-1}}{m r_m} <= \infty$$

donc $s_{m-1}/s_m \to 0$, vu l'avant dernier exercice.□

On a aussi le résultat suivant, dont la preuve sort du cadre de ce cours.

Théorème 2.5.4.6 (Levy-Steinitz) *Si* x_m *est une suite de* $\mathbb{R}^n$*, alors l'ensemble des limites des séries convergentes du type* $\sum_{m=1}^{\infty} y_m$ *où* y_m *est une permutation de la suite* x_m*, ne peut contenir deux points sans contenir la droite qu'ils déterminent.*■

Exercice. *Si la série* $\sum_{m=1}^{\infty} x_m$ *converge absolument, établir que*

$$K = \left\{ \sum_{m=1}^{\infty} r_m x_m : r_m \in \{0,1\} \quad \textit{pour tout} \quad m \in \mathbb{N}_0 \right\}$$

est compact.

Suggestion. Il est clair que, pour toute suite r_m de $\{0,1\}$, la série $\sum_{m=1}^{\infty} r_m x_m$ converge absolument et que K est un borné de $\mathbb{R}^n$. Pour conclure, il suffit alors d'établir que, de toute suite $\Sigma_m = \sum_{k=1}^{\infty} r_{m,k} x_k$ de K, on peut extraire une sous-suite convergente vers un élément de K. Or, de la suite $r_{m,1}$, nous pouvons extraire une sous-suite $r_{1(m),1}$ dont tous les éléments sont égaux à un même élément r_1 de $\{0,1\}$. De même, de la suite $r_{1(m),2}$, nous pouvons extraire une sous-suite $r_{2(m),2}$ dont tous les éléments sont égaux à un même élément r_2 de $\{0,1\}$. Continuons de la sorte. Il est alors aisé d'établir que la suite $\left(\sum_{k=1}^{\infty} r_{m(m),k} x_k\right)_{m \in \mathbb{N}_0}$ converge vers $\sum_{k=1}^{\infty} r_k x_k$.∎
$\leftarrow *$

2.5.5 Critère théorique de convergence des séries

Critère 2.5.5.1 *Soient* $\sum_{m=1}^{\infty} r_m$, $\sum_{m=1}^{\infty} s_m$, $\sum_{m=1}^{\infty} x_m$ *et* $\sum_{m=1}^{\infty} y_m$ *des séries numériques réelles à termes* ≥ 0.

a) *Si la série* $\sum_{m=1}^{\infty} r_m$ *converge et s'il existe* $M \in \mathbb{N}_0$ *et* $C > 0$ *tels que* $x_m \leq C r_m$ *pour tout* $m \geq M$, *alors la série* $\sum_{m=1}^{\infty} x_m$ *converge.*

b) *Si la série* $\sum_{m=1}^{\infty} s_m$ *diverge et s'il existe* $M \in \mathbb{N}_0$ *et* $C > 0$ *tels que* $y_m \geq C s_m$ *pour tout* $m \geq M$, *alors la série* $\sum_{m=1}^{\infty} y_m$ *diverge.*

Preuve. a) résulte immédiatement du critère de Cauchy car, pour tous p, $q \in \mathbb{N}_0$ tels que $M \leq p \leq q$, il vient

$$\left| \sum_{m=p}^{q} x_m \right| = \sum_{m=p}^{q} x_m \leq C \sum_{m=p}^{q} r_m \leq C \left| \sum_{m=p}^{q} r_m \right|.$$

b) De fait, comme on a $s_m \leq y_m / C$ pour tout $m \in \mathbb{N}_0$, la convergence de la série $\sum_{m=1}^{\infty} y_m$ implique vu a) la convergence de la série $\sum_{m=1}^{\infty} s_m$.∎

Définition. Les séries $\sum_{m=1}^{\infty} r_m$ et $\sum_{m=1}^{\infty} s_m$ qui apparaissent dans le critère précédent sont appelées *séries de comparaison.*

Remarques. a) En spécifiant la série de comparaison, on obtient autant de critères de convergence des séries, appelés *critères pratiques.* Nous en donnons trois au paragraphe suivant, parmi les plus utilisés.

b) Si la suite x_m est décroissante, on trouve, dans le cours de calcul intégral, une condition nécessaire et suffisante de convergence de la série $\sum_{m=1}^{\infty} x_m$.

2.5.6 Critères pratiques de convergence des séries

a) Cas où $\sum_{m=1}^{\infty} \theta^m$ est la série de référence

Nous savons que la série géométrique $\sum_{m=1}^{\infty} \theta^m$ converge pour $\theta \in]0, 1[$ et diverge pour $\theta \in [1, +\infty[$.

Critère 2.5.6.1 (racine) *Soit $\sum_{m=1}^{\infty} x_m$ une série numérique réelle à termes ≥ 0.*

a) *S'il existe $M \in \mathbb{N}_0$ et $\theta \in]0, 1[$ tels que $\sqrt[m]{x_m} \leq \theta$ pour tout $m \geq M$, alors la série $\sum_{m=1}^{\infty} x_m$ converge.*

Cela arrive notamment si on a $\sqrt[m]{x_m} \to \theta' < 1$.

b) *S'il existe $M \in \mathbb{N}_0$ tel que $\sqrt[m]{x_m} \geq 1$ pour tout $m \geq M$, alors la série $\sum_{m=1}^{\infty} x_m$ diverge.*

Cela arrive notamment si $\sqrt[m]{x_m} \to \Theta$, avec $\Theta > 1$ ou $\Theta = 1^+$.

Preuve. C'est une application immédiate du critère théorique car

$$\sqrt[m]{x_m} \left\{ \begin{array}{c} \leq \\ \geq \end{array} \right\} \theta \Leftrightarrow x_m \left\{ \begin{array}{c} \leq \\ \geq \end{array} \right\} \theta^m.$$

De plus, si la suite $\sqrt[m]{x_m}$ converge vers $\theta' \in [0, 1[$, il existe $M \in \mathbb{N}_0$ tel que

$$\sqrt[m]{x_m} \leq \tfrac{1}{2}(1 + \theta') < 1, \quad \forall m \geq M.$$

Enfin, si on a $\sqrt[m]{x_m} \to \Theta$ avec $\Theta > 1$ ou $\Theta = 1^+$, la suite x_m ne converge pas vers 0 et, par conséquent, la série $\sum_{m=1}^{\infty} x_m$ ne converge pas.∎.

Critère 2.5.6.2 (quotient) *Soit $\sum_{m=1}^{\infty} x_m$ une série numérique à termes > 0.*

a) *S'il existe $M \in \mathbb{N}_0$ et $\theta \in]0, 1[$ tels que $x_{m+1}/x_m \leq \theta$ pour tout $m \geq M$, alors la série $\sum_{m=1}^{\infty} x_m$ converge.*

Cela arrive notamment si on a $x_{m+1}/x_m \to \theta < 1$.

b) *S'il existe $M \in \mathbb{N}_0$ tel que $x_{m+1}/x_m \geq 1$ pour tout $m \geq M$, alors la série $\sum_{m=1}^{\infty} x_m$ diverge.*

Cela arrive notamment si $x_{m+1}/x_m \to \Theta$ avec $\Theta > 1$ ou $\Theta = 1^+$.

Preuve. C'est une application immédiate du critère théorique car on a bien sûr $\theta = \theta^{m+1}/\theta^m$ et $1 = 1^{m+1}/1^m$ pour tout $m \in \mathbb{N}_0$.

De plus, si la suite x_{m+1}/x_m converge vers $\theta' \in [0, 1[$, il existe $M \in \mathbb{N}_0$ tel que

$$\frac{x_{m+1}}{x_m} \leq \tfrac{1}{2}(1 + \theta') < 1, \quad \forall m \geq M.$$

Enfin, si on a $x_{m+1}/x_m \to \Theta$ avec $\Theta > 1$ ou $\Theta = 1^+$, la suite x_m ne converge pas vers 0 et, par conséquent, la série $\sum_{m=1}^{\infty} x_m$ ne converge pas.∎

Remarque. Insistons sur le fait que si la suite $\sqrt[m]{x_m}$ converge vers 1 sans que cette limite ne soit 1^+, le critère de la racine ne permet pas de conclure. De même, si la suite x_{m+1}/x_m converge vers 1 sans que cette limite ne soit 1^+, le critère du quotient ne permet pas de conclure. De la sorte, ces critères de la racine et du quotient ne permettent pas d'établir que la série $\sum_{m=1}^{\infty} 1/m$ (resp. $\sum_{m=1}^{\infty} 1/m^2$) diverge (resp. converge).

Nous allons remédier partiellement à cet état de fait au moyen du critère de Riemann.

b) Cas où $\sum_{m=1}^{\infty} m^{-\alpha}$ est la série de référence

Définition. La *série de Riemann* d'*ordre* $\alpha \geq 0$ s'écrit $\sum_{m=1}^{\infty} m^{-\alpha}$.

Remarque. Les critères de la racine et du quotient ne permettent pas de conclure si la série de Riemann d'ordre $\alpha \geq 0$ converge ou diverge. Il faut procéder autrement.

Lemme 2.5.6.3 (principe de condensation de Cauchy) *Si la suite $r_m \in [0, +\infty[$ est décroissante, la série $\sum_{m=1}^{\infty} r_m$ converge si et seulement si la série $\sum_{m=1}^{\infty} 2^m r_{2^m}$ converge.*

Preuve. La condition est nécessaire. La suite des sommes partielles de cette série $\sum_{m=1}^{\infty} 2^m r_{2^m}$ est croissante. De plus, elle est majorée car, pour tout $M \in \mathbb{N}_0$, on a

$$\begin{aligned}\sum_{m=1}^{M} 2^m r_{2^m} &= 2\sum_{m=1}^{M} 2^{m-1} r_{2^m} \\ &\leq 2\left(r_2 + (r_3 + r_4) + \cdots + (r_{2^{M-1}+1} + \cdots + r_{2^M})\right) \leq 2\sum_{m=1}^{\infty} r_m.\end{aligned}$$

La condition est suffisante. Bien sûr, la suite des sommes partielles de cette série $\sum_{m=1}^{\infty} r_m$ est croissante. De plus, elle est majorée car, pour tout entier $M \geq 2$, il vient

$$\begin{aligned}\sum_{m=1}^{M} r_m &\leq r_1 + r_2 + (r_3 + r_4) + \cdots + (r_{2^{M-1}+1} + \cdots + r_{2^M}) \\ &\leq r_1 + r_2 + 2r_{2^1} + \cdots + 2^{M-1} r_{2^{M-1}} \leq r_1 + r_2 + \sum_{m=1}^{\infty} 2^m r_{2^m}. \blacksquare\end{aligned}$$

Théorème 2.5.6.4 *La série de Riemann d'ordre $\alpha \geq 0$ converge pour $\alpha \in]1, +\infty[$ et diverge pour $\alpha \in [0, 1]$.*

Preuve. Pour $\alpha \geq 0$, la suite numérique $(m^{-\alpha})_{m\in\mathbb{N}_0}$ est décroissante. D'où la conclusion par le principe de condensation de Cauchy car

$$\sum_{m=1}^{\infty} 2^m (2^m)^{-\alpha} = \sum_{m=1}^{\infty} \left(2^{1-\alpha}\right)^m$$

est une série géométrique qui converge si et seulement si on a $2^{1-\alpha} < 1$.■

Critère 2.5.6.5 (Riemann) *Soit $\sum_{m=1}^{\infty} x_m$ une série numérique réelle à termes ≥ 0.*

a) *S'il existe $M \in \mathbb{N}_0$, $C > 0$ et $\alpha > 1$ tels que $m^\alpha x_m \leq C$ pour tout $m \geq M$, alors la série $\sum_{m=1}^{\infty} x_m$ converge.*

Cela arrive notamment si on a $m^\alpha x_m \to C$ avec $\alpha > 1$ et $C \in [0, +\infty[$.

b) *S'il existe $M \in \mathbb{N}_0$ et $C > 0$ tels que $mx_m \geq C$ pour tout $m \geq M$, alors la série $\sum_{m=1}^{\infty} x_m$ diverge.*

Cela arrive notamment si $mx_m \to C$ avec $C \in]0, +\infty[$ ou $C = +\infty$.

Preuve. C'est une conséquence directe du critère théorique et du théorème précédent.■

2.5.7 Etude de la convergence d'une série numérique réelle

Décrivons les différentes étapes qui interviennent dans l'étude de la convergence de la série numérique réelle $\sum_{m=1}^{\infty} x_m$.

a) Vérifier le sens des x_m.

b) Etude de la convergence absolue.

On explicite la valeur de $|x_m|$.

On applique ensuite

a) soit le *critère de la racine* sous la forme pratique suivante:

la série $\sum_{m=1}^{\infty} x_m$ converge absolument si $$\sqrt[m]{\lvert x_m\rvert} \to \theta < 1.$$	*la série $\sum_{m=1}^{\infty} x_m$ diverge si* $$\sqrt[m]{\lvert x_m\rvert} \to \begin{cases} \Theta > 1 \\ 1^+ \end{cases}$$ *car alors la suite $x_m \not\to 0$.*

b) soit le *critère du quotient* sous la forme pratique suivante:

la série $\sum_{m=1}^{\infty} x_m$ *converge absolument si*

$$\frac{|x_{m+1}|}{|x_m|} \to \theta < 1.$$

la série $\sum_{m=1}^{\infty} x_m$ *diverge si*

$$\frac{|x_{m+1}|}{|x_m|} \to \left\{ \begin{array}{l} \Theta > 1 \\ 1^+ \end{array} \right.$$

car alors la suite $x_m \not\to 0$.

Le choix entre a) et b) est dicté par la forme des $|x_m|$.

Dans les cas où a) et b) ne permettent pas de conclure, on recourt au *critère de Riemann* sous la forme pratique suivante:

la série $\sum_{m=1}^{\infty} x_m$ *converge absolument s'il existe* $\alpha > 1$ *tel que*

$$m^\alpha |x_m| \to C \in [0, +\infty[\,.$$

la série $\sum_{m=1}^{\infty} |x_m|$ *diverge si*

$$m\,|x_m| \to \left\{ \begin{array}{l} C \in]0, +\infty[\\ +\infty \end{array} \right.$$

mais la série $\sum_{m=1}^{\infty} x_m$ *peut être semi-convergente.*

c) Etude de la semi-convergence

Si la série n'est pas absolument convergente et si les critères précédents ne permettent pas d'affirmer qu'elle diverge, on applique le *critère des séries alternées.*

Remarque. La convergence des séries numériques (complexes) sera abordé au paragraphe 6.6.6.

2.5.8 Exemples d'étude de séries

Remarque. Dans les exercices relatifs à l'étude de la convergence des séries numériques réelles, on est amené à utiliser la convergence de certaines suites. Rappelons les résultats suivants

$$\sqrt[m]{r} \to \left\{ \begin{array}{ll} 1^- & \text{si} \quad r \in]0,1[\\ 1^+ & \text{si} \quad r \in]1, +\infty[\end{array} \right\} \qquad \sqrt[m]{m} \to 1^+$$

$$\left(1 + \tfrac{1}{m}\right)^m \uparrow \mathrm{e} \qquad \left(1 + \tfrac{1}{m}\right)^{m+1} \downarrow \mathrm{e}$$

établis dans le cahier d'exercices.

Exercice. *La série* $\sum_{m=1}^{\infty} m^{-m}$ *converge.*

Suggestion. C'est une série numérique réelle à termes strictement positifs. Son étude est immédiate car on a $\sqrt[m]{|x_m|} = 1/m \to 0$ et dès lors, elle converge en vertu du critère de la racine. Par contre, l'application du critère du quotient est plus délicate.

Exercice. *La série $\sum_{m=1}^{\infty} 1/(m!)$ converge.*

Suggestion. Il s'agit d'une série numérique réelle à termes strictement positifs. Comme $x_{m+1}/x_m = 1/(m+1) \to 0$, le critère du quotient permet d'affirmer qu'elle converge. Par contre, le critère de la racine s'applique moins aisément.

Exercice. *Pour tous a, $b > 0$ la série $\sum_{m=1}^{\infty} 1/(am+b)$ diverge.*

Suggestion. C'est une série numérique à termes strictement positifs. Son étude est immédiate à partir du critère de Riemann car on a $m/(am+b) \to 1/a$ et ainsi la série ne converge pas. Par contre, ni le critère de la racine, ni le critère du quotient ne permettent de conclure.

Exercice. *La série $\sum_{m=1}^{\infty} 1/m^2$ converge.*

Suggestion. De fait, c'est la série de Riemann d'ordre $\alpha = 2$.

Exercice. *Etudier la convergence de la série numérique dont le terme général est donné par*

$$x_m = \frac{(a-2)^m}{2^m a^m \sqrt{2m+1}}, \quad \forall m \in \mathbb{N}_0,$$

pour toutes les valeurs possibles du paramètre réel a.

Suggestion. Pour assurer un sens au terme général, nous devons éliminer la valeur $a = 0$ du paramètre.

On a

$$|x_m| = \frac{|a-2|^m}{2^m |a|^m \sqrt{2m+1}}, \quad \forall m \in \mathbb{N}_0.$$

Vu la forme de $|x_m|$, nous allons recourir au critère du quotient après avoir écarté la valeur $a = 2$ qui donne évidemment lieu à une série absolument convergente. Comme on a alors

$$\frac{|x_{m+1}|}{|x_m|} = \frac{|a-2|}{2\,|a|} \cdot \frac{\sqrt{2m+1}}{\sqrt{2m+3}} \to \left(\frac{|a-2|}{2\,|a|}\right)^-$$

et

$$\frac{|a-2|}{2\,|a|} \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} 1 \Leftrightarrow (a-2)^2 \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} 4a^2 \Leftrightarrow 3a^2 + 4a - 4 \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} 0,$$

nous avons les résultats suivants:
$a < -2$: la série est absolument convergente.

$-2 < a < 2/3$ et $a \neq 0$: la série diverge car son terme général ne tend pas vers 0.
$a > 2/3$ et $a \neq 2$: la série est absolument convergente.

Il reste alors à envisager les cas $a = -2$ et $a = 2/3$.

D'une part, pour $a = -2$, la série s'écrit $\sum_{m=1}^{\infty}(2m+1)^{-1/2}$ et ne converge pas, vu le critère de Riemann.

D'autre part, pour $a = 2/3$, la série s'écrit $\sum_{m=1}^{\infty}(-1)^m(2m+1)^{-1/2}$; il s'agit d'une série alternée et comme on a $(2m+1)^{-1/2} \downarrow 0$, elle est convergente.

Chapitre 3

Fonctions

3.1 Généralités sur les fonctions

3.1.1 Définitions

Définitions. Une *fonction* définie sur $A \subset \mathbb{R}^n$ est une application de A dans $\mathbb{C}$; c'est donc une loi f qui, à tout $x \in A$, associe un nombre complexe $f(x)$. On note explicitement

$$f\colon A \to \mathbb{C} \quad x \mapsto f(x)$$

mais on trouve aussi des notations telles que

$$f\colon A \to \mathbb{C} \;\text{ ou }\; f\colon x \mapsto f(x)$$

si le contexte est clair. On appelle A l'*ensemble de définition de f*, x la *variable*, $f(x)$ la *valeur de f en x* et $\{\, f(x) : x \in A\}$ l'*ensemble de variation de f sur A* ou l'*ensemble des valeurs de f sur A*.

Remarque. En théorie, on accorde une signification différente aux notations

a) f (= fonction = loi qui ...),

b) $f(x)$ (= nombre complexe représentant la valeur de f en x).

En pratique cependant, on dit bien souvent “la fonction $f(x)$”; il s'agit là d'un abus consacré par l'usage.

Définitions. Une fonction $f : A \to \mathbb{C}$ est *réelle* (resp. *positive*; *négative*) si $f(x)$ est un nombre réel (resp. ≥ 0; ≤ 0) pour tout $x \in A$.

Si on désire insister sur le fait qu'une fonction f n'est pas réelle, on dit que f est une fonction *à valeurs complexes*.

Exemples. Dans le chapitre précédent, nous avons déjà introduit quelques fonctions remarquables, à savoir

$$\begin{array}{lllllll} [\cdot]_j & : & \mathbb{R}^n & \to & \mathbb{R} & x & \mapsto & [x]_j, \\ |\cdot| & : & \mathbb{R}^n & \to & \mathbb{R} & x & \mapsto & |x|, \\ \cdot_+ & : & \mathbb{R} & \to & [0,+\infty[& x & \mapsto & x_+, \\ \cdot_- & : & \mathbb{R} & \to & [0,+\infty[& x & \mapsto & x_-. \end{array}$$

Voici quelques exemples supplémentaires:

a) $0\colon \mathbb{R}^n \to \mathbb{R} \quad x \mapsto 0$. On utilise le même symbole 0 pour désigner le nombre complexe zéro et cette fonction; cela ne crée pas de confusion, la signification à donner à 0 étant rendue claire par le contexte.

b) $\chi_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R} \quad x \mapsto 1$.

c) Etant donné un nombre réel et non nul r, la loi qui, à tout point (x_0, y_0) de l'ensemble $A = \mathbb{R}^2 \setminus \{\, (x,y) : x^2 + y^2 = r^2\}$ associe le nombre $1/(r^2 - x_0^2 - y_0^2)$ est une fonction définie sur A. C'est en fait la fonction

$$\frac{1}{r^2 - ([\cdot]_1)^2 - ([\cdot]_2)^2}$$

mais on préfère bien sûr recourir à la remarque précédente et parler tout simplement de la fonction

$$\frac{1}{r^2 - x^2 - y^2}.$$

Définition. Si f est une fonction définie sur $A \subset \mathbb{R}^n$, le *graphe de f* est l'ensemble

$$\mathcal{G}(f) = \{\, (x, \Re f(x), \Im f(x)) : x \in A\}\,;$$

c'est donc une partie de $\mathbb{R}^{n+2}$. Si f est une fonction réelle définie sur $A \subset \mathbb{R}^n$ on appelle plutôt *graphe de f* la partie suivante de $\mathbb{R}^{n+1}$

$$\mathcal{G}(f) = \{\, (x, f(x)) : x \in A\}\,.$$

Remarque. Il s'agit de ne pas confondre graphe de f avec *graphique de f*. La donnée du graphe de f est équivalente à la donnée de la fonction f elle-même. La représentation géométrique, appelée graphique d'une fonction réelle f définie sur une partie A de $\mathbb{R}$ (parfois de $\mathbb{R}^2$), peut être très instructive dans l'étude de f mais présente deux écueils:

a) il existe des fonctions définies sur $\mathbb{R}$ pour lesquelles il est impossible de tracer un graphique. Il en est ainsi pour la fonction f définie sur $\mathbb{R}$ par

$$f(x) = \begin{cases} 0 & \text{si } x \text{ est rationnel,} \\ 1 & \text{si } x \text{ est irrationnel,} \end{cases}$$

b) il est impossible d'obtenir une représentation graphique pour une fonction réelle définie sur $A \subset \mathbb{R}^3$ ou pour une fonction (à valeurs complexes) définie sur $A \subset \mathbb{R}^2$.

C'est pourquoi on ne peut recourir à cette notion de graphique, très intuitive et très utile dans de nombreux exemples pratiques, dans une étude générale des fonctions.

3.1.2 Opérations entre fonctions

a) Relations de comparaison. Deux fonctions f et g sont *égales* si elles sont définies sur une même partie A de $\mathbb{R}^n$ et si on a $f(x) = g(x)$ pour tout $x \in A$; on écrit $f = g$.

Si ce n'est pas le cas, on dit que f et g sont *différents*, ce qu'on note $f \neq g$. Cela signifie donc que leurs ensembles de définition ne sont pas égaux ou qu'ils sont égaux à une même partie A de $\mathbb{R}^n$ mais qu'il existe $x \in A$ tel que $f(x) \neq g(x)$.

Si les fonctions f et g sont définies et réelles sur $A \subset \mathbb{R}^n$, on peut les comparer davantage:

i) f *majore* ou *est supérieur ou égal à* g, ce qu'on note

$$f \geq g \quad \text{ou} \quad g \leq f,$$

si on a $f(x) \geq g(x)$ pour tout $x \in A$;

ii) f *minore* ou *est inférieur ou égal à* g si on a $f \leq g$.

En particulier, une fonction f est positive (resp. négative) sur $A \subset \mathbb{R}^n$ si on a $f \geq 0$ (resp. $f \leq 0$) sur A.

b) Prolongement, restriction. Soient f et f' des fonctions définies respectivement sur des parties A et A' de $\mathbb{R}^n$. Alors f est *la restriction de* f' *sur* A si on a $A \subset A'$ et $f(x) = f'(x)$ pour tout $x \in A$; on écrit $f = f'|_A$ et on dit également que f' est *un prolongement de* f *sur* A'.

c) Opérations algébriques. Soient f, g, f_1, $\ldots$, f_J des fonctions définies sur $A \subset \mathbb{R}^n$, où J est un élément de $\mathbb{N}_0$, et soient c_1, $\ldots$, c_J des nombres complexes.

On définit alors les opérations algébriques élémentaires suivantes:

i) la *combinaison linéaire correspondant aux* f_j *et aux* c_j notée $\sum_{j=1}^{J} c_j f_j$ est la fonction définie en tout point x de A par $(\sum_{j=1}^{J} c_j f_j)(x) = \sum_{j=1}^{J} c_j f_j(x)$,

ii) le *produit des* f_j noté $\prod_{j=1}^{J} f_j$ est la fonction définie en tout point x de A par $(\prod_{j=1}^{J} f_j)(x) = \prod_{j=1}^{J} f_j(x)$,

iii) si $g(x)$ diffère de 0 pour tout $x \in A$, le *quotient de* f *par* g, noté f/g est la fonction définie en tout point x de A par $(f/g)(x) = f(x)/g(x)$.

Une *opération algébrique* est une opération entre fonctions qu'on peut écrire explicitement au moyen d'un nombre fini d'opérations algébriques élémentaires entre ces fonctions.

C'est ainsi que

$$\frac{f}{g} + \frac{\sum_{j=1}^{J} c_j f_j}{f - \prod_{j=1}^{J} f_j}$$

est une opération algébrique portant sur ces fonctions si on a $g(x) \neq 0$ et $f(x) - \prod_{j=1}^{J} f_j(x) \neq 0$ pour tout $x \in A$.

Comme cas particulier du produit fini, introduisons le concept de fonction séparée. Une fonction f définie sur l'intervalle I de $\mathbb{R}^n$ est *séparée* s'il existe des fonctions $f_1, \ldots, f_p$ définies sur des parties adéquates telles que

$$f(x) = f_1(x_{\nu(1)}, \ldots, x_{\nu(k(1))}) \ldots f_p(x_{\nu(k(p-1)+1)}, \ldots, x_{\nu(k(p))}), \quad \forall x \in I,$$

avec $k(p) = n$ et $\nu(j) \neq \nu(k)$ si $j \neq k$. Ainsi une fonction qui ne dépend que d'une partie des composantes de la variable est une fonction séparée car elle peut être considérée comme étant le produit d'elle-même avec la fonction qui, aux autres composantes de la variable, associe la valeur 1. En particulier, la fonction f est *à variables séparées* s'il existe des fonctions $f_1, \ldots, f_n$ définies sur des parties adéquates de $\mathbb{R}$ telles que

$$f(x) = f_1(x_1) \ldots f_n(x_n), \quad \forall x \in I.$$

d) Fonction composée, fonction de fonction. Soit J un élément de $\mathbb{N}_0$ et soient $f_1, \ldots, f_J$ des fonctions réelles définies sur une même partie A de $\mathbb{R}^n$. Pour tout $x \in A$, $(f_1(x), \ldots, f_J(x))$ est alors un point de $\mathbb{R}^J$. Cela étant, si f est une fonction définie sur une partie E de $\mathbb{R}^J$ contenant l'*ensemble de variation des f_1, $\ldots$, f_J*, c'est-à-dire si on a

$$E \supset \{ (f_1(x), \ldots, f_J(x)) : x \in A \},$$

le nombre $f(f_1(x), \ldots, f_J(x))$ est défini pour tout $x \in A$. Sous ces conditions, la fonction

$$f(f_1, \ldots, f_J) \colon A \to \mathbb{C} \quad x \mapsto f(f_1(x), \ldots, f_J(x))$$

est appelée *fonction composée de f et des f_1, $\ldots$, f_J*. C'est donc la composition $f \circ (f_1, \ldots, f_J)$ des applications $f \colon E \to \mathbb{C}$ et $(f_1, \ldots, f_J) \colon A \to E$.

Comme exemple, notons que la fonction

$$g \colon \mathbb{R}^2 \to \mathbb{R} \quad x \mapsto x_1 + (x_2)^2$$

peut être interprétée comme étant une fonction composée. Il suffit pour cela de remarquer que les fonctions $f_1 = [\cdot]_1$ et $f_2 = [\cdot]_2^2$ sont définies sur $\mathbb{R}^2$, d'introduire la fonction $f(\cdot) = [\cdot]_1 + [\cdot]_2$ sur $\mathbb{R}^2$ puis de vérifier que $f(f_1, f_2) = g$. Cet exemple trivial montre bien la raison pour laquelle on introduit cette notion: elle permet de décrire une fonction g au moyen d'opérations plus élémentaires.

Une *fonction de fonction* est une fonction composée pour laquelle $J = 1$.

3.1.3 Fonctions associées

a) Fonctions associées à une fonction. Une fonction a été définie comme étant une application à valeurs complexes; en général, elle n'est pas réelle et encore moins positive. Cependant on peut toujours l'exprimer sous la forme d'une combinaison linéaire remarquable de fonctions positives particulières.

Etant donné une fonction $f\colon A \to \mathbb{C}$ et une application $\varphi\colon \mathbb{C} \to \mathbb{C}$, la composition $\varphi \circ f\colon A \to \mathbb{C}$ est une fonction définie sur A qui est bien souvent notée $\varphi(f)$.

Cela étant, à une fonction $f\colon A \to \mathbb{C}$, on associe

i) la *partie réelle de* f, à savoir $\Re \circ f$, qu'on note plutôt $\Re f$,

ii) la *partie imaginaire de* f, à savoir $\Im \circ f$, qu'on note plutôt $\Im f$,

iii) le *module de* f, à savoir $|\cdot| \circ f$, qu'on note plutôt $|f|$,

iv) la *fonction conjuguée de* f, à savoir $\bar{\cdot} \circ f$, qu'on note plutôt $\overline{f}$.

Il est clair que $\Re f$ et $\Im f$ sont des fonctions réelles, que $|f|$ est une fonction positive et que $\overline{f}$ est une fonction à valeurs complexes sur A.

De la sorte, nous avons déjà l'égalité

$$f = \Re f + i\Im f$$

qui décrit f comme combinaison linéaire spéciale de deux fonctions réelles particulières définies sur A.

A une fonction réelle $f\colon A \to \mathbb{R}$, on associe

i) la *partie positive de* f, à savoir la fonction $(\cdot)_+ \circ f$, notée plutôt f_+,

ii) la *partie négative de* f, à savoir la fonction $(\cdot)_- \circ f$, notée plutôt f_-.

Il s'agit bien sûr de deux fonctions positives sur A telles que

$$f = f_+ - f_-, \quad f_+ = (-f)_-, \quad f_- = (-f)_+ \ \text{ et } \ |f| = f_+ + f_-.$$

Remarquons que la première de ces égalités procure f au moyen d'une combinaison linéaire spéciale de fonctions positives particulières.

Au total, si f est une fonction définie sur $A \subset \mathbb{R}^n$, on a l'identité

$$f = (\Re f)_+ - (\Re f)_- + i(\Im f)_+ - i(\Im f)_-$$

qui répond bien sûr à l'objectif visé.

b) Fonctions associées à un nombre fini de fonctions. Si J appartient à $\mathbb{N}_0$ et si $f_1, \ldots, f_J$ sont des fonctions réelles définies sur une même partie A de $\mathbb{R}^n$, on leur associe leur *enveloppe supérieure*

$$\sup\{f_1, \ldots, f_J\}\colon A \to \mathbb{R} \quad x \mapsto \sup\{f_1(x), \ldots, f_J(x)\}$$

et leur *enveloppe inférieure*

$$\inf\{f_1, \ldots, f_J\} \colon A \to \mathbb{R} \quad x \mapsto \inf\{f_1(x), \ldots, f_J(x)\}.$$

On obtient alors immédiatement le théorème de structure suivant.

Théorème 3.1.3.1 *Si J appartient à $\mathbb{N}_0$ et si f_1, ..., f_J sont des fonctions réelles définies sur la partie A de $\mathbb{R}^n$, on peut obtenir les fonctions* $\sup\{f_1, \ldots, f_J\}$ *et* $\inf\{f_1, \ldots, f_J\}$ *au moyen d'un nombre fini de combinaisons linéaires et de recours aux parties positives, aux parties négatives ou aux modules de fonctions.* ∎

Ainsi, par exemple, si J est égal à 2, on a

$$\sup\{f_1, f_2\} = f_1 + (f_2 - f_1)_+ = f_1 + (f_1 - f_2)_- = \tfrac{1}{2}\left(f_1 + f_2 + |f_1 - f_2|\right)$$

et

$$\inf\{f_1, f_2\} = f_1 - (f_2 - f_1)_- = f_1 - (f_1 - f_2)_+ = \tfrac{1}{2}\left(f_1 + f_2 - |f_1 - f_2|\right).$$

c) Fonctions associées à un ensemble de fonctions. Si $\mathcal{F}$ est un ensemble de fonctions réelles définies sur une même partie A de $\mathbb{R}^n$ et si, pour tout $x \in A$, l'ensemble $\{\, f(x) : f \in \mathcal{F}\}$ est majoré (resp. minoré), on définit l'*enveloppe supérieure de $\mathcal{F}$* (resp. l'*enveloppe inférieure de $\mathcal{F}$)* notée

$$\sup \mathcal{F} \quad \text{ou} \quad \sup_{f \in \mathcal{F}} f \qquad (\text{resp. } \inf \mathcal{F} \quad \text{ou} \quad \inf_{f \in \mathcal{F}} f)$$

comme étant la fonction définie sur A par

$$(\sup \mathcal{F})(x) = \sup\{\, f(x) : f \in \mathcal{F}\} \quad (\text{resp. } (\inf \mathcal{F})(x) = \inf\{\, f(x) : f \in \mathcal{F}\})$$

pour tout $x \in A$.

Il convient de remarquer que, dans ce cas, on n'a plus en général de théorème de structure.

3.1.4 Fonctions bornées

Définitions. Une fonction f définie sur $A \subset \mathbb{R}^n$ est *bornée sur $A' \subset A$* s'il existe $C > 0$ tel que $|f(x)| \le C$ pour tout $x \in A'$; il revient au même de dire que $\{\, f(x) : x \in A'\}$ est un borné de $\mathbb{C}$. Si A' est égal à A, on dit que f est *borné*.

Une fonction réelle f définie sur $A \subset \mathbb{R}^n$ est *majorée sur $A' \subset A$* (resp. *minorée sur $A' \subset A$*) s'il existe un nombre réel M (resp. m) tel que $f(x) \le M$ (resp. $f(x) \ge m$) pour tout $x \in A'$; il revient au même de dire que $\{\, f(x) : x \in A'\}$ est une partie majorée (resp. minorée) de $\mathbb{R}$. On définit alors *la borne supérieure* (resp. *inférieure*) *de f sur A'* comme étant la borne supérieure (resp. inférieure) de cet ensemble $\{ f(x) : x \in A'\}$ et on la note

$$\sup_{x \in A'} f(x) \quad (\text{resp. } \inf_{x \in A'} f(x)).$$

Si A' est égal à A, on dit plus simplement que f est *majoré* (resp. *minoré*).

Remarque. Une fonction réelle f définie sur $A \subset \mathbb{R}^n$ et majorée (resp. minorée) sur $A' \subset A$ peut ne pas réaliser sa borne supérieure (resp. inférieure) sur A', c'est-à-dire qu'il n'existe pas nécessairement un point $x' \in A'$ tel que

$$f(x') = \sup_{x \in A'} f(x) \quad \left(\text{resp. } \; f(x') = \inf_{x \in A'} f(x)\right).$$

Exercice. Si f et g sont des fonctions réelles et bornées sur $A \subset \mathbb{R}^n$, établir que la fonction $f - g$ est bornée sur A et que

$$\left| \sup_{x \in A} f(x) - \sup_{x \in A} g(x) \right| \leq \sup_{x \in A} |f(x) - g(x)| \, .$$

Suggestion. D'une part, $f - g$ est borné sur A car, pour tout $x \in A$, on a évidemment

$$|f(x) - g(x)| \leq \sup_{y \in A} |f(y)| + \sup_{y \in A} |g(y)| \, .$$

D'autre part, établissons l'inégalité. Bien sûr, pour tout $x \in A$,

$$f(x) - g(x) \leq |f(x) - g(x)| \leq \sup_{y \in A} |f(y) - g(y)|$$

donc

$$f(x) - \sup_{z \in A} g(z) \leq \sup_{y \in A} |f(y) - g(y)| \, .$$

De là, on tire

$$\sup_{x \in A} f(x) - \sup_{x \in A} g(x) \leq \sup_{x \in A} |f(x) - g(x)| \, .$$

Comme on peut permuter les rôles de f et g, on conclut aussitôt.

Exercice. Soient A' et A'' des parties de $\mathbb{R}^{n'}$ et $\mathbb{R}^{n''}$ respectivement. Si la fonction f définie sur $A = A' \times A''$ est réelle et bornée, établir les relations

$$\begin{aligned} \sup_{x \in A} f(x) &= \sup_{y \in A'} \sup_{z \in A''} f(y,z) \\ \inf_{x \in A} f(x) &= \inf_{y \in A'} \inf_{z \in A''} f(y,z) \\ \sup_{z \in A''} \inf_{y \in A'} f(y,z) &\leq \inf_{y \in A'} \sup_{z \in A''} f(y,z). \end{aligned}$$

Suggestion. Etablissons la relation relative aux bornes supérieures. Bien sûr, pour tous $y \in A'$ et $z \in A''$, $f(y, z)$ est majoré par $\sup_{x \in A} f(x)$. De la sorte, pour tout $y \in A'$, $\sup_{z \in A''} f(y, z)$ est majoré par $\sup_{x \in A} f(x)$ et ainsi on a déjà la majoration

$$\sup_{x \in A} f(x) \geq \sup_{y \in A'} \sup_{z \in A''} f(y, z).$$

Inversement, pour tout $x \in A$, on a évidemment

$$f(x) \leq \sup_{y \in A'} \sup_{z \in A''} f(y, z),$$

d'où l'autre majoration et par conséquent la conclusion.

La relation relative aux bornes inférieures s'établit de même.

Passons à l'inégalité. Pour tous $y \in A'$ et $z \in A''$, on a évidemment

$$f(y, z) \leq \sup_{z'' \in A''} f(y, z''),$$

d'où on tire de suite que, pour tout $z \in A''$, on a

$$\inf_{y' \in A'} f(y', z) \leq \inf_{y' \in A'} \sup_{z'' \in A''} f(y', z'').$$

La conclusion est alors directe.

Remarque. On ne peut pas améliorer l'inégalité dans l'exercice précédent. Ainsi la fonction

$$f : \mathbb{R}^2 \to \mathbb{R} \quad (x, y) \mapsto \begin{cases} 0 & \text{si } x \geq y \\ 1 & \text{si } x < y \end{cases}$$

est telle que

$$\sup_{x \in \mathbb{R}} \inf_{y \in \mathbb{R}} f(x, y) = 0 \quad \text{et} \quad \inf_{y \in \mathbb{R}} \sup_{x \in \mathbb{R}} f(x, y) = 1.$$

3.1.5 Zéros d'une fonction

Définitions. Soit f une fonction définie sur $\mathbb{R}^n$.

Un *zéro de* f est un point x de $\mathbb{R}^n$ tel que $f(x) = 0$. L'ensemble d'annulation de f est l'ensemble

$$\mathcal{O}_f = \{\, x \in \mathbb{R}^n : f(x) = 0 \,\}.$$

Un zéro de f est *identique* s'il appartient à $\mathcal{O}_f^\circ$, c'est-à-dire s'il est le centre d'une boule dont tous les éléments sont des zéros de f. Un zéro non identique est dit *accidentel*. Le *support de* f est l'ensemble

$$\operatorname{supp}(f) = \{\, x \in \mathbb{R}^n : f(x) \neq 0 \,\}^-.$$

On a donc

$$\operatorname{supp}(f) = \mathbb{R}^n \setminus \mathcal{O}_f^\circ.$$

Dès lors, si J appartient à $\mathbb{N}_0$, si $f_1, \ldots, f_J$ sont des fonctions définies sur $\mathbb{R}^n$ et si $c_1, \ldots, c_J$ appartiennent à $\mathbb{C}$,

$$\operatorname{supp}\left(\sum_{j=1}^{J} c_j f_j\right) \subset \bigcup_{j=1}^{J} \operatorname{supp}(f_j) \quad \text{et} \quad \operatorname{supp}\left(\prod_{j=1}^{J} f_j\right) \subset \bigcap_{j=1}^{J} \operatorname{supp}(f_j).$$

Quelques exemples simples montrent qu'on ne peut pas améliorer ces inclusions. Considérons les fonctions f et g définies sur $\mathbb{R}$ par

$$f(x) = \left\{\begin{array}{ll} 0 & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{array}\right\} \quad \text{et} \quad g(x) = \left\{\begin{array}{ll} 1 & \text{si } x \leq 0 \\ 0 & \text{si } x > 0. \end{array}\right\}$$

On a alors

$$\operatorname{supp}(f) = [0, +\infty[\quad \text{et} \quad \operatorname{supp}(g) =]-\infty, 0]$$

ainsi que

$$\operatorname{supp}(f+g) = \mathbb{R}, \quad \operatorname{supp}(f-f) = \emptyset \quad \text{et} \quad \operatorname{supp}(fg) = \emptyset.$$

3.1.6 Fonctions caractéristiques

Définition. La *fonction caractéristique d'une partie A de* $\mathbb{R}^n$ est la fonction

$$\chi_A \colon \mathbb{R}^n \to \mathbb{R} \quad x \mapsto \left\{\begin{array}{ll} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{array}\right\}.$$

Si A et A' sont des parties de $\mathbb{R}^n$, on vérifie de suite que

$$\begin{array}{rclrcl} \chi_\emptyset & = & 0, & \chi_{\mathbb{R}^n \setminus A} & = & \chi_{\mathbb{R}^n} - \chi_A, \\ A \subset A' & \Leftrightarrow & \chi_A \leq \chi_{A'}, & A = A' & \Leftrightarrow & \chi_A = \chi_{A'}, \\ \operatorname{supp}(\chi_A) & = & A^-. & & & \end{array}$$

De plus, si J appartient à $\mathbb{N}_0$ et si $A_1, \ldots, A_J$ sont des parties de $\mathbb{R}^n$,

$$\begin{array}{rcl} \chi_{\bigcap_{j=1}^J A_j} & = & \prod_{j=1}^{J} \chi_{A_j} = \inf\{\chi_{A_1}, \ldots, \chi_{A_J}\}, \\ \chi_{\bigcup_{j=1}^J A_j} & = & \sup\{\chi_{A_1}, \ldots, \chi_{A_J}\}. \end{array}$$

3.2 Limite des valeurs d'une fonction

3.2.1 Définition

Définitions. Soit f une fonction définie sur une partie non vide A de $\mathbb{R}^n$.

Désignons par ξ un point x_0 de A^- ou ∞, cette dernière possibilité n'étant envisagée que si A n'est pas borné. Il s'ensuit que tout voisinage de ξ est d'intersection non vide avec A.

Désignons par γ un nombre complexe ou ∞.

Alors f *converge vers* γ *si* x *converge vers* ξ (on dit aussi "*tend*" en lieu et place de "converge") si, pour tout voisinage U de γ, il existe un voisinage V de ξ tel que $f(V \cap A) \subset U$. On dit aussi que γ est la *limite de* f *pour* $x \to \xi$ et on écrit

$$\lim_{x \to \xi} f(x) = \gamma \quad \text{ou} \quad f(x) \to \gamma \quad \text{si} \quad x \to \xi.$$

Cette notion est fondamentale, aussi nous allons expliciter les différents cas qui peuvent se présenter. A cet effet, rappelons que

a) une partie A de $\mathbb{R}^n$ est un voisinage de $x_0 \in \mathbb{R}^n$ si et seulement si elle contient une boule centrée en x_0, c'est-à-dire si et seulement s'il existe un rayon $r > 0$ tel que tout $x \in \mathbb{R}^n$ vérifiant $|x - x_0| \leq r$ appartienne à A.

b) une partie A de $\mathbb{R}^n$ est un voisinage de ∞ si et seulement si elle contient le complémentaire d'une boule centrée en 0, c'est-à-dire si et seulement s'il existe un rayon $R > 0$ tel que tout $x \in \mathbb{R}^n$ vérifiant $|x| \geq R$ appartienne à A.

Dès lors, suivant que ξ est le point x_0 de $\mathbb{R}^n$ ou ∞, et que γ est le nombre complexe c ou ∞, la définition s'explicite de la manière suivante:

$$\begin{array}{lcl}
\lim\limits_{x \to x_0} f(x) = c & \Leftrightarrow & \forall \varepsilon > 0, \exists \eta > 0 \text{ tel que, } \forall x \in A \text{ vérifiant l'inégalité} \\
 & & |x - x_0| \leq \eta, \text{ on a } |f(x) - c| \leq \varepsilon, \\
\lim\limits_{x \to x_0} f(x) = \infty & \Leftrightarrow & \forall N > 0, \exists \eta > 0 \text{ tel que, } \forall x \in A \text{ vérifiant l'inégalité} \\
 & & |x - x_0| \leq \eta, \text{ on a } |f(x)| \geq N, \\
\lim\limits_{x \to \infty} f(x) = c & \Leftrightarrow & \forall \varepsilon > 0, \exists M > 0 \text{ tel que, } \forall x \in A \text{ vérifiant l'inégali-} \\
 & & \text{té } |x| \geq M, \text{ on a } |f(x) - c| \leq \varepsilon, \\
\lim\limits_{x \to \infty} f(x) = \infty & \Leftrightarrow & \forall N > 0, \exists M > 0 \text{ tel que, } \forall x \in A \text{ vérifiant l'inéga-} \\
 & & \text{lité } |x| \geq M, \text{ on a } |f(x)| \geq N.
\end{array}$$

Le cas où la fonction f est réelle, est particulier.

D'une part, dans $\mathbb{R}$, nous avons introduit les concepts de $+\infty$ et $-\infty$, et nous en avons défini la notion de voisinage: une partie de $\mathbb{R}$ est un voisinage de $+\infty$ (resp. $-\infty$) si et seulement si elle contient un intervalle du type $[N, +\infty[$ (resp. $]-\infty, N]$)

avec $N > 0$. Dès lors, si f est une fonction réelle sur A, suivant que ξ est le point x_0 de $\mathbb{R}^n$ ou ∞, et que γ désigne $+\infty$ ou $-\infty$, la définition s'explicite selon:

$$\begin{aligned}
\lim_{x\to x_0} f(x) = +\infty &\Leftrightarrow \forall N > 0, \exists \eta > 0 \text{ tel que, } \forall x \in A \text{ vérifiant l'inégalité } \\ &\quad |x - x_0| \leq \eta, \text{ on a } f(x) \geq N, \\
\lim_{x\to x_0} f(x) = -\infty &\Leftrightarrow \forall N > 0, \exists \eta > 0 \text{ tel que, } \forall x \in A \text{ vérifiant l'inégalité } \\ &\quad |x - x_0| \leq \eta, \text{ on a } f(x) \leq -N, \\
\lim_{x\to \infty} f(x) = +\infty &\Leftrightarrow \forall N > 0, \exists M > 0 \text{ tel que, } \forall x \in A \text{ vérifiant l'inéga-} \\ &\quad \text{lité } |x| \geq M, \text{ on a } f(x) \geq N, \\
\lim_{x\to \infty} f(x) = -\infty &\Leftrightarrow \forall N > 0, \exists M > 0 \text{ tel que, } \forall x \in A \text{ vérifiant l'inéga-} \\ &\quad \text{lité } |x| \geq M, \text{ on a } f(x) \leq -N.
\end{aligned}$$

D'autre part, on peut préciser la manière suivant laquelle la fonction réelle f converge si x tend vers ξ: si f est une fonction réelle sur A et si r est un nombre réel, f *tend* (ou *converge*) *vers* r^+ (resp. r^-) *si* x *tend* (ou *converge*) *vers* ξ si les deux conditions suivantes sont réalisées:

a) f tend vers r si x tend vers ξ,

b) il existe un voisinage V de ξ tel que $f(A \cap V) \subset [r, +\infty[$ (resp. $]-\infty, r]$).

On écrit alors

$$\lim_{x\to\xi} f(x) = r^+ \text{ (resp. } \lim_{x\to\xi} f(x) = r^-)$$

et on dit que r^+ (resp. r^-) est la *limite de* f *en* ξ. Suivant que ξ est un point x_0 de $\mathbb{R}^n$ ou ∞, on a donc explicitement:

$$\begin{aligned}
\lim_{x\to x_0} f(x) = r^+ &\Leftrightarrow \forall \varepsilon > 0, \exists \eta > 0 \text{ tel que, } \forall x \in A \text{ vérifiant l'inégalité } \\ &\quad |x - x_0| \leq \eta, \text{ on a } 0 \leq f(x) - r \leq \varepsilon, \\
\lim_{x\to x_0} f(x) = r^- &\Leftrightarrow \forall \varepsilon > 0, \exists \eta > 0 \text{ tel que, } \forall x \in A \text{ vérifiant l'inégalité } \\ &\quad |x - x_0| \leq \eta, \text{ on a } 0 \leq r - f(x) \leq \varepsilon, \\
\lim_{x\to \infty} f(x) = r^+ &\Leftrightarrow \forall \varepsilon > 0, \exists M > 0 \text{ tel que, } \forall x \in A \text{ vérifiant l'inégalité } \\ &\quad |x| \geq M, \text{ on a } 0 \leq f(x) - r \leq \varepsilon, \\
\lim_{x\to \infty} f(x) = r^- &\Leftrightarrow \forall \varepsilon > 0, \exists M > 0 \text{ tel que, } \forall x \in A \text{ vérifiant l'inégalité } \\ &\quad |x| \geq M, \text{ on a } 0 \leq r - f(x) \leq \varepsilon.
\end{aligned}$$

On introduit aussi la notion de limite restreinte.

Définitions. Soit f une fonction définie sur une partie non vide A de $\mathbb{R}^n$. Soit A' une partie non vide de A. Soit ξ un point de $(A')^-$ ou ∞, cette dernière possibilité n'étant envisagée que si A' n'est pas borné. Soit enfin α un nombre complexe c ou ∞. Si f est une fonction réelle, on introduit aussi les valeurs $\alpha = r^+$, $\alpha = r^-$, $\alpha = +\infty$ et $\alpha = -\infty$ avec $r \in \mathbb{R}$.

Alors f *tend vers* α *si* x *tend vers* ξ *dans* A' si la restriction de f à A' tend vers α si x tend vers ξ, c'est-à-dire si, pour tout voisinage U de α dans $\mathbb{C}$ (resp. dans $\mathbb{R}$ si on a $\alpha = r^+$, $\alpha = r^-$, $\alpha = +\infty$ ou $\alpha = -\infty$), il existe un voisinage V de ξ dans $\mathbb{R}^n$ tel que $f(A' \cap V) \subset U$. On écrit

$$\lim_{x \to \xi,\, x \in A'} f(x) = \alpha.$$

En particulier, si n est égal à 1 et si $\xi = x_0$ appartient à l'adhérence de l'ensemble $A \cap]x_0, +\infty[$ (resp. $A \cap]-\infty, x_0[$), on introduit l'expression "*f tend vers α si x tend vers x_0 par la droite*" pour "f tend vers α si x tend vers x_0 dans $A \cap]x_0, +\infty[$" (resp. "*f tend vers α si x tend vers x_0 par la gauche*" pour "f tend vers α si x tend vers x_0 dans $A \cap]-\infty, x_0[$"). On écrit

$$\lim_{x \to x_0^+} f(x) = \alpha \quad \text{(resp.} \quad \lim_{x \to x_0^-} f(x) = \alpha)$$

et on dit que f *admet une limite à droite* (resp. *à gauche*) *en* x_0.

De même, si n est égal à 1 et si, pour tout $N > 0$, l'ensemble $A \cap]N, +\infty[$ (resp. $A \cap]-\infty, N[$) n'est pas vide, on introduit l'expression "*f tend vers α si x tend vers $+\infty$*" pour "f tend vers α si x tend vers ∞ dans $A \cap]0, +\infty[$" (resp. "*f tend vers α si x tend vers $-\infty$*" pour "f tend vers α si x tend vers ∞ dans $A \cap]-\infty, 0[$"). On écrit

$$\lim_{x \to +\infty} f(x) = \alpha \quad \text{(resp.} \quad \lim_{x \to -\infty} f(x) = \alpha).$$

Exercice. *Etablir que*

$$\lim_{(x,y) \to (0,0)} \frac{x^3 + y^3}{x^2 + y^2} = 0.$$

Suggestion. On remarque tout d'abord que $(x^3 + y^3)/(x^2 + y^2)$ est une fonction définie sur $\mathbb{R}^2 \setminus \{(0,0)\}$. Dès lors, le point $(0,0)$ peut être considéré comme étant un point ξ car il appartient à l'adhérence de cet ensemble $\mathbb{R}^2 \setminus \{(0,0)\}$. Cela étant, les majorations $|x^3| \le |(x,y)|^3$, $|y^3| \le |(x,y)|^3$ et $x^2 + y^2 = |(x,y)|^2$ permettent d'affirmer que, pour tout $\varepsilon > 0$, il existe $\eta = \varepsilon/2 > 0$ tel que, pour tout $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$ vérifiant $|(x,y)| \le \eta$, on a

$$\left| \frac{x^3 + y^3}{x^2 + y^2} \right| \le 2\,|(x,y)| \le \varepsilon.$$

3.2.2 Critère par les suites

Critère 3.2.2.1 (par les suites) *Soit f une fonction définie sur une partie non vide A de $\mathbb{R}^n$. Soit ξ un point de A^- ou ∞, cette dernière possibilité n'étant* *envisagée que si A n'est pas borné. Soit enfin γ un nombre complexe ou ∞. Alors les assertions suivantes sont équivalentes:*

(a) *f tend vers γ si x tend vers ξ,*

(b) *pour toute suite x_m de A qui tend vers ξ, la suite $f(x_m)$ tend vers γ.*

Preuve. Nous allons établir l'énoncé dans le cas où ξ est un point x_0 de A^- et γ un nombre complexe c. Tous les autres cas se démontrent de manière analogue.

(a) $\Rightarrow$ (b). Vu (a), pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$(x \in A,\ |x - x_0| \leq \eta) \Rightarrow |f(x) - c| \leq \varepsilon.$$

Comme il existe un entier M tel que $m \geq M$ entraîne $|x_m - x_0| \leq \eta$, il vient donc $|f(x_m) - c| \leq \varepsilon$ et on conclut aussitôt.

(b) $\Rightarrow$ (a). Si ce n'est pas le cas, il existe $\varepsilon > 0$ tel que, pour tout $\eta > 0$, il existe $x \in A$ tel que $|x - x_0| \leq \eta$ et $|f(x) - c| \geq \varepsilon$. Dès lors, pour tout $m \in \mathbb{N}_0$, il existe $x_m \in A$ tel que $|x_m - x_0| \leq 1/m$ et $|f(x_m) - c| \geq \varepsilon$. D'où une contradiction car cette suite x_m est une suite de A qui converge vers x_0 alors que la suite $f(x_m)$ ne converge pas vers c.∎

Proposition 3.2.2.2 *Soit f une fonction définie sur une partie non vide A de $\mathbb{R}^n$ et soit ξ un point de A^- ou ∞, cette dernière possibilité n'étant envisagée que si A n'est pas borné.*

a) *Il existe une suite x_m de A qui converge vers ξ.*

b) *Si ξ appartient à A et si f admet une limite si x tend vers ξ, cette limite est égale à $f(\xi)$.*

c) *Si, pour toute suite x_m de A qui converge vers ξ, la suite $f(x_m)$ converge ou converge vers l'infini, alors toutes ces limites sont égales et f tend vers cette valeur si x tend vers ξ.*

Preuve. a) est immédiat et connu.

b) De fait, la suite x_m définie par $x_m = \xi$ pour tout $m \in \mathbb{N}_0$ est une suite de A qui converge vers ξ et on a évidemment $f(x_m) \to f(\xi)$.

c) De fait, si les suites x'_m et x''_m de A convergent vers ξ et si les suites $f(x'_m)$ et $f(x''_m)$ convergent respectivement vers γ' et γ'', alors la suite x_m définie par

$$x_{2m-1} = x'_m \text{ et } x_{2m} = x''_m, \quad \forall m \in \mathbb{N}_0,$$

est une suite de A qui converge vers ξ. Cela étant, la suite $f(x_m)$ converge alors que les suites $f(x'_m)$ et $f(x''_m)$ en sont des sous-suites, ce qui suffit.∎

Remarque. Le critère par les suites fournit des moyens commodes pour prouver que f n'admet pas de limite en ξ: il suffit pour cela de trouver

a) soit une suite x_m de A qui converge vers ξ et telle que la suite $f(x_m)$ ne converge pas,

b) soit deux suites x_m et y_m de A qui convergent vers ξ et telles que les suites $f(x_m)$ et $f(y_m)$ convergent vers des limites différentes.

Ainsi la fonction $\chi_{]0,+\infty[}$ n'admet pas de limite en 0. En effet, la suite $(-1)^m/m$ converge vers 0 alors que la suite $\chi_{]0,+\infty[}((-1)^m/m)$ ne converge pas.

Exercice. *Etablir que la fonction $xy/(x^2+y^2)$ n'admet pas de limite à l'origine.*

Suggestion. On remarque tout d'abord que $xy/(x^2+y^2)$ est une fonction définie sur $A = \mathbb{R}^2 \setminus \{(0,0)\}$. Comme le point $(0,0)$ appartient à l'adhérence de A, l'énoncé a un sens. De plus, pour toute suite $r_m \in \mathbb{R} \setminus \{0\}$ convergeant vers 0 et tout $r > 0$, la suite (r_m, rr_m) est une suite de A qui converge vers $(0,0)$ et telle que

$$\frac{r_m r r_m}{r_m^2 + r^2 r_m^2} \to \frac{r}{1+r^2}.$$

D'où la conclusion en considérant deux valeurs différentes de r.

3.2.3 Critère par les limites restreintes

Proposition 3.2.3.1 *Si f est une fonction définie sur la partie non vide A de $\mathbb{R}^n$, si $\{A_1, \ldots, A_J\}$ est une partition finie de A et si, pour tout $j \in \{1, \ldots, J\}$, la restriction de f à A_j converge vers γ si x tend vers ξ, alors f converge vers γ si x tend vers ξ.*

Preuve. Etablissons le cas où ξ est un point x_0 de $\mathbb{R}^n$ et γ un nombre complexe c; les autres cas se démontrent de manière analogue.

Pour tout $\varepsilon > 0$, il existe $\eta_1, \ldots, \eta_J > 0$ tels que

$$(x \in A_j,\ |x - x_0| \le \eta_j) \Rightarrow |f(x) - c| \le \varepsilon.$$

Dès lors, en posant $\eta = \inf\{\eta_1, \ldots, \eta_J\}$, il vient

$$(x \in A,\ |x - x_0| \le \eta) \Rightarrow |f(x) - c| \le \varepsilon.$$

D'où la conclusion.∎

Corollaire 3.2.3.2 *Si f est une fonction définie sur $A \subset \mathbb{R}$, si x_0 n'appartient pas à A et si γ est à la fois limite à droite et à gauche de f en x_0, alors f converge vers γ si x tend vers x_0.*

Preuve. De fait, les ensembles

$$A_1 = \{ x \in A : x > x_0 \} \quad \text{et} \quad A_2 = \{ x \in A : x < x_0 \}$$

constituent une partition de A et dire que f converge vers γ à droite ou à gauche de x_0 revient à dire que la restriction de f à A_1 ou à A_2 respectivement converge vers γ.∎

Corollaire 3.2.3.3 *Si f est une fonction définie sur $A \subset \mathbb{R}$, si x_0 appartient à A et si $f(x_0)$ est à la fois limite à droite et à gauche de f en x_0, alors f converge vers $f(x_0)$ si x tend vers x_0.*∎

3.2.4 Critère de Cauchy

Dans les applications, la fonction f, son ensemble de définition et le point ξ sont toujours donnés. Mais, en général, il n'en est pas de même pour γ. Nous nous retrouvons donc en présence d'un problème comparable à celui posé par la convergence des suites.

Critère 3.2.4.1 (Cauchy) *Soit f une fonction définie sur la partie non vide A de $\mathbb{R}^n$ et soit ξ un point de A^- ou ∞, cette dernière possibilité n'étant envisagée que si A n'est pas borné. Alors les assertions suivantes sont équivalentes:*

(a) $\lim_{x \to \xi} f(x)$ *existe et est fini,*

(b) *pour toutes suites x_m et y_m de A qui convergent vers ξ,*

$$|f(x_m) - f(y_m)| \to 0,$$

(c) *pour tout $\varepsilon > 0$, il existe un voisinage V de ξ tel que*

$$\sup_{x,y \in V \cap A} |f(x) - f(y)| \leq \varepsilon.$$

Preuve. (a) $\Rightarrow$ (b). Supposons que f converge vers $c \in \mathbb{C}$ si x tend vers ξ. Vu le critère par les suites, on a $|f(x_m) - c| \to 0$ et $|f(y_m) - c| \to 0$, ce qui suffit.

(b) $\Rightarrow$ (c) s'établit comme (b) $\Rightarrow$ (a) dans le critère par les suites.

(c) $\Rightarrow$ (a). De fait, si la suite x_m de A converge vers ξ, l'assertion (c) établit que la suite $f(x_m)$ est de Cauchy, donc converge. D'où la conclusion à partir de la proposition 3.2.2.2.∎

3.2.5 Propriétés de la limite

Les propriétés de la limite des valeurs d'une fonction dépendent bien sûr des fonctions considérées et des propriétés des limites des suites convergentes.

Formulons l'hypothèse générale suivante, valable dans tout ce paragraphe.

Hypothèse. Dans ce paragraphe, sauf mention explicite du contraire,

a) J désigne un élément de $\mathbb{N}_0$,

b) f, g, f_1, …, f_J sont des fonctions définies sur une même partie A non vide de $\mathbb{R}^n$,

c) ξ est un élément de A^- ou ∞, cette dernière possibilité n'étant envisagée que si A n'est pas borné,

d) on a

$$\lim_{x\to\xi} f(x) = a, \quad \lim_{x\to\xi} g(x) = b$$

et

$$\lim_{x\to\xi} f_j(x) = a_j, \quad \forall j \in \{1, \ldots, J\}.$$

Théorème 3.2.5.1 a) *Si c_1, …, c_J, a_1, …, a_J sont des nombres complexes, il vient*

$$\lim_{x\to\xi} \sum_{j=1}^{J} c_j f_j(x) = \sum_{j=1}^{J} c_j a_j.$$

Si c_1, …, c_J, a_1, …, a_{J-1} sont des nombres complexes, si a_J est égal à ∞ et si c_J diffère de 0, alors il vient

$$\lim_{x\to\xi} \sum_{j=1}^{J} c_j f_j(x) = \infty.$$

b) *Si a_1, …, a_J sont des nombres complexes, il vient*

$$\lim_{x\to\xi} \prod_{j=1}^{J} f_j(x) = \prod_{j=1}^{J} a_j.$$

Si les a_j diffèrent tous de 0 et si l'un au moins est égal à ∞, il vient

$$\lim_{x\to\xi} \prod_{j=1}^{J} f_j(x) = \infty.$$

c) *Si a est un nombre complexe, si b diffère de 0 et si g ne s'annule en aucun point de A, on a*

$$\lim_{x\to\xi} \frac{f(x)}{g(x)} = \frac{a}{b} \quad \textit{ou} \quad 0$$

suivant que b est un nombre complexe ou ∞.

Preuve. Cela résulte immédiatement des propriétés correspondantes des suites convergentes.∎

Théorème 3.2.5.2 *Si J appartient à $\mathbb{N}_0$, si $f_1, \ldots, f_J$ sont des fonctions réelles sur $A \subset \mathbb{R}^n$, si f est une fonction définie sur une partie A' de $\mathbb{R}^J$ contenant l'ensemble de variation des $f_1, \ldots, f_J$ et si on a*

$$\lim_{y \to (a_1, \ldots, a_J)} f(y) = c \quad ou \quad \lim_{y \to \infty} f(y) = c$$

suivant que les a_j sont tous des nombres réels ou non, alors il vient

$$\lim_{x \to \xi} f(f_1(x), \ldots, f_J(x)) = c.$$

Preuve. Si la suite x_m de A converge vers ξ, la suite $(f_1(x_m), \ldots, f_J(x_m))$ est en effet une suite de A' qui converge vers $(a_1, \ldots, a_J)$ ou ∞ suivant que les a_j sont tous des nombres réels ou non.∎

Proposition 3.2.5.3 *On a*

$$\begin{aligned}
\lim_{x \to \xi} \overline{f}(x) &= \begin{cases} \overline{a} \text{ si } a \in \mathbb{C}, \\ \infty \text{ si } a = \infty, \end{cases} \\
\lim_{x \to \xi} |f|(x) &= \begin{cases} |a| \text{ si } a \in \mathbb{C}, \\ \infty \text{ si } a = \infty, \end{cases} \\
\lim_{x \to \xi} \left\{ \begin{array}{c} \Re \\ \Im \end{array} \right\} f(x) &= \left\{ \begin{array}{c} \Re \\ \Im \end{array} \right\} a \text{ si } a \in \mathbb{C}.
\end{aligned}$$

De plus, si f, $f_1, \ldots, f_J$ sont des fonctions réelles et si les a, $a_1, \ldots, a_J$ sont des nombres réels, il vient

$$\begin{aligned}
\lim_{x \to \xi} \left\{ \begin{array}{c} f_+ \\ f_- \end{array} \right\} (x) &= \left\{ \begin{array}{c} a_+ \\ a_- \end{array} \right\}, \\
\lim_{x \to \xi} \left\{ \begin{array}{c} \sup \\ \inf \end{array} \right\} \{f_1(x), \ldots, f_J(x)\} &= \left\{ \begin{array}{c} \sup \\ \inf \end{array} \right\} \{a_1, \ldots, a_J\}. \blacksquare
\end{aligned}$$

Proposition 3.2.5.4 *Si a diffère de b, il existe un voisinage V de ξ tel que*

$$x \in V \cap A \Rightarrow f(x) \neq g(x).$$

En particulier, si a diffère du nombre complexe c, il existe un voisinage V de ξ tel que $c \notin f(V \cap A)$.

Si f et g sont des fonctions réelles et si $a < b$, il existe un voisinage V de ξ tel que

$$x \in V \cap A \Rightarrow f(x) < g(x).$$

En particulier, si f est une fonction réelle et si $r \in \mathbb{R}$ vérifie $a < r$ (resp. $a > r$)*, il existe un voisinage V de ξ tel que $f(V \cap A) \subset]-\infty, r[$* (resp. $]r, +\infty[$).∎

Proposition 3.2.5.5 *Si, pour tout voisinage V de ξ, il existe $x \in V \cap A$ tel que $f(x) = c$, on a $a = c$.*

Si f est une fonction réelle et si, pour tout voisinage V de ξ, il existe un point x de $V \cap A$ tel que $f(x) \leq c$, on a $a \leq c$.∎

Chapitre 4

Continuité

4.1 Définition

4.1.1 Généralités

Définitions. Soit f une fonction définie sur une partie non vide A de $\mathbb{R}^n$.

Cette fonction f est *continue en* $x_0 \in A$ si $\lim_{x \to x_0} f(x)$ existe et vaut, par conséquent, $f(x_0)$. Un tel point $x_0 \in A$ est appelé *point de continuité de* f.

Si f n'est pas continu en $x_0 \in A$, on dit que f est *discontinu en* x_0 ou que x_0 est un *point de discontinuité de* f.

On peut aussi, en recourant aux limites restreintes, introduire des notions de continuité partielle; il s'agit alors de bien spécifier les parties A' de A considérées. Ici, nous n'allons introduire que les notions suivantes: une fonction f définie sur une partie non vide A de $\mathbb{R}$ est

continue à droite en $x_0 \in A$ si $\lim_{x \to x_{0+}} f(x)$ existe et vaut $f(x_0)$,

continue à gauche en $x_0 \in A$ si $\lim_{x \to x_{0-}} f(x)$ existe et vaut $f(x_0)$.

Cela étant, le critère par les limites restreintes donne le résultat suivant: *la fonction f définie sur $A \subset \mathbb{R}$ est continue en $x_0 \in A$ si et seulement si elle est continue à droite et à gauche en x_0.*

Définition. La fonction f est *continue sur* A si elle est définie sur A et continue en tout point de A.

L'ensemble des fonctions continues sur A est noté $\mathrm{C}_0(A)$.

Par définition, on a donc les critères suivants.

Critère 4.1.1.1 (en "ε, η") *Une fonction f définie sur $A \subset \mathbb{R}^n$ est continue sur A si et seulement si, pour tout $x \in A$ et tout $\varepsilon > 0$, il existe $\eta > 0$ tel que*

$$(y \in A,\ |x - y| \leq \eta) \Rightarrow |f(x) - f(y)| \leq \varepsilon. \blacksquare$$

Critère 4.1.1.2 (par les suites) *Une fonction f définie sur $A \subset \mathbb{R}^n$ est continue sur A si et seulement si, pour tout $x \in A$ et toute suite x_m de A qui converge vers x, on a $f(x_m) \to f(x)$.*∎

Définition. On est aussi amené à considérer la notion suivante: une fonction f définie sur $A \subset \mathbb{R}^n$ est *uniformément continue sur A* si, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$(x, y \in A,\ |x - y| \leq \eta) \Rightarrow |f(x) - f(y)| \leq \varepsilon.$$

Bien sûr, on a alors le résultat que voici.

Proposition 4.1.1.3 *Toute fonction uniformément continue sur $A \subset \mathbb{R}^n$ est continue sur A.*∎

Remarques. a) La réciproque de cette proposition est fausse. Ainsi, on vérifie aisément que la fonction $f\colon\ ]0, +\infty[\ \to \mathbb{R} \quad x \mapsto 1/x$ est continue. Cependant f n'est pas uniformément continu sur $]0, +\infty[$ car on a

$$|f(1/m) - f(1/(m+1))| = 1, \quad \forall m \in \mathbb{N}_0,$$

alors que la suite $1/m - 1/(m+1)$ converge vers 0.

b) Nous en établirons cependant une réciproque partielle fort importante au paragraphe 4.2.3.

c) Ceci nous permet d'insister sur le fait qu'on ne peut permuter impunément les quantificateurs $\exists$ et $\forall$. A cet égard, remarquons qu'une permutation supplémentaire de ces quantificateurs conduirait aux fonctions f définies sur $\mathbb{R}^n$ pour lesquelles il existe $\eta > 0$ tel que, pour tout $\varepsilon > 0$, on aurait

$$(x, y \in \mathbb{R}^n,\ |x - y| \leq \eta) \Rightarrow |f(x) - f(y)| \leq \varepsilon.$$

Cependant cette notion ne présente aucun intérêt car on a tôt fait d'établir qu'il s'agit là d'une fonction constante sur $\mathbb{R}^n$.

Remarque. Soient n, n' et n'' des éléments de $\mathbb{N}_0$ tels que $n = n' + n''$. Si ν est une permutation de $\{1, \ldots, n\}$, alors, à tout $x \in \mathbb{R}^n$, on associe les points

$$x' = (x_{\nu(1)}, \ldots, x_{\nu(n')}) \in \mathbb{R}^{n'}, \quad x'' = (x_{\nu(n'+1)}, \ldots, x_{\nu(n)}) \in \mathbb{R}^{n''}.$$

Inversement, à $y \in \mathbb{R}^{n'}$ et $z \in \mathbb{R}^{n''}$, on peut associer un point x de $\mathbb{R}^n$ tel que $x' = y$ et $x'' = z$.

De plus, si I est un intervalle de $\mathbb{R}^n$, on vérifie aisément que les ensembles $I' = \{\, x' : x \in I\}$ et $I'' = \{\, x'' : x \in I\}$ sont des intervalles de $\mathbb{R}^{n'}$ et $\mathbb{R}^{n''}$ respectivement.

Il s'ensuit que si f est une fonction définie sur I, on peut lui associer une fonction F définie sur $I' \times I''$ par $F(x', x'') = f(x)$ pour tout $x \in I$.

Cela étant, on vérifie de suite que, si f est continu sur I, on a

$$\lim_{x\to x_0} f(x) = \lim_{x'\to x'_0} \lim_{x''\to x''_0} F(x',x'') = \lim_{x''\to x''_0} \lim_{x'\to x'_0} F(x',x'')$$

pour tout $x_0 \in I^\circ$. Remarquons bien que la continuité de f sur I est essentielle, ne serait-ce que pour assurer un sens aux limites

$$\lim_{x''\to x''_0} F(x',x'') \quad \text{et} \quad \lim_{x'\to x'_0} F(x',x'')$$

pour tout $x' \in I'^\circ$ et tout $x'' \in I''^\circ$ respectivement. La continuité de f en x_0 ne suffit pas!

4.1.2 Exemples fondamentaux

Voici tout d'abord un exemple pathologique qui montre bien que la continuité d'une fonction est en fait une propriété très spéciale.

Exemple. *La fonction f définie sur $]0,1[$ par*
$f(x) = 0$ si x est irrationnel,
$f(x) = 1/q$ si x est égal à p/q avec p, $q \in \mathbb{N}_0$ premiers entre eux,
est continue en tout point irrationnel de $]0,1[$ et discontinue en tout point rationnel de $]0,1[$. Pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}_0$ tel que $\varepsilon \geq 1/N$. Cela étant, il existe au plus $1 + \cdots + N$ points rationnels x de $]0,1[$ tels que $f(x) \geq \varepsilon$. Dès lors, si x_0 est un point irrationnel de $]0,1[$, il existe bien sûr un nombre $r > 0$ tel que l'intervalle $[x_0 - r, x_0 + r]$ ne contienne aucun de ces points rationnels, donc tel que $|f(x) - f(x_0)| \leq \varepsilon$ pour tout $x \in\,]0,1[$ tel que $|x_0 - x| \leq r$ et ainsi f est continu en x_0. Par contre, si y_0 est un point rationnel de $]0,1[$, on a $f(y_0) > 0$ bien qu'il existe une suite y_m de nombres irrationnels appartenant à $]0,1[$ qui converge vers y_0, donc tels que la suite $f(y_m) = 0$ converge vers 0. Cela étant, y_0 est un point de discontinuité de la fonction f.

Il convient de pouvoir décider rapidement si une fonction est continue ou non. Pour ce faire, nous allons procéder en deux étapes:

a) nous allons d'abord donner des exemples fondamentaux de fonctions continues (cette liste sera augmentée au chapitre 6, lors de l'étude des fonctions élémentaires),

b) nous allons ensuite établir des théorèmes qui permettent d'engendrer des fonctions continues à partir de fonctions continues.

Exemples fondamentaux. a) *La fonction 0 est continue sur $\mathbb{R}^n$.*

b) *La fonction $\chi_{\mathbb{R}^n}$ est continue sur $\mathbb{R}^n$.*

c) *Pour tout $j \in \{1,\ldots,n\}$, la fonction $[\cdot]_j$ est continue sur $\mathbb{R}^n$.* Cela résulte aussitôt de

$$|x_j - y_j| \leq |x - y|\,, \quad \forall x, y \in \mathbb{R}^n, \forall j \in \{1,\ldots,n\}.$$

d) *Pour toute partie non vide A de $\mathbb{R}^n$, la fonction $\mathrm{d}(\cdot, A)$ est continue sur $\mathbb{R}^n$. En particulier la fonction $|\cdot|$ est continue sur $\mathbb{R}^n$.* Cela résulte aussitôt de

$$|\mathrm{d}(x, A) - \mathrm{d}(y, A)| \leq |x - y|\,, \quad \forall x, y \in \mathbb{R}^n.$$

Remarque. On vérifie de suite que ces exemples fondamentaux de fonctions continues sont aussi des exemples de fonctions uniformément continues sur $\mathbb{R}^n$.

4.1.3 Critère

Le résultat qui suit peut être fortement généralisé.

Proposition 4.1.3.1 *Si f est une fonction définie sur $\mathbb{R}^n$, les assertions suivantes sont équivalentes:*
(a) *f est continu sur $\mathbb{R}^n$,*
(b) *l'image inverse par f de tout ouvert de $\mathbb{C}$ est ouverte,*
(c) *l'image inverse par f de tout fermé de $\mathbb{C}$ est fermée.*

Preuve. (a) $\Rightarrow$ (b). Soit Ω un ouvert de $\mathbb{C}$. Pour tout $x \in f^{-1}(\Omega)$, on a $f(x) \in \Omega$ et il existe $r > 0$ tel que la boule $\{\, z \in \mathbb{C} : |z - f(x)| \leq r\}$ soit incluse dans Ω. Cela étant, vu la continuité de f, il existe $\eta > 0$ tel que

$$(y \in \mathbb{R}^n,\ |x - y| \leq \eta) \Rightarrow |f(x) - f(y)| \leq r.$$

D'où la conclusion car ceci entraîne que la boule $\{\, y \in \mathbb{R}^n : |x - y| \leq \eta\}$ est incluse dans $f^{-1}(\Omega)$.

(b) $\Rightarrow$ (a). Pour tout $x \in \mathbb{R}^n$ et tout $\varepsilon > 0$, $\Omega = \{\, c \in \mathbb{C} : |c - f(x)| < \varepsilon\}$ est un ouvert de $\mathbb{C}$. Son image inverse étant ouverte et contenant x, il existe $\eta > 0$ tel que la boule $\{\, y \in \mathbb{R}^n : |x - y| \leq \eta\}$ soit incluse dans $f^{-1}(\Omega)$. D'où la conclusion car ceci entraîne

$$(y \in \mathbb{R}^n,\ |x - y| \leq \eta) \Rightarrow |f(x) - f(y)| \leq r.$$

(b) $\Leftrightarrow$ (c) résulte aussitôt de la formule

$$f^{-1}(A) = \mathbb{R}^n \setminus f^{-1}(\mathbb{C} \setminus A), \quad \forall A \subset \mathbb{C}.\blacksquare$$

4.2 Propriétés

4.2.1 Théorèmes de génération

Les résultats obtenus lors de l'étude de la limite d'une fonction en un point et la définition de la continuité d'une fonction permettent d'établir aussitôt les propriétés suivantes.

Théorème 4.2.1.1 (opérations algébriques) *Soit A une partie non vide de $\mathbb{R}^n$.*

a) *Toute combinaison linéaire de fonctions continues sur A est une fonction continue sur A.*

b) *Tout produit fini de fonctions continues sur A est une fonction continue sur A.*

c) *Tout quotient de deux fonctions continues sur A est une fonction continue sur A pour autant que le dénominateur ne s'annule en aucun point de A.*∎

Exercice. *Etablir que si F_0 et F_1 sont deux fermés non vides et disjoints de $\mathbb{R}^n$, alors il existe une fonction f continue sur $\mathbb{R}^n$, à valeurs dans $[0,1]$ et telle que $f(F_0) = \{0\}$ et $f(F_1) = \{1\}$.* On dit que $\mathbb{R}^n$ est un *espace normal.*

Suggestion. Les fonctions $\mathrm{d}(\cdot, F_0)$ et $\mathrm{d}(\cdot, F_1)$ sont réelles, positives et continues sur $\mathbb{R}^n$. De plus, comme F_0 et F_1 sont disjoints, on a évidemment $\mathrm{d}(x, F_0) + \mathrm{d}(x, F_1) > 0$ pour tout $x \in \mathbb{R}^n$. Cela étant, la fonction

$$f : \mathbb{R}^n \to \mathbb{R} \quad x \mapsto \frac{\mathrm{d}(x, F_0)}{\mathrm{d}(x, F_0) + \mathrm{d}(x, F_1)}$$

convient.

Théorème 4.2.1.2 (fonctions composées) *Si J appartient à $\mathbb{N}_0$, si $f_1, \ldots, f_J$ sont des fonctions réelles et continues sur la partie non vide A de $\mathbb{R}^n$ et si f est une fonction continue sur une partie de $\mathbb{R}^J$ contenant l'ensemble de variation des fonctions $f_1, \ldots, f_J$, alors la fonction composée $f(f_1, \ldots, f_J)$ est continue sur A.*∎

Le cas particulier où J est égal à 1 donne le résultat relatif aux fonctions de fonction.

Théorème 4.2.1.3 (fonctions de fonction) *Si f est une fonction réelle et continue sur la partie non vide A de $\mathbb{R}^n$ et si g est une fonction continue sur une partie de $\mathbb{R}$ contenant $\{ f(x) : x \in A\}$, alors la fonction de fonction $g(f)$ est continue sur A.*∎

Proposition 4.2.1.4 (fonctions associées) *Soit A une partie non vide de $\mathbb{R}^n$.*

a) *Si f est une fonction continue sur A, alors les fonctions $\overline{f}$, $\Re f$, $\Im f$ et $|f|$ sont continues sur A.*

b) *Si f est une fonction réelle et continue sur A, alors les fonctions f_+ et f_- sont continues sur A.*

c) *Si J appartient à $\mathbb{N}_0$ et si les fonctions $f_1, \ldots, f_J$ sont réelles et continues sur A, alors les fonctions $\sup\{f_1, \ldots, f_J\}$ et $\inf\{f_1, \ldots, f_J\}$ sont continues sur A.*∎

Proposition 4.2.1.5 *Si f est une fonction continue sur la partie A de $\mathbb{R}^n$ et si A' est une partie non vide de A, alors la restriction $f|_{A'}$ de f à A' est une fonction continue sur A'.*∎

4.2.2 Théorème des valeurs intermédiaires

On pourrait justifier l'introduction de la notion de la continuité d'une fonction par le résultat suivant et ses généralisations, tant ses conséquences aussi bien pratiques que théoriques sont importantes.

Théorème 4.2.2.1 (valeurs intermédiaires) *Si la fonction f est réelle et continue sur l'intervalle ouvert (borné ou non) $]\alpha, \beta[$ de $\mathbb{R}$ et si les limites*

$$A = \lim_{x \to \alpha^+} f(x) \quad \textit{et} \quad B = \lim_{x \to \beta^-} f(x)$$

(où $\alpha^+ = -\infty$ si $\alpha = -\infty$ et $\beta^- = +\infty$ si $\beta = +\infty$) existent (A et B pouvant être égaux à $+\infty$ ou à $-\infty$) et diffèrent, alors, pour tout nombre réel R compris strictement entre A et B, il existe $r \in]\alpha, \beta[$ tel que $f(r) = R$.

Preuve. Nous allons établir ce résultat dans le cas $A < R < B$; le cas $A > R > B$ s'établit de même ou en appliquant l'autre cas à la fonction $-f$.

Considérons l'ensemble $E = \{ x \in]\alpha, \beta[: f(x) \leq R\}$.

D'une part, E n'est pas vide. De fait, s'il était vide, on aurait $f(x) > R$ pour tout $x \in]\alpha, \beta[$ et dès lors, on devrait avoir $A \geq R$.

D'autre part, E est majoré et sa borne supérieure x_0 est telle que $x_0 < \beta$. De fait, si E n'est pas majoré par un nombre strictement inférieur à β, il existe une suite $x_m \in E$ qui converge vers β: dès lors, on devrait avoir $B \leq R$ car la suite $f(x_m)$ converge vers B.

Cela étant, pour conclure, il suffit de prouver que $f(x_0) = R$. D'une part, il existe une suite $x_m \in E$ qui converge vers x_0 car x_0 est la borne supérieure de E; on en déduit aussitôt l'inégalité $f(x_0) \leq R$. D'autre part, il existe évidemment une suite $y_m \in]\alpha, \beta[$ qui décroît strictement vers x_0 car on a $x_0 < \beta$; on en déduit aussitôt l'inégalité $f(x_0) \geq R$.

D'où la conclusion.∎

Exercice. Etablir que toute fonction réelle, continue et injective sur un intervalle I de $\mathbb{R}$ est strictement monotone sur I.

Suggestion. Il suffit bien sûr d'établir que, pour tous a, $b \in I$ tel que $a < b$, une telle fonction f est strictement monotone sur $[a, b]$. Or l'injectivité de f donne $f(a) < f(b)$ ou $f(b) < f(a)$. Supposons avoir $f(a) < f(b)$; l'autre cas se traîte de même. Le théorème des valeurs intermédiaires donne alors $f(a) < f(c) < f(b)$ pour tout $c \in]a, b[$, ce qui suffit.

Exercice. *Si $f \in \mathrm{C}_0([0,1])$ est réel et tel que $f(0) = f(1)$, établir que, pour tout $m \in \mathbb{N}_0$, il existe x_m, $y_m \in [0,1]$ tels que $y_m - x_m = 1/m$ et $f(x_m) = f(y_m)$.*

Suggestion. Posons $I = [0, 1 - 1/m]$ et introduisons la fonction

$$g \colon I \to \mathbb{R} \quad x \mapsto f(x + 1/m) - f(x).$$

Il suffit de prouver qu'il existe $x \in I$ tel que $g(x) = 0$. Comme g est réel et continu sur I, si ce n'est pas le cas, le théorème des valeurs intermédiaires affirme que g garde un signe constant sur I. Cela étant, si, par exemple, on a $g(x) > 0$ pour tout $x \in I$, il vient

$$f(1) - f(0) = g(0) + g(1/m) + \cdots + g((m-1)/m) > 0,$$

d'où une contradiction.

Remarque. $* \to$ Cependant, *pour tout $r \in \,]0, 1[\, \setminus \{\, 1/m : m \in \mathbb{N}_0\}$, il existe une fonction réelle $f \in \mathrm{C}_0([0,1])$ telle que $f(0) = f(1)$ mais pour laquelle il n'existe pas de points x, $y \in [0,1]$ tels que $y - x = r$ et $f(y) = f(x)$.* Remarquons que la fonction réelle

$$g \colon \mathbb{R} \to \mathbb{R} \quad x \mapsto \frac{\sin^2(\pi x/r)}{\sin^2(\pi/r)}$$

est continue sur $\mathbb{R}$, périodique de période r et telle que $g(0) = 0$ et $g(1) \neq 0$. Cela étant, la fonction $f(x) = g(x) - g(1)x$ convient: s'il existe des points x, $y \in [0,1]$ tels que $y - x = r$ et $f(y) = f(x)$, il vient $y = x + r$ et

$$g(x) - g(1)x = g(x + r) - g(1)x - g(1)r.$$

D'où la contradiction $r = 0$. (Cet exemple est dû à P. R. Halmos.) $\leftarrow *$

Exercice. *Etablir que toute fonction réelle $f \in \mathrm{C}_0([a,b])$ à valeurs dans $[a,b]$ a un point fixe* — c'est-à-dire qu'il existe $x \in [a,b]$ tel que $f(x) = x$. (Théorème de Brouwer, cas $n = 1$.)

Suggestion. Si f n'a pas de point fixe, la fonction g définie sur $[a,b]$ par $g(x) = f(x) - x$ est continue, à valeurs réelles, telle que $g(a) > 0$ et n'a pas de zéro. Vu le théorème des valeurs intermédiaires, g est donc à valeurs strictement positives et on doit avoir $f(b) - b > 0$, ce qui est impossible.$\square$

4.2.3 Fonctions continues sur un compact

Les fonctions continues sur un compact jouissent de propriétés fort intéressantes.

Lemme 4.2.3.1 *Si f est une fonction continue sur le compact non vide K de $\mathbb{R}^n$, alors, de toute suite c_m de $f(K)$, on peut extraire une sous-suite convergente vers un élément de $f(K)$.*

Preuve. Pour tout $m \in \mathbb{N}_0$, il existe un point $x_m \in K$ tel que $f(x_m) = c_m$. De la suite x_m du compact K, on peut alors extraire une sous-suite convergeant vers un point de K, soit $x_{k(m)} \to x$. Cela étant, il suffit de remarquer que $c_{k(m)} = f(x_{k(m)})$ est une sous-suite de la suite c_m, qui converge vers $f(x)$.∎

Théorème 4.2.3.2 (image continue d'un compact) *Si une fonction f est continue sur le compact non vide K de $\mathbb{R}^n$, alors son image $f(K)$ est compacte dans $\mathbb{C}$.*

Preuve. Vu le lemme, il est clair que $f(K)$ est fermé car il doit contenir la limite de toutes ses suites convergentes. Vu le lemme, $f(K)$ est borné: sinon $f(K)$ contiendrait une suite c_m telle que $|c_m| \geq m$ pour tout $m \in \mathbb{N}_0$, dont on ne pourrait pas extraire une sous-suite convergeant vers un élément de $f(K) \subset \mathbb{C}$.∎

Exercice. *Si f est une fonction réelle, injective et continue sur l'ouvert* (resp. *le compact*) *A de $\mathbb{R}$, alors la fonction inverse f^{-1} est continue sur $f(A)$.*

Suggestion. Si A est ouvert et si r appartient à $f(A)$, il existe un point x de A tel que $f(x) = r$ et $\varepsilon > 0$ tel que le compact $[x - \varepsilon, x + \varepsilon]$ soit inclus dans A. On est ainsi ramené au cas où A est compact. Si A est compact, il suffit de noter que tout fermé F de A est compact et tel que $(f^{-1})^{-1}(F) = f(F)$ est compact.

Théorème 4.2.3.3 (limites atteintes) a) *Si la fonction f est continue sur le compact non vide K de $\mathbb{R}^n$ et s'il existe une suite x_m de K telle que la suite $f(x_m)$ converge vers $c \in \mathbb{C}$, alors il existe $x_0 \in K$ tel que $f(x_0) = c$.*

b) *Si la fonction f est réelle et continue sur le compact non vide K de $\mathbb{R}^n$, alors il existe x_0, $y_0 \in K$ tels que*

$$f(x_0) = \inf \{ f(x) : x \in K \} \quad \textit{et} \quad f(y_0) = \sup \{ f(x) : x \in K \} .$$

Preuve. C'est une conséquence directe du lemme.∎

Théorème 4.2.3.4 (continuité uniforme) *Toute fonction continue sur un compact est uniformément continue sur ce compact.*

Preuve. Soit f une fonction continue sur le compact K de $\mathbb{R}^n$. Si f n'est pas uniformément continu sur K, il existe $\varepsilon > 0$ tel que, pour tout $m \in \mathbb{N}_0$, il existe x_m, $y_m \in K$ tels que

$$|x_m - y_m| \leq 1/m \quad \text{et} \quad |f(x_m) - f(y_m)| \geq \varepsilon.$$

Cela étant, de la suite x_m du compact K, on peut extraire une sous-suite $x_{k(m)}$ qui converge dans K; soit x_0 sa limite. Comme on a

$$\left|y_{k(m)} - x_0\right| \leq \left|y_{k(m)} - x_{k(m)}\right| + \left|x_{k(m)} - x_0\right|, \quad \forall m \in \mathbb{N}_0,$$

la sous-suite $y_{k(m)}$ converge également vers x_0. Cela étant, il vient

$$f(x_{k(m)}) - f(y_{k(m)}) = \big(f(x_{k(m)}) - f(x_0)\big) - \big(f(y_{k(m)})) - f(x_0)\big) \to 0,$$

ce qui est contradictoire.∎

Chapitre 5

Dérivabilité et différentiabilité

5.1 Dérivabilité

5.1.1 Cas $n = 1$

Rappelons qu'une fonction f définie sur un intervalle ouvert I de $\mathbb{R}$ est *dérivable en* $x \in I$ si la limite

$$[\mathrm{D}f]_x = \lim_{h \to 0,\, h \in \mathbb{R}_0} \frac{f(x+h) - f(x)}{h}$$

existe et est finie, auquel cas cette limite est appelée *dérivée de* f *en* x et est aussi notée $[\mathrm{D}f](x)$ ou $\mathrm{D}f(x)$. Cette fonction est *dérivable sur* I si elle est dérivable en tout point x de I; $\mathrm{D}f$ apparaît alors comme une fonction définie sur I, appelée *dérivée de* f *sur* I.

Une fonction f définie sur un intervalle fermé $[a, b]$ de $\mathbb{R}$ est *dérivable en* a^+ (resp. *en* b^-) si la limite

$$\lim_{h \to 0,\, h > 0} \frac{f(a+h) - f(a)}{h} \quad [\text{resp. } \lim_{h \to 0,\, h > 0} \frac{f(b-h) - f(b)}{-h}]$$

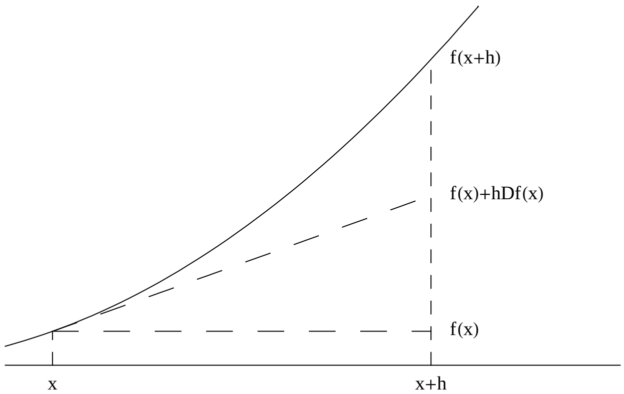
existe et est finie, auquel cas cette limite est appelée *dérivée de* f *en* a^+ (resp. *en* b^-) et est notée $[\mathrm{D}f]_{a^+}$ (resp. $[\mathrm{D}f]_{b^-}$) mais on trouve aussi la notation $[\mathrm{D}f]_a$ (resp.$[\mathrm{D}f]_b$) si aucune ambiguïté ne peut en résulter. Cette fonction est *dérivable sur* $[a, b]$ si elle est dérivable sur $]a, b[$, en a^+ et en b^-; dans ce cas, $\mathrm{D}f$ est une fonction sur $[a, b]$, appelée *dérivée de* f *sur* $[a, b]$.

Remarques. a) Soit f une fonction définie sur un intervalle I de $\mathbb{R}$. La continuité de f en $x \in I$ exprime que le reste $R(h) = f(x+h) - f(x)$ tend vers 0 si h tend vers 0 et vérifie $x + h \in I$. La dérivabilité de f en x s'interprète comme suit: le reste

$$r(h) = f(x+h) - f(x) - h \cdot [\mathrm{D}f]_x$$

non seulement tend vers 0 si $h \to 0$ avec $x + h \in I$, mais encore on a $r(h)/h \to 0$. On a donc une meilleure approximation de $f(x + h)$ à partir de $f(x)$ et de $[\mathrm{D}f]_x$ pour une fonction dérivable en x, qu'à partir de $f(x)$ pour une fonction continue en x.

b) Le nombre $(f(x+h)-f(x))/h$ est le coefficient angulaire de la droite déterminée par les points $(x, f(x))$ et $(x + h, f(x + h))$ dans le cas d'une fonction réelle. L'interprétation graphique est la suivante.



5.1.2 Fonctions dérivables

Définitions. Fixons $k \in \{1, \ldots, n\}$. Une fonction f définie sur l'ouvert Ω de $\mathbb{R}^n$ est *dérivable par rapport à x_k en $x \in \Omega$* si la limite

$$[\mathrm{D}_k f]_x = \lim_{h\to 0,\, h\in\mathbb{R}_0} \frac{f(x + he_k) - f(x)}{h}$$

existe et est finie, auquel cas cette limite est appelée *dérivée de f par rapport à x_k en x* et est aussi notée

$$[\mathrm{D}_k f](x), \quad [\mathrm{D}_{x_k} f]_x, \quad [\mathrm{D}_{x_k} f](x), \quad \left[\frac{\partial}{\partial x_k} f\right]_x, \quad \left[\frac{\partial f}{\partial x_k}\right]_x, \cdots$$

Cette fonction est *dérivable par rapport à x_k sur Ω* si elle est dérivable par rapport à x_k en tout $x \in \Omega$; $\mathrm{D}_k f$ apparaît alors comme une fonction sur Ω, appelée *dérivée de f par rapport à x_k*.

Une fonction f définie sur un ouvert Ω de $\mathbb{R}^n$ est *dérivable sur Ω* si, pour tout $k \in \{1, \ldots, n\}$, elle est dérivable par rapport à x_k sur Ω. Elle est *continûment dérivable sur Ω* si elle est dérivable sur Ω et si, pour tout $k \in \{1, \ldots, n\}$, $\mathrm{D}_k f$ est une fonction continue sur Ω.

Remarque. Soient f une fonction définie sur l'ouvert Ω de $\mathbb{R}^n$, x un point de Ω et k un élément de $\{1, \ldots, n\}$. On vérifie de suite que

$$\Omega_{x,k} = \{\, y \in \mathbb{R} : (x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_n) \in \Omega\}$$

est un ouvert de $\mathbb{R}$ qui contient x_k. Si on pose

$$g(y) = f(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_n)$$

pour tout $y \in \Omega_{x,k}$, il est clair que g est une fonction définie sur $\Omega_{x,k}$, que f est dérivable par rapport à x_k en x si et seulement si g est dérivable en x_k et qu'on a alors $[\mathrm{D}_k f]_x = [\mathrm{D}g]_{x_k}$.

La liaison entre la continuité et la dérivabilité est assez complexe.

a) *La continuité n'entraîne pas la dérivabilité.* Ainsi $|\cdot|$ est une fonction continue sur $\mathbb{R}^n$ qui n'est pas dérivable à l'origine. $* \rightarrow$ On peut même construire une fonction continue sur $\mathbb{R}$ qui n'est dérivable en aucun point de $\mathbb{R}$. $\leftarrow *$

b) *Si la fonction f est dérivable sur l'ouvert Ω de $\mathbb{R}^n$, alors, pour tout $x \in \Omega$ et tout $k \in \{1, \ldots, n\}$, on a*

$$\lim_{h \to 0,\, h \in \mathbb{R}_0} f(x + he_k) = f(x).$$

De fait, il existe $r > 0$ tel que la boule de centre x et de rayon r soit incluse dans Ω. Cela étant, pour tout $h \in \,]-r, 0[\cup]0, r[$, on a

$$f(x + he_k) - f(x) = h \cdot \frac{f(x + he_k) - f(x)}{h}$$

et, si $h \neq 0$ converge vers 0, cette expression converge vers $0 \cdot [\mathrm{D}_k f]_x = 0$, ce qui suffit.

En particulier, nous avons établi le résultat suivant.

Théorème 5.1.2.1 *Toute fonction dérivable sur un ouvert de $\mathbb{R}$ est continue sur cet ouvert.*∎

c) *Cependant si n est strictement supérieur à 1, la dérivabilité d'une fonction sur un ouvert de $\mathbb{R}^n$ n'entraîne pas sa continuité sur cet ouvert.* Ainsi la fonction

$$f \colon \mathbb{R}^2 \to \mathbb{R} \quad (x, y) \mapsto \begin{cases} 0 & \text{si } (x, y) = (0, 0) \\ xy/(x^2 + y^2) & \text{si } (x, y) \neq (0, 0) \end{cases}$$

est dérivable sur $\mathbb{R}^2$ (à vérifier directement) mais n'est pas continue à l'origine (cf. p. 98).

d) $* \rightarrow$ *La dérivabilité d'une fonction sur un ouvert de $\mathbb{R}^n$ n'entraîne pas que cette fonction soit continûment dérivable sur cet ouvert.* Ainsi la fonction

$$f \colon \mathbb{R} \to \mathbb{R} \quad x \mapsto \begin{cases} 0 & \text{si } x = 0 \\ x^2 \cdot \sin(1/x) & \text{si } x \neq 0 \end{cases}$$

est dérivable sur $\mathbb{R}$: sur $\mathbb{R} \setminus \{0\}$, on a bien sûr

$$[\mathrm{D}f]_x = 2x \cdot \sin(1/x) - \cos(1/x)$$

et, en 0, il vient

$$[\mathrm{D}f]_0 = \lim_{h \to 0,\, h \in \mathbb{R}_0} \frac{1}{h} h^2 \cdot \sin(1/h) = 0.$$

Cependant sa dérivée $\mathrm{D}f$ n'est pas continue en 0 car la suite

$$[\mathrm{D}f]_{1/(m\pi)} = \frac{2}{m\pi} \cdot \sin(m\pi) - \cos(m\pi) = (-1)^{m+1}$$

ne converge pas. $\leftarrow *$

e) $* \rightarrow$ Au paragraphe 5.1.7, nous établissons que *toute fonction continûment dérivable sur un ouvert de $\mathbb{R}^n$ est continue sur cet ouvert.* $\leftarrow *$

Exercice. Etablir qu'il n'existe pas de fonction réelle $f \in \mathrm{C}_0(\mathbb{R}^2)$ qui ne s'annule qu'en $(0,0)$ et dont la dérivée $[\mathrm{D}_1 f]_0$ existe et diffère de 0.

Suggestion. Vu le théorème des valeurs intermédiaires, une telle fonction f garde un signe constant sur $\{\,(x,0) : x > 0\}$ et sur $\{\,(x,0) : x < 0\}$. De plus, ces signes doivent différer car $[\mathrm{D}_1 f]_0$ existe et diffère de 0. Cela étant, la fonction g définie sur $[0,\pi]$ par $g \colon \theta \mapsto f(\cos(\theta), \sin(\theta))$ est continue et telle que $g(0)g(\pi) < 0$. Vu le théorème des valeurs intermédiaires à nouveau, il existe alors $\theta_0 \in \,]0,\pi[$ où g s'annule, ce qui constitue une contradiction.

5.1.3 Exemples fondamentaux

a) **Fonctions constantes.** *La fonction $\chi_{\mathbb{R}^n}$ est dérivable sur $\mathbb{R}^n$ et, pour tout $k \in \{1,\ldots,n\}$, on a $\mathrm{D}_k \chi_{\mathbb{R}^n} = 0$.*

b) **Fonctions composantes.** *Pour tout $j \in \{1,\ldots,n\}$, la fonction $[\cdot]_j$ est dérivable sur $\mathbb{R}^n$ et telle que*

$$\mathrm{D}_k[\cdot]_j = \begin{cases} 0 & \textit{si } k \neq j, \\ \chi_{\mathbb{R}^n} & \textit{si } k = j. \end{cases}$$

Il suffit de noter que, pour tout $x \in \mathbb{R}^n$ et tout $h \in \mathbb{R}_0$, on a

$$\frac{[x + he_k]_j - [x]_j}{h} = \begin{cases} 0 & \text{si } k \neq j, \\ 1 & \text{si } k = j. \end{cases}$$

Remarque. Par contre, si A est une partie non vide de $\mathbb{R}^n$, on ne peut affirmer que la fonction $d(\cdot, A)$ soit dérivable sur $\mathbb{R}^n$. Ainsi pour $A = \{0\}$, on obtient $|\cdot|$ et cette fonction n'est pas dérivable en 0 car

$$\lim_{h \to 0,\, h > 0} \frac{|0 + he_k| - |0|}{h} = 1 \quad \text{et} \quad \lim_{h \to 0,\, h < 0} \frac{|0 + he_k| - |0|}{h} = -1$$

5.1.4 Théorèmes de génération

Convention. Dans ce paragraphe, J *désigne un entier supérieur ou égal à* 1*;* f, g, f_1, ..., f_J *sont des fonctions dérivables sur un même ouvert* Ω *de* $\mathbb{R}^n$ *et* c_1, ..., c_J *sont des nombres complexes.*

Vis-à-vis des opérations algébriques, on a les résultats suivants.

Théorème 5.1.4.1 *Toute combinaison linéaire de fonctions dérivables sur un ouvert* Ω *de* $\mathbb{R}^n$ *est dérivable sur* Ω *et, pour tout* $k \in \{1, \ldots, n\}$,

$$\mathrm{D}_k \left(\sum_{j=1}^{J} c_j f_j \right) = \sum_{j=1}^{J} c_j \mathrm{D}_k f_j.$$

Preuve. Cela résulte directement de la propriété relative à la limite des combinaisons linéaires de fonctions.∎

Théorème 5.1.4.2 *Tout produit fini de fonctions dérivables sur un ouvert* Ω *de* $\mathbb{R}^n$ *est dérivable sur* Ω *et, pour tout* $k \in \{1, \ldots, n\}$,

$$\begin{aligned} \mathrm{D}_k(f_1 \cdots f_J) &= (\mathrm{D}_k f_1) \cdot f_2 \cdots f_J + \cdots + f_1 \cdots f_{j-1} \cdot (\mathrm{D}_k f_j) \cdot f_{j+1} \cdots f_J \\ &\quad + \cdots + f_1 \cdots f_{J-1} \cdot (\mathrm{D}_k f_J). \end{aligned}$$

En particulier, si les f_j *diffèrent de* 0 *en tout point de* Ω, *alors, pour tout entier* $k \in \{1, \ldots, n\}$,

$$\frac{\mathrm{D}_k(f_1 \ldots f_J)}{f_1 \ldots f_J} = \frac{\mathrm{D}_k f_1}{f_1} + \cdots + \frac{\mathrm{D}_k f_J}{f_J}.$$

Preuve. Un raisonnement par récurrence montre aisément qu'il suffit de prouver ce théorème pour $J = 2$. Or on a

$$\begin{aligned} &\frac{f_1(x + he_k) \cdot f_2(x + he_k) - f_1(x) \cdot f_2(x)}{h} \\ &= \frac{f_1(x + he_k) - f_1(x)}{h} \cdot f_2(x + he_k) + f_1(x) \cdot \frac{f_2(x + he_k) - f_2(x)}{h} \end{aligned}$$

pour tout h de module suffisamment petit et on voit aisément que le second membre de cette égalité converge vers $(\mathrm{D}_k f_1) \cdot f_2 + f_1 \cdot (\mathrm{D}_k f_2)$ si $h \to 0$, ce qui suffit.∎

Théorème 5.1.4.3 *Le quotient de deux fonctions dérivables sur un ouvert* Ω *de* $\mathbb{R}^n$ *dont le dénominateur ne s'annule en aucun point de* Ω *est une fonction dérivable sur* Ω *et telle que, pour tout* $k \in \{1, \ldots, n\}$,

$$\mathrm{D}_k \frac{f}{g} = \frac{(\mathrm{D}_k f) \cdot g - f \cdot (\mathrm{D}_k g)}{g^2}.$$

Preuve. Vu le théorème précédent, il suffit évidemment d'établir que $1/g$ est dérivable sur Ω et tel que, pour tout $k \in \{1, \ldots, n\}$,

$$\mathrm{D}_k \frac{1}{g} = -\frac{\mathrm{D}_k g}{g^2}.$$

Or, pour tout $h \neq 0$ de module suffisamment petit, il vient

$$\frac{1}{h}\left(\frac{1}{g(x+he_k)} - \frac{1}{g(x)}\right) = -\frac{g(x+he_k)-g(x)}{h} \cdot \frac{1}{g(x)} \cdot \frac{1}{g(x+he_k)};$$

la conclusion s'ensuit aussitôt.∎

Ces trois énoncés donnent lieu au résultat suivant.

Théorème 5.1.4.4 (opérations algébriques) *Toute opération algébrique effectuée sur des fonctions dérivables sur un ouvert Ω de $\mathbb{R}^n$ est une fonction dérivable sur Ω pour autant que les dénominateurs éventuels ne s'annulent en aucun point de Ω. De plus, sa dérivée par rapport à x_k est donnée par une opération algébrique où n'interviennent que les fonctions et leurs dérivées par rapport à x_k.*∎

Exercice. Si A est une matrice de dimension $J \times J$ dont les éléments sont des fonctions dérivables sur l'ouvert Ω de $\mathbb{R}^n$, établir que $\mathrm{dtm}(A)$ est dérivable sur Ω et tel que

$$\begin{aligned}
\mathrm{D}_k \mathrm{dtm}(A) &= \mathrm{dtm}(\mathrm{D}_k c_1, c_2, \ldots, c_J) + \ldots + \mathrm{dtm}(c_1, \ldots, c_{J-1}, \mathrm{D}_k c_J) \\
&= \sum_{j,l=1}^{J} \mathcal{A}_{j,l} \cdot \mathrm{D}_k a_{j,l}
\end{aligned}$$

pour tout $k \in \{1, \ldots, n\}$, en désignant par c_j la $j-$ème colonne de A, par $a_{j,l}$ l'élément (j, l) de A et par $\mathcal{A}_{j,l}$ le mineur de $a_{j,l}$.

Exercice. Si $f_1, \ldots, f_N$ sont des fonctions dérivables sur l'ouvert Ω de $\mathbb{R}^n$ et si $g_1, \ldots, g_M$ sont des fonctions dérivables sur Ω qui ne s'annulent en aucun point de Ω, établir que $(f_1 \cdots f_N)/(g_1 \cdots g_M)$ est une fonction dérivable sur Ω et que, pour tout $k \in \{1, \ldots, n\}$, on a

$$\mathrm{D}_k \frac{f_1 \cdots f_N}{g_1 \cdots g_M} = \frac{f_1 \cdots f_N}{g_1 \cdots g_M} \cdot \left(\sum_{j=1}^{N} \frac{\mathrm{D}_k f_j}{f_j} - \sum_{j=1}^{M} \frac{\mathrm{D}_k g_j}{g_j}\right).$$

Les propriétés de dérivabilité relatives aux fonctions associées à une fonction dérivable sont assez pauvres.

Théorème 5.1.4.5 (fonctions associées) *Une fonction f définie sur l'ouvert Ω de $\mathbb{R}^n$ est dérivable sur Ω si et seulement si $\overline{f}$ est dérivable sur Ω, ce qui a lieu si et seulement si $\Re f$ et $\Im f$ sont dérivables sur Ω; de plus, on a alors*

$$\mathrm{D}_k\overline{f} = \overline{\mathrm{D}_k f}, \quad \mathrm{D}_k(\Re f) = \Re(\mathrm{D}_k f) \quad et \quad \mathrm{D}_k(\Im f) = \Im(\mathrm{D}_k f).\blacksquare$$

Remarque. Insistons sur le fait que la proposition "si f est dérivable sur l'ouvert Ω, alors $|f|$ est dérivable sur Ω" est fausse comme le montre le contre-exemple de la fonction $|\cdot|$ sur $\mathbb{R}$. De là, les propositions "si f est une fonction réelle et dérivable sur l'ouvert Ω, alors f_+ et f_- sont dérivables sur Ω" et "si $f_1, \ldots, f_J$ sont des fonctions réelles et dérivables sur l'ouvert Ω, alors les fonctions $\sup\{f_1, \ldots, f_J\}$ et $\inf\{f_1, \ldots, f_J\}$ sont dérivables sur Ω" sont également fausses.

Théorème 5.1.4.6 *Si f est une fonction dérivable sur l'ouvert Ω de $\mathbb{R}^n$, alors la restriction de f à tout ouvert Ω' de $\mathbb{R}^n$ inclus dans Ω est une fonction dérivable sur Ω' et, pour tout $k \in \{1, \ldots, n\}$, $\mathrm{D}_k(f|_{\Omega'})$ est égal à la restriction de $\mathrm{D}_k f$ à Ω'.*$\blacksquare$

Dans le cadre des théorèmes de génération des fonctions dérivables, il nous reste à examiner le cas des fonctions composées et des fonctions de fonction. A cet effet, nous allons d'abord établir le théorème des accroissements finis.

5.1.5 Théorème des accroissements finis

Présentons d'abord un résultat qui sera nettement généralisé plus tard au paragraphe 7.2.2 relatif aux extrema d'une fonction.

Lemme 5.1.5.1 *Si f est une fonction réelle et dérivable sur l'intervalle $]a, b[$ de $\mathbb{R}$ et si $x_0 \in]a, b[$ est tel que $f(x_0) \geq f(x)$* (resp. $f(x_0) \leq f(x)$) *pour tout $x \in]a, b[$, alors on a $[\mathrm{D}f]_{x_0} = 0$.*

Preuve. De fait, comme l'expression $(f(x_0 + h) - f(x_0))/h$ est définie pour tout $h \in]a - x_0, 0[\cup]0, b - x_0[$ et change de signe suivant que $h > 0$ ou $h < 0$, sa limite pour $h \to 0$ doit être 0.$\blacksquare$

Voici ensuite un résultat qui apparaîtra comme étant un cas particulier du théorème des accroissements finis dans $\mathbb{R}$.

Théorème 5.1.5.2 (Rolle) *Si la fonction f est réelle et continue sur l'intervalle compact $[a, b]$ de $\mathbb{R}$, dérivable sur $]a, b[$ et vérifie $f(a) = f(b)$, alors il existe $x_0 \in]a, b[$ tel que $[\mathrm{D}f]_{x_0} = 0$.*

Preuve. Bien sûr, si f est constant sur $[a,b]$, tout point de $]a,b[$ convient.

Si f n'est pas constant sur $[a,b]$, la borne supérieure ou la borne inférieure de la fonction réelle et continue f sur le compact $[a,b]$ diffère de $f(a)=f(b)$ et est réalisée en un point $x_0 \in]a,b[$. D'où la conclusion par le lemme précédent.∎

Remarque. En procédant par l'absurde, on déduit directement du théorème de Rolle le résultat suivant: si f est une fonction réelle et dérivable sur un intervalle ouvert de $\mathbb{R}$ et si $\mathrm{D}f$ ne s'annule en aucun point de cet intervalle, alors, pour tout nombre réel r, l'équation $f(x)=r$ n'admet qu'une solution au plus dans cet intervalle.

Théorème 5.1.5.3 (accroissements finis dans $\mathbb{R}$) *Si la fonction f est réelle et continue sur l'intervalle $[a,b]$ de $\mathbb{R}$ et est dérivable sur $]a,b[$, alors il existe $h \in]0,b-a[$ tel que*

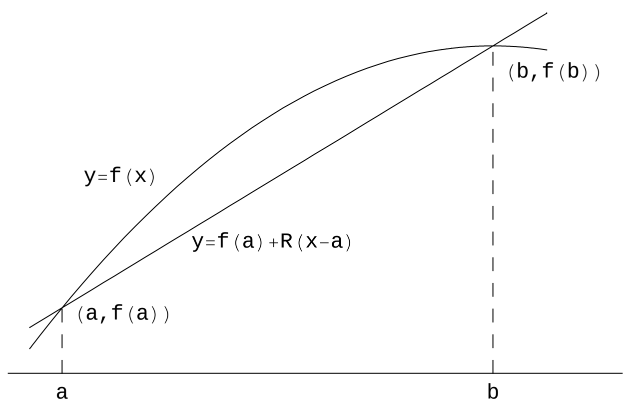
$$f(b) = f(a) + (b-a)[\mathrm{D}f]_{a+h}.$$

Preuve. Considérons la fonction

$$F\colon\ [a,b] \to \mathbb{R} \quad x \mapsto f(x) - f(a) - R\cdot(x-a)$$

où R est le coefficient angulaire de la droite déterminée par les points $(a,f(a))$ et $(b,f(b))$, c'est-à-dire que $R=(f(b)-f(a))/(b-a)$. On vérifie de suite que cette fonction F est réelle et continue sur $[a,b]$, dérivable sur $]a,b[$ et de dérivée égale à $\mathrm{D}f-R$. Comme on a aussi $F(a)=0=F(b)$, le théorème de Rolle affirme l'existence d'un point $x_0 \in]a,b[$ tel que $[\mathrm{D}F]_{x_0}=0$, c'est-à-dire qu'il existe $h \in]0,b-a[$ tel que $x_0=a+h$ donc tel que $[\mathrm{D}f]_{a+h}-R=0$. D'où la conclusion.∎

Remarque. La droite d'équation $y=f(a)+R(x-a)$ est en fait la droite déterminée par les points $(a,f(a))$ et $(b,f(b))$. De la sorte, l'interprétation géométrique de F est claire. De plus, le théorème des accroissements finis revient à dire qu'il existe un point de $]a,b[$ où $\mathrm{D}f$ est égal à R.



Corollaire 5.1.5.4 *Si la fonction f est réelle et dérivable sur l'intervalle I de $\mathbb{R}$, alors, pour tout $x \in I$ et tout accroissement $h \in \mathbb{R} \setminus \{0\}$ tel que $x + h \in I$, il existe un accroissement auxiliaire h' compris strictement entre 0 et h tel que*

$$f(x + h) = f(x) + h[\mathrm{D}f]_{x+h'}.$$

Preuve. Comme f est dérivable sur I, f est continu sur I.

Cela étant, si on a $h > 0$, c'est trivial en considérant les points $a = x$ et $b = x+h$. Si, par contre, on a $h < 0$, on considère les points $a = x + h$ et $b = x$; il existe alors $h'' \in \,]0, -h[$ tel que

$$f(x) = f(x + h) - h[\mathrm{D}f]_{x+h+h''},$$

c'est-à-dire $h' = h + h'' \in \,]h, 0[$ tel que $f(x + h) = f(x) + h[\mathrm{D}f]_{x+h'}$.∎

Exercice. Etablir que l'hypothèse "f réel" est indispensable dans l'énoncé du théorème des accroissements finis.

Suggestion. Un contre-exemple est donné par la fonction $f \colon x \mapsto 1/(x - i)$ qui est dérivable sur $\mathbb{R}$ et telle que

$$f(1) - f(0) = \frac{1}{1 + i} \quad \text{et} \quad [\mathrm{D}f]_x = -\frac{1}{(x - i)^2}$$

alors que les solutions de l'équation $x^2 - 2ix - 1 = -1 - i$ ne sont pas des nombres réels.

Exercice. Si la fonction f est dérivable sur l'intervalle $]0, r[$ et si sa dérivée est bornée sur cet intervalle, établir que $\lim_{x\to 0^+} f(x)$ existe et est fini.

Suggestion. Remarquer qu'il suffit d'établir cette propriété dans le cas où f est une fonction réelle. Poser $C = \sup \{\, |\mathrm{D}f(x)| : x \in \,]0, r[\}$ puis remarquer que, pour toute suite $x_m \in \,]0, r[$ convergeant vers 0, la suite $f(x_m)$ est de Cauchy car, pour tous r, $s \in \mathbb{N}_0$, le théorème des accroissements finis fournit un point $x_{r,s}$ compris strictement entre x_r et x_s tel que

$$|f(x_r) - f(x_s)| = \left|(x_r - x_s)[\mathrm{D}f]_{x_{r,s}}\right| \leq C\,|x_r - x_s|\,.$$

Exercice. Soit x_0 un point de l'intervalle I de $\mathbb{R}$. Si f est une fonction continue sur I et dérivable sur $I \setminus \{x_0\}$ et si $\lim_{x\to x_0} \mathrm{D}f$ existe, est fini et vaut a, établir que f est dérivable en x_0 et que $[\mathrm{D}f]_{x_0} = a$.

Suggestion. Se ramener d'abord au cas d'une fonction réelle puis appliquer le théorème des accroissements finis dans $\mathbb{R}$ pour justifier les égalités

$$\lim_{h\to 0,\, h\in\mathbb{R}_0} \frac{1}{h}(f(x_0 + h) - f(x_0)) = \lim_{h\to 0,\, h\in\mathbb{R}_0} \frac{h}{h}[\mathrm{D}f]_{x_0+h'} = a.$$

Théorème 5.1.5.5 (accroissements finis dans $\mathbb{R}^n$) *Si une fonction f est réelle et dérivable sur un ouvert Ω de $\mathbb{R}^n$, alors, pour tout $x \in \Omega$ et tout accroissement $h \in \mathbb{R}^n$ tel que la boule $\{ y : |x-y| \leq |h| \}$ soit incluse dans Ω, il existe des accroissements auxiliaires $h^{(1)}, \ldots, h^{(n)} \in \mathbb{R}^n$ tels que $\left|h^{(1)}\right|, \ldots, \left|h^{(n)}\right| \leq |h|$ et*

$$f(x+h) = f(x) + \sum_{j=1}^{n} h_j [\mathrm{D}_j f]_{x+h^{(j)}}.$$

Preuve. Comme, pour tout $j \in \{1, \ldots, n\}$,

$$(x_1, \ldots, x_{j-1}, x_j + h_j, \ldots, x_n + h_n)$$

est un point de Ω, nous pouvons écrire

$$\begin{aligned} f(x+h) - f(x) &= f(x_1+h_1, \ldots, x_n+h_n) - f(x_1, \ldots, x_n) \\ &= f(x_1+h_1, \ldots, x_n+h_n) - f(x_1, x_2+h_2, \ldots, x_n+h_n) \\ &\quad + \ldots \\ &\quad + f(x_1, \ldots, x_{j-1}, x_j+h_j, \ldots, x_n+h_n) \\ &\qquad - f(x_1, \ldots, x_j, x_{j+1}+h_{j+1}, \ldots, x_n+h_n) \\ &\quad + \ldots \\ &\quad + f(x_1, \ldots, x_{n-1}, x_n+h_n) - f(x_1, \ldots, x_n). \end{aligned}$$

Considérons le j-ème terme du second membre. Si h_j est égal à 0, ce terme est égal à 0 et peut s'écrire

$$h_j [\mathrm{D}_j f]_{x+h^{(j)}}$$

avec $h^{(j)} = (0, \ldots, 0, h_j, h_{j+1}, \ldots, h_n)$. Si h_j diffère de 0, on vérifie de suite que la fonction

$$F_j(y) = f(x_1, \ldots, x_{j-1}, y, x_{j+1}+h_{j+1}, \ldots, x_n+h_n)$$

est définie, réelle et continue sur l'intervalle compact $[x_j - |h_j|, x_j + |h_j|]$ qui contient les points x_j et $x_j + h_j$, et que cette fonction F_j est dérivable sur l'intervalle $]x_j - |h|, x_j + |h|[$. Vu le théorème des accroissements finis dans $\mathbb{R}$, il existe alors h'_j strictement compris entre 0 et h_j tel que

$$F_j(x_j + h_j) - F_j(x_j) = h_j [\mathrm{D}F_j]_{x_j + h'_j}$$

donc tel que ce $j-$ème terme du second membre soit égal à

$$h_j [\mathrm{D}_j f]_{x+h^{(j)}} \quad \text{avec} \quad h^{(j)} = (0, \ldots, 0, h'_j, h_{j+1}, \ldots, h_n).$$

D'où la conclusion.∎

5.1.6 Fonctions composées, fonctions de fonction

Théorème 5.1.6.1 (fonctions composées) *Si J appartient à $\mathbb{N}_0$, si $f_1, \ldots,$ f_J sont des fonctions réelles et dérivables sur l'ouvert Ω de $\mathbb{R}^n$ et si f est une fonction continûment dérivable sur un ouvert ω de $\mathbb{R}^J$ contenant $\{\,(f_1(x), \ldots, f_J(x)) : x \in \Omega\}$, alors $f(f_1, \ldots, f_J)$ est une fonction dérivable sur Ω et, pour tout $k \in \{1, \ldots, n\}$,*

$$[\mathrm{D}_k f(f_1, \ldots, f_J)]_x = \sum_{j=1}^{J} [\mathrm{D}_j f]_{(f_1(x), \ldots, f_J(x))} \cdot [\mathrm{D}_k f_j]_x.$$

Preuve. Etablissons tout d'abord ce résultat sous l'hypothèse supplémentaire que f soit réel sur ω. Soit x un point de Ω. Le point $X = (f_1(x), \ldots, f_J(x))$ appartient à ω; il est donc le centre d'une boule b incluse dans ω: il existe $R > 0$ tel que $b = \{\, Y : |X - Y| \leq R\} \subset \omega$. Pour tout $j \in \{1, \ldots, n\}$, nous savons que $f_j(x + he_k) \to f_j(x)$ si $h \in \mathbb{R}$ tend vers 0. Il existe donc $r > 0$ tel que, pour tout $h \in \mathbb{R} \setminus \{0\}$ vérifiant $|h| \leq r$, on a $|f_j(x + he_k) - f_j(x)| \leq R/J$. Dès lors, pour tout $h \in \mathbb{R} \setminus \{0\}$ tel que $|h| \leq r$, le point

$$X + H = (f_1(x + he_k), \ldots, f_J(x + he_k))$$

vérifie $|H| \leq R$. Dès lors, le théorème des accroissements finis dans $\mathbb{R}^J$ procure des accroissements auxiliaires $H^{(1)}, \ldots, H^{(J)} \in \mathbb{R}^J$, de module majoré par $|H|$ et tels que

$$f(X + H) - f(X) = \sum_{j=1}^{J} H_j [\mathrm{D}_j f]_{X + H^{(j)}}.$$

Pour $h \to 0$, on a $H^{(j)} \to 0$ pour tout $j \leq J$. Il s'ensuit que

$$\begin{aligned} &\frac{f(f_1(x + he_k), \ldots, f_J(x + he_k)) - f(f_1(x), \ldots, f_J(x))}{h} \\ &= \sum_{j=1}^{J} \frac{f_j(x + he_k) - f_j(x)}{h} \cdot [\mathrm{D}_j f]_{(f_1(x), \ldots, f_J(x)) + H^{(j)}} \end{aligned}$$

converge vers

$$\sum_{j=1}^{J} [\mathrm{D}_k f_j]_x \cdot [\mathrm{D}_j f]_{(f_1(x), \ldots, f_J(x))}$$

si $h \to 0$, vu la continuité des dérivées de f sur ω.

Déduisons-en la propriété dans le cas général où f est une fonction à valeurs complexes. Les fonctions $\Re f$ et $\Im f$ sont alors des fonctions réelles et continûment dérivables sur ω et on peut leur appliquer la première partie de cette preuve. Il vient ainsi successivement

$$f(f_1, \ldots, f_J) = (\Re f)(f_1, \ldots, f_J) + i(\Im f)(f_1, \ldots, f_J)$$

et

$$\begin{aligned}
&[\mathrm{D}_k(\Re f)(f_1, \ldots, f_J)]_x + i[\mathrm{D}_k(\Im f)(f_1, \ldots, f_J)]_x \\
&= \sum_{j=1}^{J} [\mathrm{D}_j(\Re f)]_{(f_1(x),\ldots,f_J(x))} \cdot [\mathrm{D}_k f_j]_x + i \sum_{j=1}^{J} [\mathrm{D}_j(\Im f)]_{(f_1(x),\ldots,f_J(x))} \cdot [\mathrm{D}_k f_j]_x \\
&= \sum_{j=1}^{J} [\mathrm{D}_j(\Re f + i\Im f)]_{(f_1(x),\ldots,f_J(x))} \cdot [\mathrm{D}_k f_j]_x.
\end{aligned}$$

D'où la conclusion.∎

En considérant la valeur $J = 1$, le résultat précédent s'applique bien sûr aux fonctions de fonction mais il ne s'appelle pas théorème de dérivation des fonctions de fonction car il est possible d'affaiblir sensiblement l'hypothèse portant sur f.

Théorème 5.1.6.2 (fonctions de fonction) *Si la fonction f est réelle et dérivable sur l'ouvert Ω de $\mathbb{R}^n$ et si g est une fonction dérivable sur un ouvert ω de $\mathbb{R}$ contenant $\{ f(x) : x \in \Omega\}$, alors la fonction de fonction $g(f)$ est dérivable sur Ω et*

$$[\mathrm{D}_k g(f)]_x = [\mathrm{D}g]_{f(x)} \cdot [\mathrm{D}_k f]_x.$$

Preuve. Si x est un point de Ω, il existe $r > 0$ tel que $\{ y : |x - y| \leq r\}$ soit inclus dans Ω. Cela étant, pour tout $h \in \mathbb{R} \setminus \{0\}$ vérifiant $|h| \leq r$, on peut écrire

$$\frac{g(f(x + he_k)) - g(f(x))}{h} = \Phi(h) \cdot \frac{f(x + he_k) - f(x)}{h}$$

si on définit $\Phi(h)$ par

$$\Phi(h) = \begin{cases} \dfrac{g(f(x + he_k)) - g(f(x))}{f(x + he_k) - f(x)} & \text{si } f(x + he_k) \neq f(x) \\ [\mathrm{D}g]_{f(x)} & \text{sinon.} \end{cases}$$

Remarquons alors que, pour $h \to 0$, la fonction $\Phi(h)$ converge vers $[\mathrm{D}g]_{f(x)}$. De fait, dans l'ensemble $\Omega_1 = \{ h \neq 0 : f(x + he_k) \neq f(x)\}$, cela est vrai si 0 appartient à $(\Omega_1)^-$ et, dans $\Omega_2 = \{ h \neq 0 : f(x + he_k) = f(x)\}$, cela est encore vrai pour autant que $0 \in (\Omega_2)^-$. La conclusion s'ensuit aisément.∎

Voici deux applications intéressantes du théorème de dérivation des fonctions composées.

Application. Le théorème de dérivation des fonctions composées permet de calculer les dérivées d'une fonction réelle f sur un ouvert Ω de $\mathbb{R}^n$, sans la connaître explicitement, si on sait qu'elle vérifie une identité du type $g(x, f(x)) = 0$ sur Ω.

En effet, si on sait qu'on peut appliquer le théorème de dérivation des fonctions composées à $g(x, f(x))$ sur Ω, il vient

$$0 = \mathrm{D}_k g(x, f(x)) = \sum_{j=1}^{n+1} [\mathrm{D}_j g]_{(x,f(x))} \cdot [\mathrm{D}_k f_j]_x$$

en posant $f_j(x) = x_j$ pour $j = 1, \ldots, n$ et $f_{n+1} = f$. On a donc

$$[\mathrm{D}_k f]_x = -\frac{[\mathrm{D}_k g]_{(x,f(x))}}{[\mathrm{D}_{n+1} g]_{(x,f(x))}}$$

pour autant que cette division puisse se faire.

Application. Si f est une fonction réelle et dérivable sur l'ouvert Ω de $\mathbb{R}$, si g est une fonction dérivable sur un ouvert Ω' de $\mathbb{R}$ contenant $\{\, f(x) : x \in \Omega\}$ et si on a $g(f(x)) = x$ pour tout $x \in \Omega$, on obtient que $[\mathrm{D}g]_{f(x)} \cdot [\mathrm{D}f]_x = 1$ sur Ω.

Notons encore que le même procédé fonctionnne pour un système d'équations de la forme

$$\left\{ \begin{array}{rcl} g_1(f_1(x), \ldots, f_J(x)) & = & 0 \\ \ldots & & \\ g_J(f_1(x), \ldots, f_J(x)) & = & 0 \end{array} \right.$$

sur un ouvert Ω de $\mathbb{R}^n$. Dans un tel cas, on détermine simultanément $\mathrm{D}_k f_1, \ldots, \mathrm{D}_k f_J$ en dérivant chacune des équations de ce système par rapport à x_k et en résolvant le nouveau système obtenu, qui est un système de J équations linéaires en les inconnues $\mathrm{D}_k f_1, \ldots, \mathrm{D}_k f_J$.

Exercice. $* \rightarrow$ *Sachant que f est une fonction continue sur $[-1, 1]$ et dérivable sur $]-1, 1[$, de dérivée donnée par $[\mathrm{D}f]_x = (1 - x^2)^{-1/2}$ sur $]-1, 1[$, où la fonction $g\colon x \mapsto f(2x/(1 + x^2))$ est-elle définie? continue? dérivable? Que vaut sa dérivée?*

Suggestion. On sait que $\sqrt{\cdot}$ est une fonction continue sur $[0, +\infty[$ et dérivable sur $]0, +\infty[$, de dérivée donnée par $[\mathrm{D}\sqrt{\cdot}]_x = 1/(2\sqrt{x})$ sur $]0, +\infty[$.

a) Ensemble de définition. L'ensemble de définition A de g est donné par

$$A = \left\{\, x \in \mathbb{R} : -1 \leq 2x/(1 + x^2) \leq 1 \right\} = \mathbb{R}.$$

b) Le théorème de continuité des fonctions de fonction donne aussitôt la continuité de g sur $\mathbb{R}$.

c) Ensemble de dérivabilité. La fonction g est alors dérivable sur

$$A' = \left\{ x \in \mathbb{R} : -1 < 2x/(1+x^2) < 1 \right\} =]-\infty, -1[\cup]-1, 1[\cup]1, +\infty[\,.$$

d) Calcul de la dérivée de g. Il vient successivement

$$\begin{aligned} \mathrm{D}g(x) &= \left[\frac{1}{\sqrt{1-x^2}}\right]_{2x/(1+x^2)} \cdot \mathrm{D}\frac{2x}{1+x^2} = \frac{1}{\sqrt{(1-x^2)^2}} \cdot \frac{2(1-x^2)}{1+x^2} \\ &= \begin{cases} \dfrac{2}{1+x^2} & \text{sur }]-1, 1[, \\ \dfrac{-2}{1+x^2} & \text{sur }]-\infty, -1[\cup]1, +\infty[\,. \quad \leftarrow * \end{cases} \end{aligned}$$

Exercice. *Si f est une fonction réelle et dérivable sur l'intervalle ouvert $]\alpha, \beta[$ (borné ou non) de $\mathbb{R}$ et si les limites $\lim_{x\to\alpha^+} \mathrm{D}f$ et $\lim_{x\to\beta^-} \mathrm{D}f$ existent et sont égales respectivement à A et B (les cas A, $B \in \{-\infty, +\infty\}$ n'étant pas exclus), alors, pour tout nombre réel R compris strictement entre A et B, il existe $r \in]\alpha, \beta[$ tel que $[\mathrm{D}f]_r = R$.*

Suggestion. Résolvons le cas $A < R < B$, l'autre s'établit de manière analogue. Il existe alors a, $b \in]\alpha, \beta[$ tels que $[\mathrm{D}f]_a < R < [\mathrm{D}f]_b$, puis $h > 0$ tel que $a + h$, $b + h \in]\alpha, \beta[$ et

$$\frac{f(a+h) - f(a)}{h} < R < \frac{f(b+h) - f(b)}{h}.$$

Cela étant, comme la fonction

$$g\colon\ [a, b] \to \mathbb{R} \quad x \mapsto \frac{f(x+h) - f(x)}{h}$$

est continue et réelle sur $[a, b]$, et dérivable sur $]a, b[$, le théorème des valeurs intermédiaires donne l'existence d'un point $x_0 \in]a, b[$ tel que $g(x_0) = R$. On a donc $f(x_0 + h) - f(x_0) = hR$ alors que le théorème des accroissements finis procure l'existence de $h' \in]0, h[$ tel que $f(x_0 + h) - f(x_0) = h[\mathrm{D}f]_{x_0+h'}$, ce qui suffit.

5.1.7 Continuité des fonctions continûment dérivables

Théorème 5.1.7.1 *Toute fonction continûment dérivable sur un ouvert de $\mathbb{R}^n$ y est continue.*

Preuve. Soit f une fonction continûment dérivable sur l'ouvert Ω de $\mathbb{R}^n$.

Etablissons tout d'abord l'énoncé sous l'hypothèse supplémentaire que f soit réel. Soit x un point de Ω. Pour tout $h \in \mathbb{R}^n$ tel que la boule $\{ y : |x - y| \leq |h| \}$

soit incluse dans Ω, le théorème des accroissements finis procure des accroissements auxiliaires $h^{(1)}, \ldots, h^{(n)} \in \mathbb{R}^n$ tels que $\left|h^{(k)}\right| \leq |h|$ pour tout $k \in \{1, \ldots, n\}$ et

$$f(x+h) - f(x) = \sum_{k=1}^{n} h_k [\mathrm{D}_k f]_{x+h^{(k)}}.$$

Si h tend vers 0, les accroissements $h^{(k)}$ tendent aussi vers 0 et ainsi, vu la continuité des dérivées de f sur Ω, les $[\mathrm{D}_k f]_{x+h^{(k)}}$ tendent vers $[\mathrm{D}_k f]_x$. Dès lors, on obtient

$$f(x+h) - f(x) \to \sum_{k=1}^{n} 0 \cdot [\mathrm{D}_k f]_x = 0$$

si $h \to 0$; d'où la conclusion.

Le cas général d'une fonction à valeurs complexes est alors immédiat car on vérifie directement que $\Re f$ et $\Im f$ sont des fonctions réelles et continûment dérivables sur Ω donc continues sur Ω, et ainsi $f = \Re f + i\Im f$ est continu sur Ω.∎

Théorème 5.1.7.2 (dérivabilité uniforme) *Si la fonction f est continûment dérivable sur l'ouvert Ω de $\mathbb{R}^n$, alors, pour tout $k \in \{1, \ldots, n\}$ et tout compact $K \subset \Omega$, on a*

$$\sup_{x \in K} \left| \frac{f(x+he_k) - f(x)}{h} - [\mathrm{D}_k f]_x \right| \to 0 \quad \textit{si} \quad h \to 0.$$

Preuve. Bien sûr, nous pouvons supposer f réel, quitte à considérer $\Re f$ et $\Im f$ séparément.

Cela étant, soit r un nombre réel strictement positif, quelconque si Ω est égal à $\mathbb{R}^n$ et vérifiant $0 < r < \mathrm{d}(K, \mathbb{R}^n \setminus \Omega)$ si Ω diffère de $\mathbb{R}^n$. Pour tout $x \in K$ et tout $h \in \mathbb{R}$ tel que $0 < |h| < r$, le théorème des accroissement finis dans $\mathbb{R}$ procure un accroissement auxiliaire $h^{(k)}(x) \in \mathbb{R}$ tels que $\left|h^{(k)}(x)\right| \leq |h|$ et

$$\begin{aligned} \sup_{x \in K} \left| \frac{f(x+he_k) - f(x)}{h} - [\mathrm{D}_k f]_x \right| &= \sup_{x \in K} \left| [\mathrm{D}_k f]_{x+h^{(k)}(x)e_k} - [\mathrm{D}_k f]_x \right| \\ &\leq \sup_{x \in K,\, |x-y| \leq |h|} \left| [\mathrm{D}_k f]_y - [\mathrm{D}_k f]_x \right|. \end{aligned}$$

Pour conclure, il suffit alors de noter que, vu la continuité donc la continuité uniforme de chacune des dérivées de f sur le compact $\{ z : \mathrm{d}(z, K) \leq r\}$, cette majorante tend vers 0 si h tend vers 0.∎

5.1.8 Théorème de l'ouvert connexe

Définition. Un ouvert Ω de $\mathbb{R}^n$ est *connexe* s'il n'existe pas de partition de Ω en deux ouverts non vides.

La notion d'ouvert connexe est assez malaisée à manipuler; aussi, dans une première étude, nous allons admettre les exemples suivants: ils sont établis dans le cours de deuxième candidature.

Exemples. a) *Un ouvert de $\mathbb{R}$ est connexe si et seulement s'il est égal à un intervalle ouvert ou à $\emptyset$.*

b) *Dans $\mathbb{R}^n$, tout intervalle ouvert est un ouvert connexe.*

c) *Dans $\mathbb{R}^n$, toute boule ouverte est un ouvert connexe.*

Théorème 5.1.8.1 (ouvert connexe) *Deux fonctions dérivables sur un ouvert connexe de $\mathbb{R}^n$ sont égales à une constante additive près si et seulement si leurs dérivées sont égales chacune à chacune.*

Explicitement, *si f et g sont deux fonctions dérivables sur un ouvert connexe Ω de $\mathbb{R}^n$, alors il existe $c \in \mathbb{C}$ tel que $f - g = c\chi_\Omega$ si et seulement si on a $\mathrm{D}_k f = \mathrm{D}_k g$ sur Ω pour tout $k \in \{1, \ldots, n\}$.*

Preuve. La condition est évidemment nécessaire et ne nécessite d'ailleurs pas que l'ouvert Ω soit connexe.

La condition est suffisante. En considérant la différence $f - g$ des deux fonctions, on est ramené à prouver que *si la fonction f est dérivable sur un ouvert connexe Ω de $\mathbb{R}^n$ et vérifie $\mathrm{D}_k f = 0$ pour tout $k \in \{1, \ldots, n\}$, alors il existe $c \in \mathbb{C}$ tel que $f = c\chi_\Omega$.*

Quitte à considérer $\Re f$ et $\Im f$ séparément, nous pouvons supposer f réel.

Etant donné $x \in \Omega$, considérons l'ensemble

$$\omega_x = \{\, y \in \Omega : f(y) = f(x) \} \,.$$

Bien sûr, ω_x n'est pas vide car il contient x. De plus, ω_x est ouvert: de fait, si y appartient à ω_x, il appartient à Ω et il existe donc $r > 0$ tel que $\{\, z : |z - y| \leq r \} \subset \Omega$. Pour conclure que ω_x est ouvert, il suffit alors de prouver que ω_x contient la boule fermée de centre y et de rayon r. Or, pour tout $h \in \mathbb{R}^n$ tel que $|h| \leq r$, le théorème des accroissements finis peut être appliqué: il existe des points $h^{(1)}, \ldots, h^{(n)} \in \mathbb{R}^n$ tels que $\left|h^{(k)}\right| \leq |h|$ pour tout $k \in \{1, \ldots, n\}$ et

$$f(y + h) - f(y) = \sum_{k=1}^{n} h_k [\mathrm{D}_k f]_{y+h^{(k)}} = 0.$$

Cela étant, soit x_0 un point de Ω. Si ω_{x_0} n'est pas égal à Ω, l'ensemble

$$\omega' = \{\, x \in \Omega : f(x) \neq f(x_0)\} = \bigcup_{x \in \Omega \setminus \omega_{x_0}} \omega_x$$

est un ouvert non vide, disjoint de ω_{x_0} et tel que $\omega_{x_0} \cup \omega' = \Omega$.

D'où une contradiction.∎

5.1.9 Espace $\mathrm{C}_p(\Omega)$

Définitions. a) Etant donné une fonction f définie sur un ouvert de $\mathbb{R}^n$ et un entier $p \in \mathbb{N}_0$, on appelle *dérivée partielle* d'*ordre* p de f, toute expression de la forme

$$\mathrm{D}_{\nu_1} \ldots \mathrm{D}_{\nu_p} f \qquad \text{avec} \;\; \nu_1, \ldots, \nu_p \in \{1, \ldots, n\}$$

si elle existe. Ce symbole est à comprendre comme suit: l'existence de $\mathrm{D}_{\nu_1}\mathrm{D}_{\nu_2}f$ signifie que f est dérivable par rapport à x_{ν_2} et que $\mathrm{D}_{\nu_2}f$ est dérivable par rapport à x_{ν_1}; cela étant, $\mathrm{D}_{\nu_1}\mathrm{D}_{\nu_2}f$ est égal à la dérivée par rapport à x_{ν_1} de $\mathrm{D}_{\nu_2}f$.

Les fonctions dérivables sur un ouvert Ω de $\mathbb{R}^n$ coïncident donc avec les fonctions qui admettent toutes les dérivées partielles d'ordre 1.

b) Soient p un élément de $\mathbb{N}_0$ et Ω un ouvert de $\mathbb{R}^n$. Une fonction est *p fois continûment dérivable sur* Ω (on dit aussi qu'elle est C_p *sur* Ω) si elle est définie sur Ω, si elle admet toutes les dérivées partielles d'ordre $q \leq p$ sur Ω et si chacune de ces dérivées est continue sur Ω.

Les fonctions continûment dérivables sur un ouvert Ω de $\mathbb{R}^n$ coïncident donc avec les fonctions une fois continûment dérivables sur Ω.

c) Une fonction f est *infiniment continûment dérivable sur un ouvert* Ω *de* $\mathbb{R}^n$ (on dit aussi qu'elle est C_∞ *sur* Ω) si elle est p fois continûment dérivable sur Ω quel que soit $p \in \mathbb{N}_0$.

Notations. a) Pour tout $p \in \mathbb{N}_0$ et tout ouvert Ω de $\mathbb{R}^n$, $\mathrm{C}_p(\Omega)$ désigne l'ensemble des fonctions p fois continûment dérivables sur Ω.

b) Pour tout ouvert Ω de $\mathbb{R}^n$, $\mathrm{C}_\infty(\Omega)$ désigne l'ensemble des fonctions infiniment continûment dérivables sur Ω.

Remarques. Soit Ω un ouvert de $\mathbb{R}^n$.

a) Si f appartient à $\mathrm{C}_p(\Omega)$, on a évidemment $f \in \mathrm{C}_q(\Omega)$ pour tout entier $q \leq p$.

b) Si f appartient à $\mathrm{C}_p(\Omega)$, $\mathrm{D}^\alpha f$ appartient à $\mathrm{C}_{p-q}(\Omega)$ pour toute dérivée D^α d'ordre $q \leq p$.

c) On a bien sûr

$$\mathrm{C}_\infty(\Omega) = \bigcap_{p \in \mathbb{N}_0} \mathrm{C}_p(\Omega).$$

Voici à présent un critère d'appartenance à l'espace $\mathrm{C}_p(\Omega)$.

Critère 5.1.9.1 *Etant donné $p \in \mathbb{N}_0$ et un ouvert Ω de $\mathbb{R}^n$, une fonction f définie sur Ω appartient à $\mathrm{C}_p(\Omega)$ si et seulement si toutes ses dérivées partielles d'ordre $q \leq p$ existent, les dérivées partielles d'ordre p étant continues sur Ω.*

Preuve. La condition est évidemment nécessaire.

La condition est suffisante. Si g désigne une dérivée partielle de f d'ordre $p-1$, g est une fonction continûment dérivable donc continue sur Ω. D'où la conclusion au moyen d'un nombre fini d'applications successsives de ce raisonnement.∎

Le cas $n = 1$ est particulier. D'une part, ce qui précède s'applique bien sûr si Ω est un ouvert de $\mathbb{R}$; on omet cependant le qualificatif "partiel" dans ce cas. D'autre part, on peut aussi considérer le cas d'un intervalle fermé.

Définitions. Etant donné $p \in \mathbb{N}_0$, une fonction f définie sur l'intervalle fermé $I = [a, b]$ de $\mathbb{R}$ est

a) *p fois dérivable sur I* si f, $\mathrm{D}f$, ..., $\mathrm{D}^p f$ existent sur I, où on a bien sûr posé $\mathrm{D}^1 f = \mathrm{D}f$ et $\mathrm{D}^{k+1} f = \mathrm{D}(\mathrm{D}^k f)$ pour tout $k \geq 1$,

b) *p fois continûment dérivable sur I* si elle est p fois dérivable sur I et telle que f, $\mathrm{D}f$, ..., $\mathrm{D}^p f \in \mathrm{C}_0(I)$,

c) *infiniment continûment dérivable sur I* si elle est q fois continûment dérivable sur I pour tout $q \in \mathbb{N}_0$.

Notations. Etant donné un intervalle fermé $[a, b]$ de $\mathbb{R}$ et $p \in \mathbb{N}_0$,

$$\mathrm{C}_p([a, b]) \quad \text{et} \quad \mathrm{C}_\infty([a, b])$$

désignent respectivement les ensembles des fonctions p fois continûment dérivables sur $[a, b]$ d'une part et infiniment continûment dérivables sur $[a, b]$ d'autre part.

Théorème 5.1.9.2 (structure) *Une fonction f définie sur l'intervalle fermé $[a, b]$ de $\mathbb{R}$ appartient à $\mathrm{C}_p([a, b])$ avec $p \in \mathbb{N}_0$ si et seulement si c'est la restriction à $[a, b]$ d'un élément de $\mathrm{C}_p(\mathbb{R})$.*

Preuve. La condition est nécessaire. Il suffit de vérifier que la fonction

$$F\colon \mathbb{R} \to \mathbb{C} \quad x \mapsto \begin{cases} \displaystyle\sum_{k=0}^{p} \frac{1}{k!}[\mathrm{D}^k f]_a (x-a)^k & \text{si } x < a \\ f(x) & \text{si } a \leq x \leq b \\ \displaystyle\sum_{k=0}^{p} \frac{1}{k!}[\mathrm{D}^k f]_b (x-b)^k & \text{si } b < x \end{cases}$$

convient, ce qui est direct.

La condition est évidemment suffisante.∎

Remarque. Le résultat précédent est aussi valable dans le cas $p = \infty$ (cf. le cours de deuxième candidature).

5.1.10 Théorème d'interversion des dérivées

Les dérivées partielles d'un élément de $\mathrm{C}_p(\Omega)$ sont très aisées à manipuler grâce au théorème qui va suivre et dont le lemme suivant constitue une version améliorée dans le cas des dérivées partielles d'ordre 2.

Lemme 5.1.10.1 *Soit f une fonction définie sur l'ouvert Ω de $\mathbb{R}^n$ avec $n \geq 2$. Si les dérivées $\mathrm{D}_j f$, $\mathrm{D}_k f$ et $\mathrm{D}_k \mathrm{D}_j f$ existent sur Ω et si $\mathrm{D}_k \mathrm{D}_j f$ est continu sur Ω, alors $\mathrm{D}_j \mathrm{D}_k f$ existe sur Ω et est égal à $\mathrm{D}_k \mathrm{D}_j f$.*

Preuve. Bien sûr, nous pouvons supposer f réel, quitte à considérer séparément $\Re f$ et $\Im f$, puis $f = \Re f + i\Im f$.

Etant donné $x_0 \in \Omega$, il existe $r > 0$ tel que $x_0 + \xi e_j + \eta e_k \in \Omega$ pour tous ξ, $\eta \in \mathbb{R}$ tels que $|\xi|$, $|\eta| \leq r$.

Pour tous ξ, $\eta \in \mathbb{R}$ tels que $|\xi|$, $|\eta| \leq r$, introduisons l'expression

$$A(\xi, \eta) = (f(x_0 + \xi e_j + \eta e_k) - f(x_0 + \xi e_j)) - (f(x_0 + \eta e_k) - f(x_0)) \,.$$

Remarquons que la fonction F définie sur $]-r, r[$ par

$$F(t) = f(x_0 + te_j + \eta e_k) - f(x_0 + te_j)$$

est réelle et dérivable. Dès lors, pour tous $\xi \in \,]-r, 0[\, \cup \,]0, r[$, une application du théorème des accroissements finis donne l'existence de $\xi' \in \mathbb{R}$ tel que $|\xi'| \leq |\xi|$ et

$$A(\xi, \eta) = \xi \left([\mathrm{D}_j f]_{x_0 + \xi' e_j + \eta e_k} - [\mathrm{D}_j f]_{x_0 + \xi' e_j} \right) .$$

Comme la fonction G définie sur $]-r, r[$ par

$$G(t) = [\mathrm{D}_j f]_{x_0 + \xi' e_j + te_k}$$

est réelle et dérivable, une deuxième application du théorème des accroissements finis donne l'existence de $\eta' \in \mathbb{R}$ tel que $|\eta'| \leq |\eta|$ et

$$A(\xi, \eta) = \xi \eta [\mathrm{D}_k \mathrm{D}_j f]_{x_0 + \xi' e_j + \eta' e_k}.$$

Cela étant, vu la continuité de $\mathrm{D}_k\mathrm{D}_j f$ sur Ω, nous obtenons

$$\lim_{\xi,\eta\to 0;\,\xi,\eta\neq 0} \frac{1}{\xi\eta} A(\xi,\eta) = [\mathrm{D}_k\mathrm{D}_j f]_{x_0}.$$

Pour tout $\varepsilon > 0$, il existe donc $\delta > 0$ tel que

$$\left| \frac{1}{\xi\eta} A(\xi,\eta) - [\mathrm{D}_k\mathrm{D}_j f]_{x_0} \right| \leq \varepsilon$$

pour tous ξ, $\eta \in \mathbb{R} \setminus \{0\}$ tels que $|\xi|$, $|\eta| \leq \delta$. On en déduit que

$$\left| \frac{1}{\xi} \left([\mathrm{D}_k f]_{x_0+\xi e_j} - [\mathrm{D}_k f]_{x_0} \right) - [\mathrm{D}_k\mathrm{D}_j f]_{x_0} \right| \leq \varepsilon$$

pour tout $\xi \in \mathbb{R} \setminus \{0\}$ tel que $|\xi| \leq \delta$, ce qui permet de conclure aussitôt.∎

Théorème 5.1.10.2 (interversion des dérivées) *Si une fonction appartient à* $\mathrm{C}_p(\Omega)$*, alors toutes ses dérivées partielles d'ordre* $q \leq p$ *qui ne diffèrent que par l'ordre dans lequel on dérive, sont égales.*

Preuve. Cela résulte aussitôt du lemme précédent car toute permutation est égale au produit d'un nombre fini de transpositions.∎

Notations. a) Ce théorème permet d'alléger considérablement la notation des dérivées partielles d'ordre $q \leq p$ des éléments de $\mathrm{C}_p(\Omega)$. En effet, si f appartient à $\mathrm{C}_p(\Omega)$ et si on a $q \leq p$, une dérivée partielle d'ordre q de f, à savoir $\mathrm{D}_{\nu_1} \ldots \mathrm{D}_{\nu_q} f$ où apparaît α_1 fois D_1, ..., α_n fois D_n, peut s'écrire

$$\mathrm{D}_1^{\alpha_1} \ldots \mathrm{D}_n^{\alpha_n} f \quad \text{ou même} \quad \mathrm{D}^\alpha f$$

s'il n'est pas nécessaire de spécifier les $\alpha_1, \ldots, \alpha_n \in \mathbb{N}$.

b) Parfois, afin d'unifier les notations, on pose $\mathrm{D}^0 f = f$ pour tout $f \in \mathrm{C}_p(\Omega)$.

Remarque. Dans le lemme, la continuité de la dérivée d'ordre 2 *est indispensable.* Considérons la fonction

$$f(x,y) = \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2} & \text{si } (x,y) \neq (0,0), \\ 0 & \text{si } (x,y) = (0,0). \end{cases}$$

Bien sûr, elle est dérivable sur $\mathbb{R}^2$, avec

$$\begin{aligned} [\mathrm{D}_1 f]_{(x,y)} &= \begin{cases} \dfrac{x^4y+4x^2y^3-y^5}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0), \\ 0 & \text{si } (x,y) = (0,0), \end{cases} \\ [\mathrm{D}_2 f]_{(x,y)} &= \begin{cases} \dfrac{-xy^4-4x^3y^2+x^5}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0), \\ 0 & \text{si } (x,y) = (0,0). \end{cases} \end{aligned}$$

Cela étant, $\mathrm{D}_1 f$ est dérivable par rapport à y et $\mathrm{D}_2 f$ est dérivable par rapport à x sur $\mathbb{R}^2 \setminus \{(0,0)\}$ et même en $(0,0)$ vu que

$$\begin{aligned}
[\mathrm{D}_2\mathrm{D}_1 f]_{(0,0)} &= \lim_{h\to 0,\, h\neq 0} \frac{[\mathrm{D}_1 f]_{(0,h)} - 0}{h} &= \lim_{h\to 0,\, h\neq 0} \frac{-h^5}{h^5} &= -1,\\
[\mathrm{D}_1\mathrm{D}_2 f]_{(0,0)} &= \lim_{h\to 0,\, h\neq 0} \frac{[\mathrm{D}_2 f]_{(h,0)} - 0}{h} &= \lim_{h\to 0,\, h\neq 0} \frac{h^5}{h^5} &= 1.
\end{aligned}$$

D'où la nécessité de la continuité portant sur la dérivée d'ordre deux.

Remarque. $* \rightarrow$ On peut établir que la fonction f définie sur $\mathbb{R}^2$ par

$$f(x,y) = \begin{cases} \frac{x^2}{y}\mathrm{e}^{-x^2/y} & \text{si } y > 0 \\ 0 & \text{si } y \leq 0 \end{cases}$$

est telle que $f(\cdot, y) \in \mathrm{C}_\infty(\mathbb{R})$ pour tout $y \in \mathbb{R}$ ainsi que $f(x,\cdot) \in \mathrm{C}_\infty(\mathbb{R})$ pour tout $x \in \mathbb{R}$. Cependant elle n'est pas continue sur $\mathbb{R}^2$. $\leftarrow *$

5.1.11 Génération des éléments de $\mathrm{C}_p(\Omega)$

Voici tout d'abord des exemples fondamentaux d'éléments de $\mathrm{C}_p(\mathbb{R}^n)$. Ils sont immédiats. Nous introduisons d'autres exemples fondamentaux au chapitre suivant.

Exemples fondamentaux. a) *Pour tout $c \in \mathbb{C}$, $c\chi_{\mathbb{R}^n}$ appartient à $\mathrm{C}_\infty(\mathbb{R}^n)$.*

b) *Pour tout $j \in \{1, \ldots, n\}$, $[\cdot]_j$ appartient à $\mathrm{C}_\infty(\mathbb{R}^n)$.*

Théorème 5.1.11.1 (génération) *Soient p un élément de $\mathbb{N}_0 \cup \{\infty\}$ et Ω un ouvert de $\mathbb{R}^n$.*

a) *Toute combinaison linéaire d'éléments de $\mathrm{C}_p(\Omega)$ appartient à $\mathrm{C}_p(\Omega)$.*

b) *Tout produit fini d'éléments de $\mathrm{C}_p(\Omega)$ appartient à $\mathrm{C}_p(\Omega)$.*

c) *Tout quotient de deux éléments de $\mathrm{C}_p(\Omega)$ appartient à $\mathrm{C}_p(\Omega)$ pour autant que le dénominateur ne s'annule en aucun point de Ω.*

d) *Si f appartient à $\mathrm{C}_p(\Omega)$, alors $\overline{f}$, $\Re f$ et $\Im f$ appartiennent à $\mathrm{C}_p(\Omega)$.*

e) *Si J appartient à $\mathbb{N}_0$, si $f_1, \ldots, f_J$ sont réels et appartiennent à $\mathrm{C}_p(\Omega)$, et si f appartient à $\mathrm{C}_p(\omega)$ où ω est un ouvert de $\mathbb{R}^J$ contenant l'ensemble de variation de $(f_1, \ldots, f_J)$ sur Ω, alors la fonction composée $f(f_1, \ldots, f_J)$ appartient à $\mathrm{C}_p(\Omega)$.*

f) *Si f appartient à $\mathrm{C}_p(\Omega)$, alors la restriction de f à tout ouvert Ω' de $\mathbb{R}^n$ inclus dans Ω appartient à $\mathrm{C}_p(\Omega')$.*

Preuve. Cela résulte directement des théorèmes de génération des fonctions dérivables et des fonctions continues, en procédant par récurrence.

Etablissons par exemple l'énoncé relatif aux fonctions composées, connu sous le nom de *théorème de dérivation multiple des fonctions composées.*

Si on a $p = 1$, le théorème de dérivation des fonctions composées affirme que $f(f_1, \ldots, f_J)$ est une fonction dérivable et que, pour tout $k \leq n$,

$$\begin{aligned}[\mathrm{D}_k f(f_1, \ldots, f_J)]_x &= \sum_{j=1}^{J} [\mathrm{D}_j f]_{(f_1(x),\ldots,f_J(x))} \cdot [\mathrm{D}_k f_j]_x \\ &= \sum_{j=1}^{J} [(\mathrm{D}_j f)(f_1, \ldots, f_J)]_x \cdot [\mathrm{D}_k f_j]_x.\end{aligned}$$

Dès lors le théorème de continuité des fonctions composées permet de conclure.

Supposons l'énoncé établi pour $p = 1, \ldots, p'$ et passons au cas $p = p' + 1$. Vu ce qui précède, $f(f_1, \ldots, f_J)$ appartient déjà à $\mathrm{C}_1(\Omega)$ et, vu les égalités que nous venons d'établir dans le cas $p = 1$, ses dérivées partielles d'ordre 1 se présentent sous la forme d'une somme de produits de fonctions appartenant à $\mathrm{C}_{p'}(\Omega)$, donc appartiennent à $\mathrm{C}_{p'}(\Omega)$, ce qui suffit.∎

Remarque. Il n'est guère possible de donner des formules aisées à manipuler, donnant les dérivées partielles d'ordre $q \leq p$ des éléments de $\mathrm{C}_p(\Omega)$ engendrés par des éléments de $\mathrm{C}_p(\Omega)$, au moyen des dérivées partielles de ces dernières fonctions. On a cependant les résultats suivants.

Proposition 5.1.11.2 *Pour toute combinaison linéaire* $\sum_{j=1}^{J} c_j f_j$ *d'éléments de* $\mathrm{C}_p(\Omega)$ *et toute dérivée* D^α *d'ordre* $q \leq p$*, on a*

$$\mathrm{D}^\alpha \left(\sum_{j=1}^{J} c_j f_j \right) = \sum_{j=1}^{J} c_j \mathrm{D}^\alpha f_j.∎$$

Théorème 5.1.11.3 (formule de Leibniz) *Si* Ω *est un ouvert de* $\mathbb{R}$ *et si* f *et* g *appartiennent à* $\mathrm{C}_p(\Omega)$*, on a*

$$\mathrm{D}^p(fg) = \sum_{k=0}^{p} \mathrm{C}_p^k \, \mathrm{D}^k f \cdot \mathrm{D}^{p-k} g.$$

Preuve. Procédons par récurrence sur p. Bien sûr la formule est correcte pour $p = 1$. Pour conclure, il suffit alors de prouver que, si elle est vraie pour p, elle l'est encore pour $p + 1$. Or, si f et g appartiennent à $\mathrm{C}_{p+1}(\Omega)$, ils appartiennent notamment à $\mathrm{C}_p(\Omega)$ et on peut écrire

$$\begin{aligned}
\mathrm{D}^{p+1}(fg) = \mathrm{D}(\mathrm{D}^p(fg)) &= \mathrm{D}\left(\sum_{k=0}^{p} \mathrm{C}_p^k \, \mathrm{D}^k f \cdot \mathrm{D}^{p-k} g\right) \\
&= \sum_{k=0}^{p} \mathrm{C}_p^k \, \mathrm{D}^{k+1} f \cdot \mathrm{D}^{p-k} g + \sum_{k=0}^{p} \mathrm{C}_p^k \, \mathrm{D}^k f \cdot \mathrm{D}^{p-k+1} g \\
&= \mathrm{C}_p^0 \, \mathrm{D}^0 f \cdot \mathrm{D}^{p+1} g + \sum_{k=1}^{p} \left(\mathrm{C}_p^k + \mathrm{C}_p^{k-1}\right) \mathrm{D}^k f \cdot \mathrm{D}^{p+1-k} g + \mathrm{C}_p^p \, \mathrm{D}^{p+1} f \cdot \mathrm{D}^0 g,
\end{aligned}$$

d'où la conclusion vu la formule du triangle de Pascal

$$\mathrm{C}_p^{k-1} + \mathrm{C}_p^k = \mathrm{C}_{p+1}^k, \quad \forall k \in \{1, \ldots, p\}.\blacksquare$$

Règle pratique. Pour utiliser aisément la formule de Leibniz, on procède comme suit: on construit le tableau

$$\begin{array}{lllll}
\mathrm{C}_p^0 & \mathrm{C}_p^1 & \mathrm{C}_p^2 & \ldots & \mathrm{C}_p^p \\
f & \mathrm{D} f & \mathrm{D}^2 f & \ldots & \mathrm{D}^p f \\
\mathrm{D}^p g & \mathrm{D}^{p-1} g & \mathrm{D}^{p-2} g & \ldots & g
\end{array}$$

La règle est alors d'"effectuer les produits dans les colonnes et sommer les fonctions obtenues de la sorte".

Proposition 5.1.11.4 *Si Ω est un ouvert de $\mathbb{R}$, si f et g appartiennent à $\mathrm{C}_p(\Omega)$ et si g diffère de 0 en tout point de Ω, alors on a*

$$\mathrm{D}^p \frac{f}{g} = \frac{F}{g^{p+1}}$$

où F désigne le déterminant de la matrice obtenue en remplaçant dans la matrice

$$\begin{pmatrix}
g & & & & \\
\mathrm{D} g & g & & & \\
\mathrm{D}^2 g & \mathrm{C}_2^1 \mathrm{D} g & g & & \\
\vdots & \vdots & \vdots & \ddots & \\
\mathrm{D}^p g & \mathrm{C}_p^1 \mathrm{D}^{p-1} g & \mathrm{C}_p^2 \mathrm{D}^{p-2} g & \ldots & g
\end{pmatrix}$$

la dernière colonne par le vecteur $(f, \mathrm{D} f, \ldots, \mathrm{D}^p f)$.

Preuve. Posons $h = f/g$. Il vient $f = gh$ avec f, g, $h \in \mathrm{C}_p(\Omega)$ et la formule de Leibniz donne

$$\mathrm{D}^k f = \sum_{j=0}^{k} \mathrm{C}_k^j \, \mathrm{D}^j h \cdot \mathrm{D}^{k-j} g$$

pour tout $k \in \{0, \ldots, p\}$. Cela se met aisément sous forme matricielle

$$\begin{pmatrix} f \\ \mathrm{D}f \\ \mathrm{D}^2 f \\ \vdots \\ \mathrm{D}^p f \end{pmatrix} = \begin{pmatrix} g & & & & \\ \mathrm{D}g & g & & & \\ \mathrm{D}^2 g & \mathrm{C}_2^1 \mathrm{D}g & g & & \\ \vdots & \vdots & \vdots & \ddots & \\ \mathrm{D}^p g & \mathrm{C}_p^1 \mathrm{D}^{p-1} g & \mathrm{C}_p^2 \mathrm{D}^{p-2} g & \ldots & g \end{pmatrix} \begin{pmatrix} h \\ \mathrm{D}h \\ \mathrm{D}^2 h \\ \cdots \\ \mathrm{D}^p h \end{pmatrix},$$

d'où la conclusion par la règle de Cramer.■

Proposition 5.1.11.5 *Si f appartient à $\mathrm{C}_p(\Omega)$, il vient*

$$\mathrm{D}^\alpha \overline{f} = \overline{\mathrm{D}^\alpha f}, \quad \mathrm{D}^\alpha(\Re f) = \Re\,(\mathrm{D}^\alpha f) \quad \textit{et} \quad \mathrm{D}^\alpha(\Im f) = \Im\,(\mathrm{D}^\alpha f)$$

pour toute dérivée d'ordre $q \leq p$.■

En ce qui concerne les dérivées partielles des fonctions composées, les formules se compliquent très rapidement. Ainsi, si l'hypothèse du théorème de dérivation multiple des fonctions composées est vérifiée pour $p = 2$, il vient successivement

$$[\mathrm{D}_k f(f_1, \ldots, f_J)]_x = \sum_{j=1}^{J} [(\mathrm{D}_j f)(f_1, \ldots, f_J)]_x \cdot [\mathrm{D}_k f_j]_x$$

et

$$\begin{aligned} [\mathrm{D}_l \mathrm{D}_k f(f_1, \ldots, f_J)]_x &= \sum_{i=1}^{J} \sum_{j=1}^{J} [(\mathrm{D}_i \mathrm{D}_j f)(f_1, \ldots, f_J)]_x \cdot [\mathrm{D}_l f_i]_x \cdot [\mathrm{D}_k f_j]_x \\ &\quad + \sum_{j=1}^{J} [(\mathrm{D}_j f)(f_1, \ldots, f_J)]_x \cdot [\mathrm{D}_l \mathrm{D}_k f_j]_x. \end{aligned}$$

5.1.12 Exercices

Exercice. *Si Ω est un ouvert de $\mathbb{R}$ et si les fonctions f, g appartiennent à $\mathrm{C}_p(\Omega)$, établir la formule*

$$g \cdot \mathrm{D}^p f = \sum_{k=0}^{p} (-1)^k \mathrm{C}_p^k \, \mathrm{D}^{p-k}(f \mathrm{D}^k g).$$

Suggestion. De fait on a successivement

$$\begin{aligned}\sum_{k=0}^{p}(-1)^k\mathrm{C}_p^k\,\mathrm{D}^{p-k}(f\mathrm{D}^k g) &= \sum_{k=0}^{p}(-1)^k\mathrm{C}_p^k\sum_{j=0}^{p-k}\mathrm{C}_{p-k}^j\mathrm{D}^j f\cdot\mathrm{D}^{p-j}g\\ =\ \sum_{j=0}^{p}\mathrm{D}^j f\cdot\mathrm{D}^{p-j}g\cdot\sum_{k=0}^{p-j}(-1)^k\mathrm{C}_p^k\mathrm{C}_{p-k}^j &= \sum_{j=0}^{p}\mathrm{C}_p^j\,\mathrm{D}^j f\cdot\mathrm{D}^{p-j}g\cdot\delta_{j,p}.\end{aligned}$$

Pour obtenir la deuxième égalité, on a permuté les deux sommes selon

$$\left\{\begin{array}{l}0\le k\le p\\ 0\le j\le p-k\end{array}\right\}\Leftrightarrow\left\{\begin{array}{l}0\le j\le p\\ 0\le k\le p\\ j\le p-k\end{array}\right\}\Leftrightarrow\left\{\begin{array}{l}0\le j\le p\\ 0\le k\le p-j\end{array}\right\}.$$

Pour obtenir la troisième égalité, il suffit de noter que

$$\mathrm{C}_p^k\mathrm{C}_{p-k}^j=\frac{p!}{k!\,(p-k)!}\cdot\frac{(p-k)!}{j!\,(p-k-j)!}=\mathrm{C}_p^j\mathrm{C}_{p-j}^k$$

donc

$$\sum_{k=0}^{p-j}(-1)^k\mathrm{C}_p^k\mathrm{C}_{p-k}^j=\mathrm{C}_p^j\sum_{k=0}^{p-j}(-1)^k\mathrm{C}_{p-j}^k=\left\{\begin{array}{ll}1 & \text{si } j=p\\ 0 & \text{si } j\ne p\end{array}\right\}.$$

Exercice. *Si f appartient à $\mathrm{C}_p(\mathbb{R}\setminus\{0\})$, établir que*

$$x^{-p-1}[\mathrm{D}^p f]_{1/x}=(-1)^p\mathrm{D}^p\big(x^{p-1}f(1/x)\big),\quad\forall x\in\mathbb{R}\setminus\{0\}.$$

Suggestion. Procédons par récurrence sur p. Pour $p=1$, on vérifie directement que la formule proposée est correcte. Supposons à présent qu'elle est correcte pour p et établissons qu'elle l'est encore pour $p+1$. Il vient successivement

$$\begin{aligned}&(-1)^{p+1}\mathrm{D}^{p+1}(x^p f(1/x))\\ &\quad=\ (-1)^{p+1}\mathrm{D}^p\big(px^{p-1}f(1/x)-x^{p-2}[\mathrm{D}f]_{1/x}\big)\\ &\quad=\ -p(-1)^p\mathrm{D}^p\big(x^{p-1}f(1/x)\big)+(-1)^p\mathrm{D}^p\big(x^{p-1}[x\mathrm{D}f]_{1/x}\big)\\ &\quad=\ -px^{-p-1}[\mathrm{D}^p f]_{1/x}+x^{-p-1}[\mathrm{D}^p(x\mathrm{D}f)]_{1/x}\\ &\quad=\ x^{-p-1}x^{-1}[\mathrm{D}^{p+1}f]_{1/x},\end{aligned}$$

ce qui suffit. Pour établir la dernière égalité, on a utilisé la formule de Leibniz selon

$$[\mathrm{D}^p(x\mathrm{D}f)]_{1/x}=[x\mathrm{D}^{p+1}f]_{1/x}+p[\mathrm{D}^p f]_{1/x}.$$

Exercices. $*\rightarrow$ a) *Vérifier que* $\Delta\ln(x^2+y^2)^{1/2}=0$ *sur* $\mathbb{R}^2\setminus\{0\}$.
b) *Vérifier que* $\Delta(x^2+y^2+z^2)^{-1/2}=0$ *sur* $\mathbb{R}^3\setminus\{0\}$. $\leftarrow *$

5.1.13 Formule de Taylor limitée

Théorème 5.1.13.1 (Taylor limité dans $\mathbb{R}$) *Si la fonction f est réelle et appartient à $\mathrm{C}_{p-1}([a,b]) \cap \mathrm{C}_p(]a,b[)$, alors il existe $r \in]a,b[$ tel que*

$$f(b) = f(a) + \frac{b-a}{1!}[\mathrm{D}f]_a + \cdots + \frac{(b-a)^{p-1}}{(p-1)!}[\mathrm{D}^{p-1}f]_a + \frac{(b-a)^p}{p!}[\mathrm{D}^p f]_r.$$

En particulier, on a l'énoncé suivant, connu lui aussi sous le nom de formule de Taylor limitée dans $\mathbb{R}$: *si $f \in \mathrm{C}_p(]a,b[)$ est réel, alors, pour tout $x \in]a,b[$ et tout accroissement $h \in \mathbb{R}$ tel que $x+h \in]a,b[$, il existe $\theta \in]0,1[$ tel que*

$$f(x+h) = f(x) + \frac{h}{1!}[\mathrm{D}f]_x + \cdots + \frac{h^{p-1}}{(p-1)!}[\mathrm{D}^{p-1}f]_x + \frac{h^p}{p!}[\mathrm{D}^p f]_{x+\theta h}.$$

Preuve. Tout revient à déterminer le nombre R défini par

$$f(b) = f(a) + \frac{b-a}{1!}[\mathrm{D}f]_a + \cdots + \frac{(b-a)^{p-1}}{(p-1)!}[\mathrm{D}^{p-1}f]_a + \frac{(b-a)^p}{p!}R.$$

Pour ce faire, considérons la fonction F définie sur $[a,b]$ par

$$F(y) = f(b) - f(y) - \frac{b-y}{1!}[\mathrm{D}f]_y - \cdots - \frac{(b-y)^{p-1}}{(p-1)!}[\mathrm{D}^{p-1}f]_y - \frac{(b-y)^p}{p!}R.$$

On vérifie de suite que cette fonction F est réelle et continue sur $[a,b]$, dérivable sur $]a,b[$ et que sa dérivée est donnée par

$$[\mathrm{D}F]_y = \frac{(b-y)^{p-1}}{(p-1)!}\left(R - [\mathrm{D}^p f]_y\right).$$

Comme on a de plus $F(a) = F(b) = 0$, le théorème de Rolle procure l'existence de $r \in]a,b[$ tel que $[\mathrm{D}F]_r = 0$, ce qui suffit.

Le cas particulier est immédiat. ∎

Passons au cas de l'espace $\mathbb{R}^n$.

Théorème 5.1.13.2 (Taylor limité dans $\mathbb{R}^n$) *Si Ω est un ouvert de $\mathbb{R}^n$, si $f \in \mathrm{C}_p(\Omega)$ est réel et si le segment d'extrémités x et $x+h$ (à savoir l'ensemble*

$\{x+th : t \in [0,1]\}$) est inclus dans Ω, alors il existe $\theta \in]0,1[$ tel que

$$\begin{aligned} f(x+h) &= f(x) + \frac{1}{1!}\sum_{k=1}^{n} h_k [\mathrm{D}_k f]_x \\ &+ \frac{1}{2!}\sum_{k_1=1}^{n}\sum_{k_2=1}^{n} h_{k_1}h_{k_2}[\mathrm{D}_{k_1}\mathrm{D}_{k_2} f]_x + \cdots \\ &+ \frac{1}{(p-1)!}\sum_{k_1=1}^{n}\cdots\sum_{k_{p-1}=1}^{n} h_{k_1}\cdots h_{k_{p-1}}[\mathrm{D}_{k_1}\cdots\mathrm{D}_{k_{p-1}} f]_x \\ &+ \frac{1}{p!}\sum_{k_1=1}^{n}\cdots\sum_{k_p=1}^{n} h_{k_1}\cdots h_{k_p}[\mathrm{D}_{k_1}\cdots\mathrm{D}_{k_p} f]_{x+\theta h}. \end{aligned}$$

Preuve. Comme le segment $\{x+th : t \in [0,1]\}$ est un compact inclus dans Ω, on vérifie de suite qu'il existe $\varepsilon > 0$ tel que $x+th$ appartienne à Ω pour tout $t \in]-\varepsilon, 1+\varepsilon[$. Cela étant, la fonction

$$F\colon\ ]-\varepsilon,1+\varepsilon[\to \mathbb{R} \quad t \mapsto f(x_1+th_1,\ldots,x_n+th_n)$$

est réelle et appartient à $\mathrm{C}_p(]-\varepsilon,1+\varepsilon[)$. Dès lors, vu la formule de Taylor limitée dans $\mathbb{R}$, il existe $\theta \in]0,1[$ tel que

$$F(1) = F(0) + \frac{1}{1!}[\mathrm{D}F]_0 + \cdots + \frac{1}{(p-1)!}[\mathrm{D}^{p-1}F]_0 + \frac{1}{p!}[\mathrm{D}^p F]_\theta.$$

Pour conclure, il suffit alors de prouver que, pour tout $l \leq p$, on a

$$[\mathrm{D}^l F]_t = \sum_{k_1=1}^{n}\cdots\sum_{k_l=1}^{n} h_{k_1}\cdots h_{k_l}[\mathrm{D}_{k_1}\cdots\mathrm{D}_{k_l} f]_{x+th}$$

pour tout $t \in]0,1[$. Procédons par récurrence. Pour $l=1$, cela résulte directement du théorème de dérivation des fonctions composées. De plus, si cette formule est correcte pour $j = 1, \ldots, l$ avec $l < p$, elle est correcte également pour $l+1$ car, pour tous $k_1, \ldots, k_l \in \{1,\ldots,n\}$ fixés, on a

$$\mathrm{D}_t[\mathrm{D}_{k_1}\cdots\mathrm{D}_{k_l} f]_{x+th} = \sum_{k_{l+1}=1}^{n} h_{k_{l+1}}[\mathrm{D}_{k_1}\cdots\mathrm{D}_{k_{l+1}} f]_{x+th}$$

vu le théorème de dérivation des fonctions composées.

D'où la conclusion.∎

Remarques. a) Dans le cas $p = 1$, on introduit la notion de *gradient* d'une fonction $f \in \mathrm{C}_p(\Omega)$: c'est l'application

$$\mathrm{grad}\colon \mathrm{C}_p(\Omega) \to \mathrm{C}_{p-1}(\Omega)^n \quad f \mapsto (\mathrm{D}_1 f, \ldots, \mathrm{D}_n f).$$

Au moyen du gradient, la formule de Taylor limitée s'écrit

$$f(x+h) = f(x) + \langle h, [\mathrm{grad}(f)]_{x+\theta h} \rangle .$$

b) Dans le cas $p = 2$, on introduit la notion de *matrice hessienne* d'une fonction $f \in \mathrm{C}_2(\Omega)$: c'est la matrice H_f de dimension $n \times n$ définie sur Ω par

$$(\mathrm{H}_f)_{j,k} = \mathrm{D}_j \mathrm{D}_k f.$$

Au moyen du gradient et de la matrice hessienne, la formule de Taylor limitée s'écrit

$$f(x+h) = f(x) + \langle h, [\mathrm{grad}(f)]_x \rangle + \frac{1}{2} \langle h, [\mathrm{H}_f]_{x+\theta h} h \rangle .$$

Exercice. Si f est une fonction continûment dérivable sur l'intervalle ouvert I de $\mathbb{R}^n$ et si toutes ses dérivées d'ordre 1 sont bornées sur I, établir que f est uniformément continu sur I.

Suggestion. Il existe $C > 0$ tel que $|[\mathrm{D}_k f]_x| \leq C$ pour tout $x \in I$ et tout $k \in \{1, \ldots, n\}$. Pour tous x, $y \in I$, le segment d'extrémités x et y étant inclus dans I, la formule de Taylor limitée à l'ordre 1 donne de suite $|f(x) - f(y)| \leq nC\,|x - y|$, ce qui suffit.

Exercice. Si $f \in \mathrm{C}_{p+1}(]a, b[)$ est réel et si $[\mathrm{D}^{p+1} f]_x$ diffère de 0 en $x \in]a, b[$, établir que le $\theta \in]0, 1[$ qui figure dans la formule de Taylor limitée à l'ordre p en x tend vers $1/(p+1)$ si $h \to 0$.

Suggestion. Comme la fonction $\mathrm{D}^p f$ appartient à $\mathrm{C}_1(]a, b[)$ et est réelle, il existe $\theta' \in]0, 1[$ tel que

$$[\mathrm{D}^p f]_{x+\theta h} = [\mathrm{D}^p f]_x + \theta h [\mathrm{D}^{p+1} f]_{x+\theta\theta' h}.$$

Il vient donc

$$f(x+h) = f(x) + \frac{h}{1!}[\mathrm{D} f]_x + \ldots + \frac{h^p}{p!}[\mathrm{D}^p f]_x + \frac{\theta h^{p+1}}{p!}[\mathrm{D}^{p+1} f]_{x+\theta\theta' h},$$

d'où on tire

$$\frac{\theta}{p!}[\mathrm{D}^{p+1} f]_{x+\theta\theta' h} = \frac{1}{(p+1)!}[\mathrm{D}^{p+1} f]_{x+\theta'' h}$$

si θ'' correspond au développement de Taylor limité à $p + 1$. D'où la conclusion en faisant tendre h vers 0, vu la continuité de $\mathrm{D}^{p+1} f$ sur $]a, b[$.

Proposition 5.1.13.3 *Si la fonction* $f \in \mathrm{C}_p(\Omega)$ *est telle que* $f(x_m) \to 0$ *et* $\mathrm{D}^\alpha f(x_m) \to 0$ *pour toute suite* x_m *de* Ω *qui converge vers un point de* $\Omega^\bullet$ *ou* ∞, *et toute dérivée* D^α *d'ordre* $\leq p$, *alors la fonction* F *définie sur* $\mathbb{R}^n$ *par*

$$F(x) = \begin{cases} f(x), & \forall x \in \Omega, \\ 0, & \forall x \in \mathbb{R}^n \setminus \Omega, \end{cases}$$

appartient à $\mathrm{C}_p(\mathbb{R}^n)$.

Un résultat analogue pour $\mathrm{C}_\infty(\mathbb{R}^n)$ *s'en déduit aussitôt.*

Preuve. Par une récurrence aisée, on se ramène directement au cas de $\mathrm{C}_1(\Omega)$. En fait, nous avons déjà les appartenances suivantes

$$F \in \mathrm{C}_0(\mathbb{R}^n) \cap \mathrm{C}_1(\Omega) \cap \mathrm{C}_1(\mathbb{R}^n \setminus \Omega^-).$$

Il n'y a donc un problème qu'aux points de $\Omega^\bullet$. Soit x_0 un point de $\Omega^\bullet$ et soit k un élément de $\{1, \ldots, n\}$. Calculons

$$\lim_{h \to 0,\, h \in \mathbb{R}_0} \frac{F(x_0 + he_k) - F(x_0)}{h}.$$

Comme $\mathrm{D}_k F$ est identiquement nul sur $\mathbb{R}^n \setminus \Omega^-$, il suffit d'établir que cette limite existe et vaut 0. Or $F(x_0)$ est égal à 0 et

a) si $x_0 + he_k$ n'appartient pas à Ω, $F(x_0 + he_k)$ vaut 0,

b) si $x_0 + he_k$ appartient à Ω et si, par exemple, h est strictement positif, soit h_0 la borne inférieure des $r > 0$ tels que le segment ouvert d'extrémités $x_0 + re_k$ et $x_0 + he_k$ soit inclus dans Ω. Il existe alors $\theta \in]0, 1[$ tel que

$$\begin{aligned} F(x_0 + he_k) - F(x_0) &= F(x_0 + he_k) = F(x_0 + he_k) - F(x_0 + h_0 e_k) \\ &= (h - h_0)[\mathrm{D}_k F]_{x_0 + h_0 e_k + \theta(h - h_0) e_k}. \end{aligned}$$

D'où la conclusion en passant à la limite car $h \to 0$ entraîne bien sûr $h_0 \to 0$ vu que $h_0 \in [0, h[$.∎

5.2 Fonctions réelles sur un intervalle de $\mathbb{R}$

5.2.1 Variation

Remarques. a) Si f est une fonction strictement monotone sur $A \subset \mathbb{R}$ et si $x, y \in A$ sont tels que $f(x) = f(y)$, on a nécessairement $x = y$.

b) Soit A une partie de $\mathbb{R}$. Si $f \in \mathrm{C}_0(A^-)$ est réel et si f est monotone sur A, alors f est aussi monotone sur A^-, vu le théorème de passage à la limite dans les inégalités.

Les limites des valeurs d'une fonction monotone sur un intervalle ouvert de $\mathbb{R}$ (borné ou non) sont régies par le résultat suivant.

Théorème 5.2.1.1 *Si f est une fonction réelle et croissante sur $]a,b[\subset \mathbb{R}$, les limites $\lim_{x\to y^-} f(x)$ et $\lim_{x\to y^+} f(x)$ existent pour tout $y \in]a,b[$ et sont finies. Si elles sont égales, f est continu en y.*

De plus, les limites $\lim_{x\to a^+} f(x)$ et $\lim_{x\to b^-} f(x)$ existent et sont finies ou égales à $-\infty$ ou $+\infty$.

Enfin, on a

$$\lim_{x\to a^+} f(x) \leq \lim_{x\to y^-} f(x) \leq f(y) \leq \lim_{x\to y^+} f(x) \leq \lim_{x\to b^-} f(x)$$

pour tout $y \in]a,b[$ et même

$$\lim_{x\to a^+} f(x) < \lim_{x\to y^-} f(x) \leq f(y) \leq \lim_{x\to y^+} f(x) < \lim_{x\to b^-} f(x)$$

pour tout $y \in]a,b[$ si f est strictement croissant sur $]a,b[$.

Il y a évidemment un *énoncé analogue pour les fonctions (strictement) décroissantes.*

Preuve. Supposons, par exemple, f croissant sur $]a,b[$. Si f est décroissant sur $]a,b[$, il suffit de procéder de même ou de considérer la fonction croissante $-f$.

Etablissons le cas relatif à y^-. Soit y_m une suite de $]a,b[$ qui converge en croissant strictement vers y. Bien sûr, la suite $f(y_m)$ est croissante et majorée par $f(y)$: elle converge donc et sa limite α est majorée par $f(y)$. Cela étant, soit x_m une suite de $]a,y[$ qui converge vers y. Pour tout $\varepsilon > 0$, il existe d'abord $m_0 \in \mathbb{N}_0$ tel que $f(y_{m_0}) + \varepsilon \geq \alpha$ puis $M \in \mathbb{N}$ tel que $y_{m_0} \leq x_m < y$ pour tout $m \geq M$. Mais, pour tout $m \geq M$, il existe $m' \in \mathbb{N}$ tel que $x_m \leq y_{m'}$. Au total, nous avons

$$m \geq M \Rightarrow \alpha - \varepsilon \leq f(y_{m_0}) \leq f(x_m) \leq f(y_{m'}) \leq \alpha \leq f(y).$$

De ce qui précède, on déduit aussitôt que

$$\lim_{x\to y^-} f(x) = \alpha \leq f(y).$$

Etablissons le cas relatif à b^-. Soit y_m une suite de $]a,b[$ qui converge en croissant strictement vers b. Bien sûr, la suite $f(y_m)$ est croissante. Si elle est majorée, elle converge vers une limite finie α et le raisonnement du cas y^- permet de conclure. Si elle n'est pas majorée, elle converge vers $+\infty$ et un raisonnement analogue à celui du cas y^- permet de conclure.

Pour les cas y^+ et a^+, il suffit de procéder de même.

De la sorte, nous avons établi l'énoncé relatif au cas d'une fonction croissante. Si f est strictement croissant, la considération de points $y' \in \,]a, y[$ et $y'' \in \,]y, b[$ permet immédiatement d'écrire

$$\lim_{x\to a^+} f(x) < f(y') < \lim_{x\to y^-} f(x) \quad \text{et} \quad \lim_{x\to y^+} f(x) < f(y'') < \lim_{x\to b^-} f(x).$$

Enfin, si les limites

$$\lim_{x\to y^-} f(x) \quad \text{et} \quad \lim_{x\to y^+} f(x)$$

sont égales, elles valent $f(y)$ et nous savons qu'alors y est un point de continuité de f.∎

La considération de la dérivée des fonctions réelles et dérivables sur un intervalle de $\mathbb{R}$ permet de déduire la monotonie de la fonction sur cet intervalle.

Théorème 5.2.1.2 *Si f est une fonction réelle et dérivable sur l'intervalle $]a, b[$ de $\mathbb{R}$ (borné ou non), alors f est croissant* (resp. *décroissant*) *sur $]a, b[$ si et seulement si on a* $\mathrm{D}f \geq 0$ (resp. $\mathrm{D}f \leq 0$) *sur* $]a, b[$.

Preuve. Etablissons le cas d'une fonction croissante; pour une fonction décroissante f, on procède de même ou on considère $-f$.

La condition est nécessaire. De fait, si f est une fonction réelle et croissante sur $]a, b[$, alors, pour tout $x \in \,]a, b[$, on a

$$[\mathrm{D}f]_x = \lim_{h\to 0,\, h\in\mathbb{R}_0} \frac{f(x+h) - f(x)}{h} \geq 0$$

puisque $f(x+h) - f(x) \geq 0$ pour tout $h \in \,]0, b-x[$.

La condition est suffisante. De fait, pour tous x, $y \in \,]a, b[$ tel que $x < y$, il existe $r \in \,]x, y[$ tel que $f(y) - f(x) = (y - x) \cdot [\mathrm{D}f]_r$, vu le théorème des accroissements finis.∎

Théorème 5.2.1.3 *Si f est une fonction réelle et dérivable sur l'intervalle $]a, b[$ de $\mathbb{R}$ (borné ou non), alors f est strictement croissant* (resp. *strictement décroissant*) *sur $]a, b[$ si et seulement si* $\mathrm{D}f$ *est* ≥ 0 (resp. ≤ 0) *sur $]a, b[$ et sans zéro identique dans $]a, b[$.*

Preuve. Ici également, établissons le cas relatif à la croissance. Pour une fonction décroissante f, on procède de même ou on considère $-f$.

La condition est nécessaire. Comme f est notamment croissant, nous avons déjà $\mathrm{D}f \geq 0$ sur $]a, b[$. De plus, si $x_0 \in \,]a, b[$ est un zéro identique de $\mathrm{D}f$, il existe bien sûr

$\varepsilon > 0$ tel que $\mathrm{D}f = 0$ sur $]x_0 - \varepsilon, x_0 + \varepsilon[$; il s'ensuit que f est constant sur l'ouvert connexe $]x_0 - \varepsilon, x_0 + \varepsilon[$. D'où la conclusion.

La condition est suffisante. Comme nous avons $\mathrm{D}f \geq 0$ sur $]a, b[$, nous savons déjà que f est croissant sur $]a, b[$. De plus, nous ne pouvons pas avoir $f(x_1) = f(x_2)$ avec x_1, $x_2 \in\,]a, b[$ tels que $x_1 < x_2$ car alors nous devrions avoir $f(x) = f(x_1)$ pour tout $x \in\,]x_1, x_2[$, donc $\mathrm{D}f = 0$ sur $]x_1, x_2[$. D'où la conclusion.∎

5.2.2 Théorème de la fonction inverse

Le résultat suivant donne de très bons renseignements sur les bijections entre intervalles de $\mathbb{R}$.

Théorème 5.2.2.1 (fonction inverse) *Toute fonction f réelle, continue et strictement croissante* (resp. *strictement décroissante*) *sur un intervalle ouvert $]a, b[$ de $\mathbb{R}$ (borné ou non) est une bijection entre $]a, b[$ et*

$$I' = \left]\lim_{x\to a^+} f(x), \lim_{x\to b^-} f(x)\right[\qquad \left(\text{resp. } \left]\lim_{x\to b^-} f(x), \lim_{x\to a^+} f(x)\right[\right).$$

Preuve. Le premier théorème du paragraphe précédent permet d'affirmer que f est une application de $]a, b[$ dans I'. Il s'agit d'une injection car f est strictement monotone sur $]a, b[$. Il s'agit aussi d'une surjection, vu le théorème des valeurs intermédiaires. D'où la conclusion.∎

La bijection inverse associée à la fonction f de l'énoncé précédent est alors donnée par la fonction g définie sur I' par $g(x') = x$ si et seulement si $f(x) = x'$. Cette fonction g est appelée *fonction inverse de f* et jouit des propriétés suivantes qui font partie du théorème de la fonction inverse et utilisent les notations de son énoncé. Les démonstrations sont établies dans le cas d'une fonction f strictement croissante; pour l'autre cas, on procède de même ou on considère $-f$.

a) *Le graphe de g est celui de f où on permute les rôles de la variable et de la valeur de la fonction.*

Preuve. De fait, on a bien sûr

$$\{\,(x', g(x')) : x' \in I'\} = \{\,(f(x), x) : x \in\,]a, b[\}\,.$$ ∎

b) *La fonction g est strictement croissante* (resp. *strictement décroissante*) *sur I'.*

Preuve. De fait, si x', $y' \in I'$ sont tels que $x' < y'$, on ne peut avoir $g(x') \geq g(y')$ car cela entrainerait $x' = f(g(x')) \geq f(g(y')) = y'$.∎

c) *La fonction g est continue sur I'.*

Preuve. Il suffit de prouver qu'étant donné un point $x' \in I'$ et des suites x'_m et y'_m qui convergent respectivement vers x'^+ et x'^-, on a $g(x'_m) \to g(x')$ et $g(y'_m) \to g(x')$. Vu b) et le premier théorème du paragraphe précédent, nous savons déjà que les suites $g(x'_m)$ et $g(y'_m)$ convergent; soient α et β leurs limites respectives. Bien sûr, on a α, $\beta \in]a, b[$. De plus, vu la continuité de f sur $]a, b[$, il vient

$$x'_m = f(g(x'_m)) \to f(\alpha) \quad \text{et} \quad y'_m = f(g(y'_m)) \to f(\beta).$$

Ceci entraîne évidemment les égalités $f(\alpha) = x' = f(\beta)$, donc la relation $\alpha = \beta$. D'où la conclusion.∎

d) *On a*

$$\lim_{x' \to (\lim_{x \to a^+} f(x))^+} g(x') = a^+ \quad \textit{et} \quad \lim_{x' \to (\lim_{x \to b^-} f(x))^-} g(x') = b^-$$

$$\left(\text{resp.} \quad \lim_{x' \to (\lim_{x \to a^+} f(x))^-} g(x') = a^+ \quad \textit{et} \quad \lim_{x' \to (\lim_{x \to b^-} f(x))^+} g(x') = b^- \right).$$

Preuve. Soit x_m une suite de $]a, b[$ qui converge vers a. On sait alors que la suite $f(x_m)$ appartient à I' et converge vers $\lim_{x \to a^+} f(x)$ donc que la suite $g(f(x_m))$ appartient à $]a, b[$ et converge vers

$$\lim_{x' \to (\lim_{x \to a^+} f(x))^+} g(x').$$

Or c'est la suite $g(f(x_m)) = x_m$. D'où la conclusion.∎

e) *Si, en outre, f est dérivable sur* $]a, b[$ (resp. *appartient à* $\mathrm{C}_p(]a, b[)$) *et si* $\mathrm{D}f$ *diffère de* 0 *en tout point de* $]a, b[$, *alors g est dérivable sur I'* (resp. *appartient à* $\mathrm{C}_p(I')$) *et est tel que*

$$[\mathrm{D}g]_{x'} = \frac{1}{[\mathrm{D}f]_{g(x')}}.$$

Preuve. Pour tout $x' \in I'$, on a

$$\lim_{h' \to 0,\, h' \in \mathbb{R}_0} \frac{g(x' + h') - g(x')}{h'} = \lim_{h' \to 0,\, h' \in \mathbb{R}_0} \frac{1}{\dfrac{f(x + h) - f(x)}{h}},$$

d'où la conclusion car $h = g(x' + h') - g(x')$ tend vers 0 si $h' \to 0$.∎

f) *La fonction f est la fonction inverse de g.*∎

Remarque. Le théorème de la fonction inverse s'étend sans problème au cas des fonctions réelles, continues et strictement monotones sur un intervalle d'un des types $[a, b]$, $]a, b]$ *ou* $[a, b[$.

5.2.3 Fonctions convexes et fonctions concaves

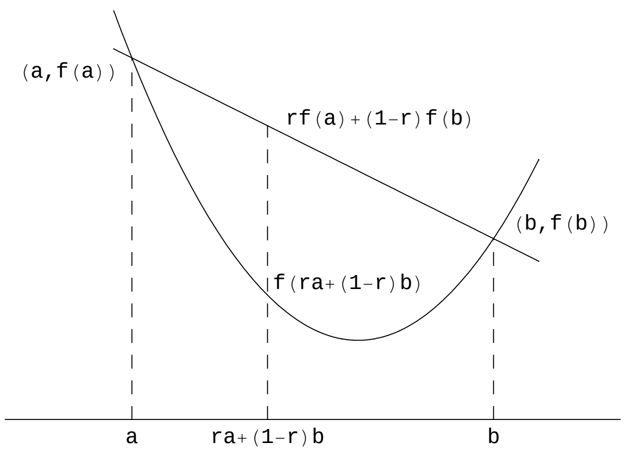
Définitions. Une fonction réelle f sur $A \subset \mathbb{R}$ est *convexe* (resp. *concave*) sur l'intervalle $I \subset A$ si, pour tous a, $b \in I$ et tout $r \in]0,1[$,

$$f(ra + (1-r)b) \leq rf(a) + (1-r)f(b)$$
$$(\text{resp. } f(ra + (1-r)b) \geq rf(a) + (1-r)f(b)).$$

L'interprétation géométrique est claire si on note que
a) les points $ra + (1-r)b$ avec $r \in [0,1]$ décrivent le segment d'extrémités a et b,
b) la droite déterminée par les points $(a, f(a))$ et $(b, f(b))$ a pour équation

$$y - f(b) = \frac{f(b) - f(a)}{b - a}(x - b)$$

et a donc $rf(a) + (1-r)f(b)$ pour ordonnée en $ra + (1-r)b$.



Remarquons aussi que la fonction réelle f sur l'intervalle I de $\mathbb{R}$ est concave si et seulement si la fonction $-f$ est convexe sur I. Ceci permet d'obtenir directement les propriétés des fonctions concaves à partir de celles des fonctions convexes.

Proposition 5.2.3.1 *Si la fonction réelle f sur l'intervalle I de $\mathbb{R}$ est convexe, alors, pour tout $J \in \mathbb{N}_0$, tous $x_1, \ldots, x_J \in I$ et tous $r_1, \ldots, r_J > 0$ tels que $\sum_{j=1}^{J} r_j = 1$, on a $\sum_{j=1}^{J} r_j x_j \in I$ et*

$$f\left(\sum_{j=1}^{J} r_j x_j\right) \leq \sum_{j=1}^{J} r_j f(x_j).$$

Enoncé analogue si on remplace “convexe” par “concave” et “$\leq$ ” par “$\geq$”.

Preuve. Procédons par récurrence. Pour $J = 2$, on retrouve la définition de la convexité. Si la propriété est vraie pour $J = 2, \ldots, K$, il suffit alors de noter que $\sum_{j=1}^{K+1} r_j x_j$ s'écrit aussi

$$(r_1 + \ldots + r_K) \sum_{j=1}^{K} \frac{r_j}{r_1 + \ldots + r_K} x_j + r_{K+1} x_{K+1} \in I$$

et de recourir à la convexité de f.∎

Critère 5.2.3.2 *Si f est une fonction réelle sur l'intervalle I de $\mathbb{R}$, les conditions suivantes sont équivalentes:*
(1) *f est convexe sur I,*
(2) *pour tout $x, y, z \in I$ tels que $x < y < z$, on a*

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(z) - f(x)}{z - x} \leq \frac{f(z) - f(y)}{z - y},$$

(3) *pour tout $x \in I$, la fonction $(f(\cdot) - f(x)/(\cdot - x)$ est croissante sur $I \setminus \{x\}$.*

Preuve. (1) $\Rightarrow$ (2). De $y = \theta x + (1 - \theta) z$ avec $\theta = (z - y)/(z - x) \in]0, 1[$, on tire

$$f(y) \leq \frac{z - y}{z - x} f(x) + \frac{y - x}{z - x} f(z),$$

c'est-à-dire

$$(z - x) f(y) \leq (z - y) f(x) + (y - x) f(z)$$

donc

$$(z - x)(f(y) - f(x)) \leq (y - x)(f(z) - f(x)),$$

c'est-à-dire la première inégalité. La seconde inégalité quant à elle a lieu si et seulement si

$$(z - y)(f(z) - f(x)) \leq (z - x)(f(z) - f(y)),$$

c'est-à-dire si et seulement si l'avant dernière inégalité a lieu, ce qui suffit.

(2) $\Rightarrow$ (3) est direct (on considérera les cas $y < z < x$, $y < x < z$ et $x < y < z$).

(3) $\Rightarrow$ (1). Pour tous $x, y \in I$ et $r \in]0, 1[$, on a $x < rx + (1 - r) y < y$ donc

$$\frac{f(x) - f(rx + (1 - r) y)}{(1 - r)(x - y)} \leq \frac{f(y) - f(rx + (1 - r) y)}{r(y - x)},$$

ce qui suffit.∎

Remarques. a) L'interprétation de $(f(y) - f(x))/(y - x)$ comme étant le coefficient angulaire de la droite passant par les points $(x, f(x))$ et $(y, f(y))$ rend cet énoncé bien plus parlant.

b) On a bien sûr un énoncé analogue avec “concave” au lieu de “convexe”, “$\geq$” au lieu de “$\leq$” et “décroissante” au lieu de “croissante”.

Théorème 5.2.3.3 *Toute fonction réelle et convexe* (resp. *concave*) *sur un intervalle ouvert de* $\mathbb{R}$ *est continue sur cet intervalle.*

Preuve. Il suffit bien sûr de traiter le cas de la convexité.

Soit f une fonction réelle et convexe sur l'intervalle ouvert I de $\mathbb{R}$. Pour tout $a \in I$ et tous x_0, x, y, $y_0 \in I$ tels que $x_0 < x < a < y < y_0$, le critère précédent donne

$$(y-a)\frac{f(x_0)-f(a)}{x_0-a} \leq f(y)-f(a) \leq (y-a)\frac{f(y_0)-f(a)}{y_0-a}$$

et

$$(x-a)\frac{f(y_0)-f(a)}{y_0-a} \leq f(x)-f(a) \leq (x-a)\frac{f(x_0)-f(a)}{x_0-a},$$

donc respectivement

$$\lim_{y\to a^+} f(y) = f(a) \quad \text{et} \quad \lim_{x\to a^-} f(x) = f(a),$$

ce qui suffit.∎

Remarque. Le résultat précédent ne s'étend pas au cas d'un intervalle non ouvert de $\mathbb{R}$. Pour s'en convaincre, il suffit, par exemple, de considérer la fonction

$$f\colon [0,1[\, \to \mathbb{R} \quad x \mapsto \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{si } 0 < x < 1 \end{cases}$$

qui est convexe sur $[0,1[$ mais non continue en 0.

Théorème 5.2.3.4 *Une fonction réelle et dérivable f sur un intervalle ouvert I de* $\mathbb{R}$ *est convexe* (resp. *concave*) *sur I si et seulement si sa dérivée est croissante* (resp. *décroissante*) *sur I.*

Preuve. Il suffit bien sûr de traiter le cas de la convexité.

La condition est nécessaire. De fait, pour tous x, $y \in I$ tels que $x < y$, il existe une suite $h_m > 0$ convergeant vers 0 et telle que $x + h_m < y$ et $y + h_m \in I$ pour tout $m \in \mathbb{N}_0$. Vu le critère, il vient alors

$$\frac{f(x+h_m)-f(x)}{h_m} \leq \frac{f(y+h_m)-f(y)}{h_m}, \quad \forall m \in \mathbb{N}_0,$$

donc $[\mathrm{D}f]_x \leq [\mathrm{D}f]_y$ par passage à la limite sur m.

La condition est suffisante. Pour tous x, $y \in I$ tels que $x < y$ et tout $r \in]0,1[$,

$$f(rx + (1-r)y) - rf(x) - (1-r)f(y)$$

est égal à

$$r(f(x + (1-r)(y-x)) - f(x)) + (1-r)(f(y + r(x-y)) - f(y)).$$

Cela étant, vu le théorème des accroissements finis, il existe θ, $\mu \in]0,1[$ tels que cette expression soit égale à

$$\begin{aligned} & r(1-r)(y-x)[\mathrm{D}f]_{x+\theta(1-r)(y-x)} + (1-r)r(x-y)[\mathrm{D}f]_{y+\mu r(x-y)} \\ = \ & r(1-r)(y-x)\left([\mathrm{D}f]_{x+\theta(1-r)(y-x)} - [\mathrm{D}f]_{y+\mu r(x-y)}\right) \end{aligned}$$

avec $x + \theta(1-r)(y-x) < y + \mu r(x-y)$, ce qui suffit.∎

Corollaire 5.2.3.5 *Une fonction réelle et deux fois dérivable sur un intervalle ouvert I de $\mathbb{R}$ est convexe* (resp. *concave*) *sur I si et seulement si sa dérivée seconde est une fonction positive* (resp. *négative*) *sur I.*∎

Proposition 5.2.3.6 *Si f est une fonction réelle et convexe sur l'intervalle ouvert $]a,b[$ de $\mathbb{R}$, alors l'ensemble A des points de $]a,b[$ où f n'est pas dérivable est dénombrable et $\mathrm{D}f$ est continu sur $]a,b[\setminus A$.*

En fait,

a) *f admet une dérivée à droite et une dérivée à gauche en tout point de $]a,b[$,*

b) *les deux fonctions "dérivée à droite" et "dérivée à gauche" sont respectivement continues à droite et à gauche,*

c) *A est l'ensemble des points de $]a,b[$ où cette dérivée à droite n'est pas continue; c'est aussi l'ensemble des points où la dérivée à gauche n'est pas continue,*

d) *sur $]a,b[\setminus A$, les dérivées à gauche et à droite coïncident.*

Preuve. Vu le critère, pour tout $x \in]a,b[$, les limites

$$f'_+(x) = \lim_{h\to 0,\, h>0} \frac{f(x+h) - f(x)}{h} \quad \text{et} \quad f'_-(x) = \lim_{h\to 0,\, h<0} \frac{f(x+h) - f(x)}{h}$$

existent et sont finies. De plus, comme pour tous nombres réels c, d et x tels que $a < c < x < d < b$, on a

$$f'_-(c) \leq f'_+(c) \leq \frac{f(x) - f(c)}{x - c} \leq \frac{f(x) - f(d)}{x - d} \leq f'_-(d) \leq f'_+(d),$$

les fonctions f'_+ et f'_- ainsi définies sur $]a, b[$ sont croissantes et telles que $f'_- \leq f'_+$. (Ce sont les fonctions appelées dérivée à droite et dérivée à gauche, dans l'énoncé.)

Etablissons à présent que f'_+ est une fonction continue à droite sur $]a, b[$. Vu la croissance de la fonction $(f(\cdot) - f(x))/(\cdot - x)$ sur $]a, x[\cup]x, b[$ quel que soit $x \in]a, b[$, on a bien sûr

$$a < c < x < y < b \Rightarrow f'_+(x) \leq \frac{f(y) - f(x)}{y - x}$$

et, comme f est continu sur $]a, b[$,

$$a < c < x < y < b \Rightarrow \lim_{x \to c^+} f'_+(x) \leq \frac{f(y) - f(c)}{y - c}$$

donc

$$a < c < b \Rightarrow \lim_{x \to c^+} f'_+(x) \leq f'_+(c),$$

ce qui suffit car f'_+ est une fonction croissante.

On établit de même que f'_- est continu à gauche sur $]a, b[$.

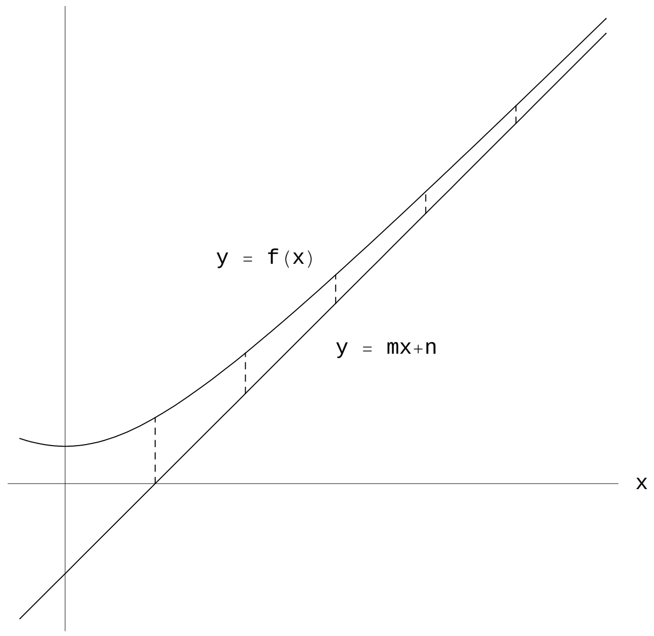
Enfin on déduit de suite des premières inégalités établies que f'_- est continu en tout point de continuité de f'_+ et inversement, avec égalité de ces deux fonctions en tous ces points. D'où la conclusion car étant une fonction croissante sur $]a, b[$, f'_+ a une partie dénombrable de $]a, b[$ pour ensemble de points de discontinuité.∎

5.2.4 Asymptote au graphe

Définition. Une *asymptote* au graphe d'une fonction réelle f sur un intervalle I non majoré (resp. non minoré) de $\mathbb{R}$ est une droite d'équation $y = mx + n$ telle que

$$\lim_{x \to +\infty} |f(x) - mx - n| = 0 \quad \text{(resp. } \lim_{x \to -\infty} |f(x) - mx - n| = 0\text{)}.$$

L'interprétation géométrique est claire.



Proposition 5.2.4.1 *Une fonction f réelle sur un intervalle I non majoré* (resp. *non minoré*) *de $\mathbb{R}$ admet la droite d'équation $y = mx + n$ pour asymptote en* $+\infty$ (resp. *en* $-\infty$) *si et seulement si*

$$\lim_{x\to+\infty} \frac{f(x)}{x} = m \quad et \quad \lim_{x\to+\infty} (f(x) - mx) = n$$
$$(\text{resp. } \lim_{x\to-\infty} \frac{f(x)}{x} = m \quad et \quad \lim_{x\to-\infty} (f(x) - mx) = n).$$

En particulier, si f admet une asymptote en $+\infty$ (resp. $-\infty$), *elle est unique.* ∎

5.3 Applications

5.3.1 Applications continues

Soit A une partie de $\mathbb{R}^p$. Une application $f\colon A \to \mathbb{R}^q$ est bien sûr caractérisée par q fonctions réelles $f_j\colon A \to \mathbb{R}$, à savoir les fonctions définies sur A par $f_j(x) = [f(x)]_j$ pour tout $j \in \{1, \ldots, p\}$ et tout $x \in A$; f_j est évidemment appelé *la j-ème composante de f.*

Définition. Une application f de $A \subset \mathbb{R}^p$ dans $\mathbb{R}^q$ est *continue en* $x \in A$ si, pour tout voisinage V de $f(x)$ dans $\mathbb{R}^q$, il existe un voisinage U de x dans $\mathbb{R}^p$ tel que $f(U \cap A) \subset V$. Elle est *continue sur* A si elle est continue en tout point de A.

Critère 5.3.1.1 *Une application f de $A \subset \mathbb{R}^p$ dans $\mathbb{R}^q$ est continue en $x \in A$* (resp. *continue sur A*) *si et seulement si chacune de ses composantes est une fonction continue en x* (resp. *continue sur A*).

Preuve. C'est une conséquence directe des inégalités

$$|x_j| \leq |x| \leq \sum_{k=1}^{n} |x_k|$$

valables pour tout $x \in \mathbb{R}^n$ et tout $j \in \{1, \dots, n\}$.∎

Corollaire 5.3.1.2 *Une application f de $A \subset \mathbb{R}^p$ dans $\mathbb{R}^q$ est continue en $x \in A$ si et seulement si on a $f(x_m) \to f(x)$ pour toute suite x_m de A qui converge vers x.*

Elle est continue sur A si et seulement si, pour toute suite x_m de A qui converge vers un point x de A, on a $f(x_m) \to f(x)$.∎

Théorème 5.3.1.3 (applications composées) *Soient f une application continue de $A \subset \mathbb{R}^p$ dans $\mathbb{R}^q$ et g une application continue de $B \subset \mathbb{R}^q$ dans $\mathbb{R}^r$. Si B contient $\{ f(x) : x \in A\}$, alors $g \circ f$ est une application continue de A dans $\mathbb{R}^r$.*

Preuve. Cela résulte aussitôt du théorème de continuité des fonctions composées, appliqué aux composantes de $g \circ f$.

Cela résulte aussi directement de la définition: soient x un point de A et W un voisinage de $g \circ f(x)$ dans $\mathbb{R}^r$. Vu la continuité de g en $f(x)$, il existe un voisinage V de $f(x)$ dans $\mathbb{R}^q$ tel que $g(V \cap B) \subset W$. Vu la continuité de f en x, il existe ensuite un voisinage U de x dans $\mathbb{R}^p$ tel que $f(U \cap A) \subset V$. Au total, on a alors $g \circ f(U \cap A) \subset g(V \cap B) \subset W$, ce qui suffit.∎

5.3.2 Applications différentiables

Remarque. Rappelons qu'une fonction f définie sur un ouvert Ω de $\mathbb{R}$ est dérivable en $x \in \Omega$ si et seulement si la limite

$$\lim_{h \to 0,\, h \neq 0} \frac{f(x+h) - f(x)}{h}$$

existe et est finie, auquel cas elle est appelée dérivée de f en x et notée $[\mathrm{D}f]_x$ ou $\mathrm{D}f(x)$. En fait, le nombre $[\mathrm{D}f]_x$ peut aussi être considéré comme étant un opérateur linéaire T de $\mathbb{C}$ dans $\mathbb{C}$ tel que

$$\lim_{h \to 0,\, h \neq 0} \frac{f(x+h) - f(x) - T \cdot h}{h} = 0.$$

Ce point de vue débouche sur une autre généralisation de la notion de dérivabilité.

Définition. Une application f de l'ouvert Ω de $\mathbb{R}^p$ dans $\mathbb{R}^q$ est *différentiable en* $x \in \Omega$ s'il existe un opérateur linéaire $T\colon \mathbb{R}^p \to \mathbb{R}^q$ tel que

$$\lim_{h \to 0,\, h \neq 0} \frac{|f(x+h) - f(x) - T \cdot h|}{|h|} = 0.$$

Elle est *différentiable sur* Ω si elle est différentiable en tout point de Ω.

Théorème 5.3.2.1 (unicité) *Si l'application f de l'ouvert Ω de $\mathbb{R}^p$ dans $\mathbb{R}^q$ est différentiable en $x \in \Omega$, il existe un et un seul opérateur linéaire $T\colon \mathbb{R}^p \to \mathbb{R}^q$ tel que*

$$\lim_{h \to 0,\, h \neq 0} \frac{|f(x+h) - f(x) - T \cdot h|}{|h|} = 0.$$

Preuve. De fait, si l'opérateur linéaire $S\colon \mathbb{R}^p \to \mathbb{R}^q$ est aussi tel que

$$\lim_{h \to 0,\, h \neq 0} \frac{|f(x+h) - f(x) - S \cdot h|}{|h|} = 0,$$

alors, pour tout $y \in \mathbb{R}^p \setminus \{0\}$, on a bien sûr

$$\frac{|Ty - Sy|}{|y|} = \lim_{t \to 0,\, t \in \mathbb{R}_0} \frac{|T(ty) - S(ty)|}{|ty|} = 0$$

donc $Ty = Sy$. La conclusion s'ensuit aussitôt.∎

Notation. Si l'application f de l'ouvert Ω de $\mathbb{R}^p$ dans $\mathbb{R}^q$ est différentiable en $x \in \Omega$, l'opérateur linéaire unique $T\colon \mathbb{R}^p \to \mathbb{R}^q$ vérifiant l'égalité

$$\lim_{h \to 0,\, h \neq 0} \frac{|f(x+h) - f(x) - T \cdot h|}{|h|} = 0$$

est appelé *différentielle de f en x* et est noté $[\mathrm{D}f]_x$, $\mathrm{D}f(x)$ ou f_{*x}. Si f est différentiable sur Ω, la *différentielle de f* est l'application $\mathrm{D}f$, notée aussi f_*.

Théorème 5.3.2.2 *Toute application différentiable d'un ouvert Ω de $\mathbb{R}^p$ dans $\mathbb{R}^q$ est continue.*

Preuve. De fait, si f est une telle application, on a

$$\begin{aligned} &\lim_{h \to 0} |f(x+h) - f(x)| \\ &\quad \leq \lim_{h \to 0} |h| \cdot \frac{|f(x+h) - f(x) - [\mathrm{D}f]_x h|}{|h|} + \lim_{h \to 0} \|[\mathrm{D}f]_x\| \cdot |h| = 0 \end{aligned}$$

en tout $x \in \Omega$, ce qui suffit.∎

Critère 5.3.2.3 *L'application f de l'ouvert Ω de $\mathbb{R}^p$ dans $\mathbb{R}^q$ est différentiable sur Ω si et seulement si chacune de ses composantes est différentiable sur Ω, auquel cas $\mathrm{D}f$ est le vecteur colonne $(\mathrm{D}f_1, \cdots, \mathrm{D}f_q)$.*

Preuve. La condition est nécessaire. Cela résulte aussitôt de la majoration

$$|f_j(x+h) - f_j(x) - < l_j, h >| \le |f(x+h) - f(x) - [\mathrm{D}f]_x h|$$

valable pour tout $j \in \{1, \cdots, q\}$ et tout $h \in \mathbb{R}^p$ de module suffisamment petit si l_j désigne la j-ème ligne de la matrice $[Df]_x$.

La condition est suffisante. Cela résulte aussitôt de la majoration

$$|f(x+h) - f(x) - Th| \le \sum_{j=1}^{q} |f_j(x+h) - f_j(x) - [\mathrm{D}f_j]_x h|$$

valable pour tout $h \in \mathbb{R}^p$ de module suffisamment petit si T désigne la matrice dont $[\mathrm{D}f_j]_x$ est la j-ème ligne pour tout $j \in \{1, \ldots, q\}$.∎

La différentiabilité d'une application est assez intimement liée à la (continue) dérivabilité de chacune de ses composantes.

Théorème 5.3.2.4 *Si l'application f de l'ouvert Ω de $\mathbb{R}^p$ dans $\mathbb{R}^q$ est différentiable, chacune de ses composantes est dérivable sur Ω et on a*

$$[\mathrm{D}f]_{x,j,k} = [\mathrm{D}_k f_j]_x, \quad \forall j \in \{1, \cdots, q\}, \quad \forall k \in \{1, \ldots, p\}, \quad \forall x \in \Omega.$$

Preuve. Pour tout $x \in \Omega$, il existe $r > 0$ tel que la boule $\{\, y \in \mathbb{R}^p : |x-y| \le r\}$ soit incluse dans Ω. Dès lors, pour tout $h \in \mathbb{R}^p \setminus \{0\}$ tel que $|h| \le r$ et tout $j \in \{1, \cdots, q\}$, on a

$$\left| \frac{f_j(x+he_k) - f_j(x)}{h} - [\mathrm{D}f_j]_{x,k} \right| \le \frac{|f(x+H) - f(x) - [\mathrm{D}f]_x H|}{|H|}$$

si on pose $H = he_k$, ce qui permet de conclure aussitôt.∎

Théorème 5.3.2.5 *Si l'application f de l'ouvert Ω de $\mathbb{R}^p$ dans $\mathbb{R}^q$ a toutes ses composantes qui appartiennent à* $\mathrm{C}_1(\Omega)$—(on dit alors que *f est une application C_1)—, alors f est différentiable sur Ω.*

Preuve. Etant donné $x \in \Omega$, soient T la matrice réelle de dimension $q \times p$ dont l'élément (j,k) est égal à $[\mathrm{D}_k f_j]_x$ et $\{\, y \in \mathbb{R}^p : |x-y| \le r\}$ une boule incluse dans Ω. Cela étant, pour tout $h \in \mathbb{R}^p \setminus \{0\}$ tel que $|h| \le r$ et tout élément j de $\{1, \cdots, q\}$, le théorème des accroissements finis procure pour tout $k \leq p$ un accroissement auxiliaire $h_{j,k}$ tel que $|h_{j,k}| \leq |h|$ et

$$f_j(x+h) - f_j(x) = \sum_{k=1}^{p} h_k [\mathrm{D}_k f_j]_{x+h_{j,k}}.$$

Dès lors, il vient

$$\begin{aligned}\frac{|f(x+h)-f(x)-Th|}{|h|} &\leq \frac{1}{|h|}\sum_{j=1}^{q}\left| f_j(x+h)-f_j(x)-\sum_{k=1}^{p} h_k[\mathrm{D}_k f_j]_x\right| \\ &\leq \frac{1}{|h|}\sum_{j=1}^{q}\sum_{k=1}^{p}|h_k|\left|[\mathrm{D}_k f_j]_{x+h_{j,k}} - [\mathrm{D}_k f_j]_x\right|\end{aligned}$$

où la majorante tend vers 0 si $h \to 0$. D'où la conclusion.∎

Remarques. Soit f une application de l'ouvert Ω de $\mathbb{R}^p$ dans $\mathbb{R}^q$.

a) Le dernier théorème donne une condition suffisante, très aisée à vérifier, pour déterminer si f est différentiable sur Ω. Il ne donne pas une condition nécessaire car nous savons qu'il existe une fonction réelle dérivable sur $\mathbb{R}$, dont la dérivée n'est pas continue sur $\mathbb{R}$.

b) L'avant dernier théorème permet de calculer très aisément la différentielle d'une application différentiable..

Exemples fondamentaux. a) *Toute application constante* $f\colon \mathbb{R}^p \to \mathbb{R}^q$ *est différentiable et telle que* $[\mathrm{D}f]_x = 0$ *pour tout* $x \in \mathbb{R}^p$.

b) *Toute application linéaire* $f : \mathbb{R}^p \to \mathbb{R}^q$ *est différentiable et telle que* $[\mathrm{D}f]_x = f$ *pour tout* $x \in \mathbb{R}^p$.

En particulier, *l'application*

$$s\colon \mathbb{R}^2 \to \mathbb{R} \quad (x,y) \mapsto x+y$$

est différentiable et telle que $[\mathrm{D}s]_{(x,y)} = s$ *pour tout* $(x,y) \in \mathbb{R}^2$.

c) *L'application*

$$p\colon \mathbb{R}^2 \to \mathbb{R} \quad (x,y) \mapsto xy$$

est différentiable et telle que $[\mathrm{D}p]_{(x,y)} = (y,x)$ *pour tout* $(x,y) \in \mathbb{R}^2$.

Théorème 5.3.2.6 (applications composées) *Si f et g sont des applications différentiables de l'ouvert Ω de $\mathbb{R}^p$ dans $\mathbb{R}^q$ et de l'ouvert ω de $\mathbb{R}^q$ dans $\mathbb{R}^r$ respectivement et si on a $f(\Omega) \subset \omega$, alors l'application composée $h = g \circ f\colon \Omega \to \mathbb{R}^r$ est différentiable et telle que $\mathrm{D}h = ((\mathrm{D}g)\circ f)\mathrm{D}f$.*

Preuve. Il suffit d'établir que, pour tout $x \in \Omega$, on a

$$\lim_{h \to 0,\, h \neq 0} \frac{1}{|h|} \left| g \circ f(x+h) - g \circ f(x) - [\mathrm{D}g]_{f(x)} [\mathrm{D}f]_x h \right| = 0.$$

Pour tout $x \in \Omega$, il existe $r > 0$ tel que $\{ y \in \mathbb{R}^p : |x - y| \leq r \} \subset \Omega$. Cela étant, pour tout $h \in \mathbb{R}^n$ tel que $|h| \leq r$, posons

$$\varphi(h) = f(x+h) - f(x) - [\mathrm{D}f]_x h.$$

L'expression

$$A = \left| g(f(x+h)) - g(f(x)) - [\mathrm{D}g]_{f(x)} [\mathrm{D}f]_x h \right|$$

est alors majorée par

$$\left| g(f(x+h)) - g(f(x)) - [\mathrm{D}g]_{f(x)} (f(x+h) - f(x)) \right| + \left| [\mathrm{D}g]_{f(x)} \varphi(h) \right|.$$

Fixons $\varepsilon > 0$. Comme l'application g est différentiable en $f(x)$, il existe $\eta > 0$ tel que $|f(x+h) - f(x)| \leq \eta$ entraîne

$$\left| g(f(x+h)) - g(f(x)) - [\mathrm{D}g]_{f(x)} (f(x+h) - f(x)) \right| \leq \varepsilon \, |f(x+h) - f(x)|.$$

Or il existe $\delta \in]0, r[$ tel que

$$|h| \leq \delta \Rightarrow |f(x+h) - f(x)| \leq \eta$$

car l'application f est continue en x. Dès lors, pour $h \in \mathbb{R}^p$ tel que $0 < |h| < \delta$, il vient

$$\frac{A}{|h|} \leq \frac{\varepsilon}{|h|} \, |f(x+h) - f(x)| + \left\| [\mathrm{D}g]_{f(x)} \right\| \frac{|\varphi(h)|}{|h|}$$

Comme on a $|\varphi(h)| / |h| \to 0$ si $h \to 0$ car f est différentiable en x, il suffit, pour conclure, de noter que

$$\frac{\varepsilon}{|h|} \, |f(x+h) - f(x)| \leq \frac{\varepsilon}{|h|} \, |\varphi(h)| + \frac{\varepsilon}{|h|} \, \| [\mathrm{D}f]_x \| \, |h| . \blacksquare$$

Théorème 5.3.2.7 (génération) *Soient f et g deux applications différentiables de l'ouvert Ω de $\mathbb{R}^p$ dans $\mathbb{R}^q$ et soit r un nombre réel.*

a) *L'application rf est différentiable et on a $\mathrm{D}(rf) = r\mathrm{D}f$.*

b) *L'application $f + g$ est différentiable et on a $\mathrm{D}(f + g) = \mathrm{D}f + \mathrm{D}g$.*

c) *Si q est égal à 1, l'application fg est différentiable et on a $\mathrm{D}(fg) = g\mathrm{D}f + f\mathrm{D}g$.*

d) *Si q est égal à 1 et si f ne s'annule en aucun point de Ω, alors $1/f$ est différentiable et on a $\mathrm{D}(1/f) = -(1/f)^2 \mathrm{D}f$.* $\blacksquare$

Chapitre 6

Fonctions élémentaires

6.1 Puissances entières

6.1.1 Définition

Définitions. Etant donné un entier $m \in \mathbb{N}$, on introduit la *puissance entière de z* de *degré m* comme étant la fonction z^m définie sur $\mathbb{C}$ par

$$z^0 = 1, \quad \forall z \in \mathbb{C}$$

et

$$z^m = z \cdot z^{m-1}, \quad \forall z \in \mathbb{C} \text{ si } m \in \mathbb{N}_0.$$

La restriction à $\mathbb{R}$ de cette fonction z^m est appelée *puissance entière de x* de *degré m* et est notée x^m. On peut aussi dire que z^m est un prolongement de x^m à $\mathbb{C}$.

Les propriétés de ces fonctions sont aisées à établir.

Théorème 6.1.1.1 a) *On a $\overline{z^m} = \overline{z}^m$ pour tout $z \in \mathbb{C}$ et tout $m \in \mathbb{N}$.*

En particulier, x^m est une fonction réelle sur $\mathbb{R}$ pour tout $m \in \mathbb{N}$. De plus, x^m est supérieur ou égal à 0 si m est pair, et a le signe de x si m est impair.

b) *On a $|z^m| = |z|^m$ pour tous $z \in \mathbb{C}$ et $m \in \mathbb{N}$.*

c) *On a $z^m z^{m'} = z^{m+m'}$ pour tous $z \in \mathbb{C}$ et $m, m' \in \mathbb{N}$.*

d) *Pour tout $m \in \mathbb{N}$, on a $z^m \in \mathrm{C}_\infty(\mathbb{R}^2)$ et, pour tout $m \in \mathbb{N}_0$,*

$$\mathrm{D}_x z^m = m z^{m-1} \quad \textit{et} \quad \mathrm{D}_y z^m = i m z^{m-1}.$$

e) *Pour tout $m \in \mathbb{N}_0$, on a $z^m \to \infty$ si $z \to \infty$.*

Plus particulièrement, on a

$$x^m \to +\infty \quad si \quad x \to \pm\infty \quad et\ si \quad m \in \mathbb{N}_0 \quad est\ pair,$$

et

$$x^m \to \pm\infty \quad si \quad x \to \pm\infty \quad et\ si \quad m \in \mathbb{N}_0 \quad est\ impair.$$ ▮

Notations. Si f est une fonction dérivable sur un ouvert Ω de $\mathbb{C}$, on introduit les opérateurs de dérivation

$$\mathrm{D}_z = \tfrac{1}{2}(\mathrm{D}_x - i\mathrm{D}_y) \quad \text{et} \quad \mathrm{D}_{\overline{z}} = \tfrac{1}{2}(\mathrm{D}_x + i\mathrm{D}_y).$$

Cela étant, les formules relatives aux dérivées de z^m donnent lieu à

$$\mathrm{D}_z z^m = mz^{m-1} = \mathrm{D}_x z^m = -i\mathrm{D}_y z^m \quad \text{et} \quad \mathrm{D}_{\overline{z}} z^m = 0$$

sur $\mathbb{C}$ pour tout $m \in \mathbb{N}$.

Remarque. Il est possible de mettre les formules de dérivation sous la forme plus symétrique suivante:

$$\mathrm{D}_z^{p<m} \frac{z^m}{m!} = \frac{z^{m-p}}{(m-p)!}, \quad \mathrm{D}_z^m \frac{z^m}{m!} = 1 \quad \text{et} \quad \mathrm{D}_z^{p>m} \frac{z^m}{m!} = 0.$$

6.1.2 Formule de Newton

Théorème 6.1.2.1 (Newton généralisé) *Pour tout entier $p \in \mathbb{N}_0$, tous nombres complexes $z_1, \ldots, z_p$ et tout $m \in \mathbb{N}_0$, on a*

$$\frac{(z_1 + \ldots + z_p)^m}{m!} = \sum_{m_1+\ldots+m_p=m} \frac{z_1^{m_1}}{m_1!} \cdots \frac{z_p^{m_p}}{m_p!}.$$

En particulier, pour tout $m \in \mathbb{N}_0$, on a la formule de Newton

$$(z_1 + z_2)^m = \sum_{k=0}^{m} \mathrm{C}_m^k z_1^k z_2^{m-k}.$$

Preuve. On peut procéder par récurrence, mais cela n'est pas aussi rapide que la méthode suivante. Si on effectue les produits dans l'expression $(z_1 + \ldots + z_p)^m/(m!)$, on obtient bien sûr une somme du type suivant

$$\sum_{m_1+\ldots+m_p=m} C_{m_1,\ldots,m_p} z_1^{m_1} \ldots z_p^{m_p}$$

et tout revient à déterminer la valeur des coefficients $C_{m_1,\ldots,m_p}$. Le calcul de la dérivée

$$\mathrm{D}_{z_1}^{m_1} \ldots \mathrm{D}_{z_p}^{m_p} \quad \text{avec} \quad m_1 + \ldots + m_p = m,$$

de ces deux expressions identiques donne aussitôt le résultat annoncé. ▮

6.2 Polynômes

6.2.1 Définition

Définitions. Un *polynôme de z* est une combinaison linéaire à coefficients complexes de puissances entières de z.

Sauf dans le cas où tous les coefficients sont nuls, il s'écrit

$$c_0 + c_1 z + \ldots + c_p z^p$$

avec $c_0, \ldots, c_p \in \mathbb{C}$ et $c_p \neq 0$; p est alors le *degré* du polynôme et $c_0, \ldots, c_p$ ses *coefficients*. Par extension, on dit que la fonction 0 est un polynôme de degré 0. Un *polynôme de x* est une combinaison linéaire à coefficients complexes de puissances entières de x. Tout polynôme de x est donc la restriction à $\mathbb{R}$ d'un polynôme de z et inversement tout polynôme de z est un prolongement à $\mathbb{C}$ d'un polynôme de x.

Définition. Si P_p est un polynôme de z, son *polynôme conjugué* est le polynôme dont les coefficients sont les conjugués des coefficients correspondants de P_p; il est noté $\overline{P_p}$ ou $(P_p)^-$. On a donc $\overline{\overline{P_p}} = P_p$ pour tout polynôme tandis que l'égalité $P_p = \overline{P_p}$ a lieu si et seulement si tous les coefficients de P_p sont réels, comme on le voit de suite.

Les propriétés générales des polynômes se déduisent aisément de celles des puissances.

Théorème 6.2.1.1 a) *Pour tout polynôme P_p et tout $z \in \mathbb{C}$, on a $\overline{P_p(z)} = \overline{P_p}(\overline{z})$.*

b) *Tout polynôme appartient à $\mathrm{C}_\infty(\mathbb{R}^2)$.*

Plus précisément, si $P_p(z)$ est un polynôme de degré p, on a

$$\mathrm{D}_z P_p = \mathrm{D}_x P_p = -i\mathrm{D}_y P_p;$$

toute dérivée de P_p est un polynôme et, pour $q > p$, on a même $\mathrm{D}_z^q P_p = 0$.

c) *Pour tout polynôme P_p de degré $p \geq 1$, on a*

$$\lim_{z \to \infty} P_p(z) = \infty.$$

Preuve. a) est immédiat.

b) découle de suite des propriétés des dérivées des combinaisons linéaires.

c) De fait, pour tout $z \neq 0$, on a

$$P_p(z) = z^p\left(c_p + \frac{c_{p-1}}{z} + \ldots + \frac{c_0}{z^p}\right)$$

où $c_p + c_{p-1}/z + \ldots + c_0/z^p$ tend vers c_p si $z \to \infty$. ∎

6.2.2 Identité de Taylor

Théorème 6.2.2.1 (identité de Taylor) *Pour tout polynôme P_p de degré $p \geq 1$ et tout nombre complexe a, on a*

$$P_p(z) = \sum_{q=0}^{p} \frac{1}{q!} \mathrm{D}_z^q P_p(a) \cdot (z-a)^q.$$

Preuve. De

$$P_p(z) = P_p((z-a)+a) = \sum_{q=0}^{p} c_q((z-a)+a)^q, \quad \forall z \in \mathbb{C},$$

on déduit de suite qu'il existe des nombres complexes $d_0, \ldots, d_p$ tels que

$$P_p(z) = \sum_{q=0}^{p} \frac{1}{q!} d_q (z-a)^q$$

et tout revient à déterminer ces coefficients d_q. Or, pour tout $j \leq p$, on a

$$\mathrm{D}_z^j P_p(z) = \sum_{q=j}^{p} \frac{1}{(q-j)!} d_q (z-a)^{q-j}, \quad \forall z \in \mathbb{C},$$

d'où on tire de suite la valeur de d_q en calculant la valeur des deux membres en a.∎

6.2.3 Etude des zéros

Définitions. Un zéro a d'un polynôme $P_p(z)$ est *α-uple* si on a

$$P_p(a) = \mathrm{D}_z P_p(a) = \ldots = \mathrm{D}_z^{\alpha-1} P_p(a) = 0 \text{ et } \mathrm{D}_z^{\alpha} P_p(a) \neq 0.$$

On dit aussi que a est un zéro de *multiplicité* α. Ce nombre α est donc un entier supérieur ou égal à 1. On vérifie de suite que cette notion n'a de sens que si le degré p du polynôme P_p est supérieur ou égal à 1.

Proposition 6.2.3.1 *Si a est un zéro α-uple du polynôme P_p, $\overline{a}$ est un zéro α-uple du polynôme $\overline{P_p}$. En particulier, si P_p est un polynôme à coefficients réels et si a est un zéro α-uple de P_p, alors $\overline{a}$ est aussi un zéro α-uple de P_p.*

Preuve. Cela résulte aussitôt de l'identité

$$\overline{\mathrm{D}^q P_p}(\overline{z}) = \overline{\mathrm{D}^q P_p(z)},$$

valable pour tout $q \in \mathbb{N}$ et tout $z \in \mathbb{C}$.∎

Proposition 6.2.3.2 *Le nombre $a \in \mathbb{C}$ est un zéro α-uple du polynôme P_p de degré p si et seulement s'il existe un polynôme $P_{p-\alpha}$ de degré $p-\alpha$ tel que*

$$P_{p-\alpha}(a) \neq 0 \quad et \quad P_p(z) = (z-a)^\alpha P_{p-\alpha}(z), \quad \forall z \in \mathbb{C}.$$

De plus, les zéros de P_p différents de a et les zéros de $P_{p-\alpha}$ sont égaux, avec même multiplicité.

Preuve. La nécessité de la condition résulte aussitôt de l'identité de Taylor exprimée en a.

La condition est suffisante. Pour tout $q \leq p$, la formule de Leibniz donne

$$\mathrm{D}_z^q P_p(z) = \sum_{k=0}^{q} C_q^k \mathrm{D}_z^k (z-a)^\alpha \cdot \mathrm{D}_z^{q-k} P_{p-\alpha}(z).$$

On en tire de suite que a est un zéro α-uple de P_p.

Ces mêmes formules permettent immédiatement d'affirmer la dernière partie de l'énoncé.∎

Lemme 6.2.3.3 (Gauss) *Soient P_p un polynôme de degré $p \geq 1$ et Ω un ouvert de $\mathbb{C}$. Si $a \in \Omega$ est tel que $|P_p(a)| = \inf_{z\in\Omega} |P_p(z)|$, alors a est un zéro de P_p.*

Preuve. Vu l'identité de Taylor, il existe un entier $k \geq 1$, un polynôme Q et un nombre complexe b non nul tels que

$$P_p(z) = P_p(a) + b(z-a)^k + (z-a)^{k+1} Q(z).$$

Si $P_p(a)$ diffère de 0, soit z_0 un nombre complexe tel que $z_0^k = -P_p(a)/b$. Comme Q est continu en a, il existe $t \in \,]0,1[$ tel que $a + tz_0 \in \Omega$ et

$$t \cdot \left| z_0^{k+1} Q(a+tz_0) \right| < |P_p(a)|.$$

Dès lors, de

$$P_p(a+tz_0) = P_p(a) - t^k P_p(a) + t^{k+1} z_0^{k+1} Q(a+tz_0),$$

on tire

$$|P_p(a+tz_0)| < (1-t^k)\,|P_p(a)| + t^k\,|P_p(a)| = |P_p(a)|,$$

ce qui est impossible.∎

Théorème 6.2.3.4 (fondamental) *Tout polynôme de degré $p \geq 1$ a p zéros si on compte leurs multiplicités.*

Plus précisément, si P_p est un polynôme de degré $p \geq 1$ et si $a_1, \ldots, a_q$ sont ses zéros de multiplicités respectives $\alpha_1, \ldots, \alpha_q$, on a $\alpha_1 + \ldots + \alpha_q = p$ et

$$P_p(z) = c_p (z-a_1)^{\alpha_1} \ldots (z-a_q)^{\alpha_q}, \quad \forall z \in \mathbb{C}.$$

Preuve. a) Etablissons tout d'abord que P_p admet un zéro.

Comme on a $P_p(z) \to \infty$ si $z \to \infty$, il existe $R > 0$ tel que

$$|z| \geq R \Rightarrow |P_p(z)| \geq |P_p(0)| + 1. \quad (*)$$

Vu le théorème des bornes atteintes, il existe alors un point z_0 de la boule compacte $\{ z : |z| \leq R\}$ tel que

$$|P_p(z_0)| = \inf_{|z|\leq R} |P_p(z)|$$

et, vu la relation (*), on doit avoir $|z_0| < R$. Vu le lemme précédent, z_0 est un zéro de P_p.

b) Si a_1 est un zéro du polynôme P_p, sa multiplicité α_1 est inférieure ou égale à p puisqu'on a $\mathrm{D}^p P_p = c_p p!$. Il existe donc $\alpha_1 \in \{1, \ldots, p\}$ ainsi qu'un polynôme $P_{p-\alpha_1}$ de degré $p - \alpha_1$ tels que

$$P_p(z) = (z - a_1)^{\alpha_1} P_{p-\alpha_1}(z), \quad \forall z \in \mathbb{C}.$$

Si α_1 est égal à p, $P_{p-\alpha_1}$ est forcément égal à c_p. Sinon le degré du polynôme $P_{p-\alpha_1}$ est supérieur ou égal à 1. Cela étant, vu la partie a) de cette démonstration, il admet au moins un zéro a_2 nécessairement différent de a_1 et, si α_2 est sa multiplicité, il vient $\alpha_2 \in \{1, \ldots, p - \alpha_1\}$ et il existe un polynôme $P_{p-\alpha_1-\alpha_2}$ tel que

$$P_p(z) = (z - a_1)^{\alpha_1}(z - a_2)^{\alpha_2} P_{p-\alpha_1-\alpha_2}(z), \quad \forall z \in \mathbb{C}.$$

En continuant de la sorte, on finit par obtenir des nombres $a_1, \ldots, a_q \in \mathbb{C}$ et des entiers $\alpha_1, \ldots, \alpha_q \geq 1$ tels que $\alpha_1 + \ldots + \alpha_q = p$ et

$$P_p(z) = (z - a_1)^{\alpha_1} \ldots (z - a_q)^{\alpha_q} P_0, \quad \forall z \in \mathbb{C},$$

où P_0 est un polynôme de degré 0, nécessairement égal à c_p.

c) On contrôle alors de suite que $a_1, \ldots, a_q$ sont des zéros de P_p de multiplicités respectives $\alpha_1, \ldots, \alpha_q$ et que P_p n'a aucun autre zéro.∎

Remarque. Si P_0 est un polynôme de degré 0, c'est une constante c_0. Si c_0 est égal à 0, tout point a de $\mathbb{C}$ est un zéro de P_0; si c_0 diffère de 0, P_0 n'a pas de zéro.

Corollaire 6.2.3.5 *Si deux polynômes P_p et P'_p de même degré $p \geq 1$ ont mêmes zéros avec mêmes multiplicités respectives, il existe un nombre complexe $c \neq 0$ tel que $P_p = cP'_p$.*

Preuve. De fait, si ces zéros sont les $a_1, \ldots, a_q$ de multiplicités respectives $\alpha_1, \ldots, \alpha_q$, on a

$$\begin{aligned} P_p(z) &= c_p(z - a_1)^{\alpha_1} \ldots (z - a_q)^{\alpha_q} \\ &= \frac{c_p}{c'_p} \cdot c'_p(z - a_1)^{\alpha_1} \ldots (z - a_q)^{\alpha_q} = \frac{c_p}{c'_p} \cdot P'_p(z). \blacksquare \end{aligned}$$

Corollaire 6.2.3.6 *Si deux polynômes P_p et P'_p de degrés $\leq p$ sont égaux en $p+1$ points distincts, ils sont égaux.*

Preuve. De fait, $P_p - P'_p$ est alors un polynôme de degré inférieur ou égal à p qui admet au moins $p+1$ zéros distincts; vu ce qui précède, $P_p - P'_p$ est nécessairement le polynôme 0. La conclusion s'ensuit aussitôt.∎

Remarque. Insistons sur le fait que nous avons établi que tout polynôme de degré $p \geq 1$ a p zéros si on compte ces zéros avec leurs multiplicités, sans donner de méthode de calcul explicite de ces zéros. En fait, de telles méthodes sont connues pour les polynômes de degré 1, 2, 3 ou 4. Par contre, pour obtenir les zéros des polynômes de degré $p \geq 5$, on recourt, en général, à des méthodes de calcul numérique qui donnent ces zéros comme limites de suites.

6.2.4 Division des polynômes

Théorème 6.2.4.1 *Si D est un polynôme de degré $q \geq 1$ et P un polynôme de degré $p \geq q$, il existe des polynômes Q de degré $p-q$ et R de degré strictement inférieur à q tels que $P = QD + R$. De plus, ces polynômes Q et R sont uniques.*

Preuve. a) Etablissons d'abord l'existence de tels polynômes Q et R.

Bien sûr, pour tout polynôme P' de degré $p' \geq q$, il existe $c \in \mathbb{C}$ tel que le polynôme $P'(z) - cz^{p'-q}D(z)$ soit un polynôme de degré $p'-1$ au plus. Cela étant, il existe $c_{p-q} \in \mathbb{C}$ tel que

$$P_{p-1}(z) = P(z) - c_{p-q}z^{p-q}D(z)$$

soit un polynôme de degré $p-1$ au plus, puis successivement des nombres c_{p-q-1}, $\ldots$, $c_0 \in \mathbb{C}$ tels que

$$\begin{gathered} P_{p-2}(z) = P_{p-1}(z) - c_{p-q-1}z^{p-q-1}D(z) \\ \cdots \\ P_{q-1} = P_q - c_0 D, \end{gathered}$$

où P_{p-2}, $\ldots$, P_{q-1} sont des polynômes de degrés majorés par $p-2$, $\ldots$, $q-1$ respectivement. D'où la conclusion car au total on a

$$P(z) = (c_0 + c_1 z + \ldots + c_{p-q}z^{p-q}) \cdot D(z) + P_{q-1}(z).$$

b) Prouvons l'unicité de Q et R.

De fait, si les polynômes Q' et R' de degrés $p-q$ et strictement inférieur à q respectivement sont tels que $P = Q'D + R'$, on obtient $(Q - Q')D = R' - R$. Cela étant, on doit avoir $Q = Q'$ sinon le degré du premier membre est strictement supérieur à celui du second. On en déduit de suite l'égalité $R = R'$.∎

Définitions. Les polynômes Q et R de l'énoncé précédent sont appelés respectivement *quotient* et *reste* de la division de P par D.

De plus, P est *divisible par* D si on a $R = 0$.

Remarque. Si $a_1, \ldots, a_q$ appartiennent à $\mathbb{C}$ et diffèrent, si $\alpha_1, \ldots, \alpha_q$ appartiennent à $\mathbb{N}_0$ et si le polynôme P est divisible par $(z-a_1)^{\alpha_1}, \ldots$ et $(z-a_q)^{\alpha_q}$, alors P est divisible par le produit $(z-a_1)^{\alpha_1} \ldots (z-a_q)^{\alpha_q}$. De fait, pour tout $j \leq q$, il existe un polynôme Q_j tel que $P(z) = (z-a_j)^{\alpha_j} Q_j(z)$ car P est divisible par $(z-a_j)^{\alpha_j}$. De là, a_j est un zéro de P, de multiplicité au moins égale à α_j et la conclusion s'ensuit aussitôt.

6.2.5 Plus grand commun diviseur

Définition. Etant donné des polynômes non identiquement nuls $P_1, \ldots, P_J$ en nombre fini, on appelle *plus grand commun diviseur* (en abrégé PGCD) de $P_1, \ldots, P_J$ tout polynôme dont les zéros a_j sont les zéros communs de $P_1, \ldots, P_J$, la multiplicité du zéro a_j étant la borne inférieure des multiplicités de a_j comme zéro de $P_1, \ldots, P_J$.

Vu un corollaire du paragraphe 6.2.3, le PGCD d'un nombre fini de polynômes $P_1, \ldots, P_J$ n'est pas unique, il n'est défini qu'à une constante multiplicative non nulle près.

Lemme 6.2.5.1 (Euler) *Si $P_1, \ldots, P_J$ sont des polynômes non identiquement nuls et en nombre fini et si D est un plus grand commun diviseur de ces polynômes, il existe des polynômes $Q_1, \ldots, Q_J$ tels que $P_1Q_1 + \ldots + P_JQ_J = D$.*

Preuve. Pour tous polynômes $Q_1, \ldots, Q_J$, $P_1Q_1 + \ldots + P_JQ_J$ est un polynôme et il existe bien sûr des polynômes $Q_1, \ldots, Q_J$ tels que cette somme diffère de 0. Il existe donc des polynômes $Q_1^0, \ldots, Q_J^0$ tels que le polynôme $d = P_1Q_1^0 + \ldots + P_JQ_J^0$ diffère de 0 et soit de degré minimum.

a) Cela étant, établissons d'abord que, pour tous polynômes $Q_1, \ldots, Q_J$, le polynôme $P_1Q_1 + \ldots + P_JQ_J$ est divisible par d.

Sinon il existe des polynômes $Q_1, \ldots, Q_J$ tels que $P_1Q_1 + \ldots + P_JQ_J$ ne soit pas divisible par d. Comme son degré est nécessairement supérieur ou égal à celui de d, il existe un polynôme Q et un polynôme R non nul et de degré strictement inférieur à celui de d tels que

$$P_1Q_1 + \ldots + P_JQ_J = Qd + R.$$

D'où une contradiction car on a alors

$$R = P_1 \cdot (Q_1 - QQ_1^0) + \ldots + P_J \cdot (Q_J - QQ_J^0).$$

b) Prouvons à présent que d est un PGCD de $P_1, \ldots, P_J$. D'une part, vu sa définition, d est évidemment divisible par D. D'autre part, D est divisible par d car vu a), chacun des P_j est divisible par d puisqu'il s'écrit

$$P_j = P_1 \cdot 0 + \ldots + P_{j-1} \cdot 0 + P_j \cdot 1 + P_{j+1} \cdot 0 + \ldots + P_J \cdot 0.$$

Il s'ensuit que d et D ont mêmes zéros, de mêmes multiplicités, ce qui suffit.

D'où la conclusion.∎

Remarque. Les polynômes $Q_1, \ldots, Q_J$ du lemme d'Euler ne sont pas uniques car on a évidemment

$$P_1 \cdot (Q_1 + c_1 P_2 \ldots P_J) + \ldots + P_J \cdot (Q_J + c_J P_1 \ldots P_{J-1}) = d$$

si $c_1, \ldots, c_J \in \mathbb{C}$ sont tels que $c_1 + \ldots + c_J = 0$.

6.3 Fractions rationnelles

6.3.1 Définition

Définition. Une *fraction rationnelle de z* est un quotient de deux polynômes de z, le degré du dénominateur étant supérieur ou égal à 1. Bien sûr, son ensemble de définition est le complémentaire dans $\mathbb{C}$ de l'ensemble des zéros du dénominateur.

Remarque. Si $R = A/B$ est une fraction rationnelle, nous savons qu'il existe des polynômes uniques Q et R' tels que $A = QB + R'$, le degré de R' étant strictement inférieur à celui de B; on a donc $R(z) = Q(z) + R'(z)/B(z)$ pour tout $z \in \mathbb{C}$ appartenant à l'ensemble de définition de R. Ici, R'/B est encore une fraction rationnelle, mais le degré de R' est strictement inférieur à celui de B. De plus, si D est un plus grand commun diviseur de R' et de B, il existe des polynômes R'' et B' tels que $R' = DR''$ et $B = DB'$, donc tels que $R(z) = Q(z) + R''(z)/B'(z)$ pour tout $z \in \mathbb{C}$ appartenant à l'ensemble de définition de R. Ici, R''/B' est une fraction rationnelle où le degré de R'' est strictement inférieur à celui de B' et où R'' et B' n'ont pas de zéro commun. Cela étant, pour étudier R, il suffit d'étudier R''/B car nous connaissons déjà les propriétés du polynôme Q.

Définition. Une fraction rationnelle $R = A/B$ est *propre* si le degré de A est strictement inférieur à celui de B et si A et B n'ont pas de zéro commun.

La remarque précédente ramène alors l'étude des fractions rationnelles à celle des polynômes (déjà effectuée) et à celle des fractions rationnelles propres.

Remarque. Soient A/B et A'/B' deux fractions rationnelles propres. Si ces deux fonctions sont égales sur l'intersection de leurs ensembles de définition, les polynômes AB' et $A'B$ prennent les mêmes valeurs en une suite de points distincts de $\mathbb{C}$ et sont donc identiques. On en déduit de suite que B et B' ont mêmes zéros de mêmes multiplicités. Il existe donc un nombre complexe c non nul tel que $B = cB'$ et, par conséquent, aussi tel que $A = cA'$. Dès lors, si $R = A/B$ est une fraction rationnelle propre, A et B sont uniques à une constante multiplicative près.

Cela étant, on peut introduire les définitions suivantes.

Définition. On appelle *pôle* d'une fraction rationnelle propre tout zéro de son dénominateur, l'*ordre* d'un pôle étant la multiplicité de ce zéro du dénominateur.

Théorème 6.3.1.1 *Soit $R = A/B$ une fraction rationnelle propre et soit $\mathcal{P}$ l'ensemble de ses pôles. Alors*

a) *R est une fonction définie sur $\mathbb{C} \setminus \mathcal{P}$,*

b) *on a*

$$\overline{R(z)} = \frac{\overline{A}(\overline{z})}{\overline{B}(\overline{z})} \quad et \qquad |R(z)| = \frac{|A(z)|}{|B(z)|}, \quad \forall z \in \mathbb{C} \setminus \mathcal{P},$$

c) *R appartient à $\mathrm{C}_\infty(\mathbb{C} \setminus \mathcal{P})$ et est tel que*

$$\mathrm{D}_z R(z) = \mathrm{D}_x R(z) = -i\mathrm{D}_y R(z), \quad \forall z \in \mathbb{C} \setminus \mathcal{P}.$$

De plus, $\mathrm{D}_z R$ est une fraction rationnelle propre qui a les mêmes pôles que R, l'ordre de chaque pôle de $\mathrm{D}_z R$ étant égal à l'ordre de ce pôle de R augmenté de 1.

d) *si b est un pôle d'ordre β de R, on a*

$$\lim_{z \to b} (z-b)^\beta R(z) = \beta! \frac{A(b)}{\mathrm{D}_z^\beta B(b)}$$

Preuve. a) et b) sont immédiats

c) Bien sûr, R appartient à $\mathrm{C}_\infty(\mathbb{C} \setminus \mathcal{P})$, comme quotient de deux fonctions appartenant à $\mathrm{C}_\infty(\mathbb{C} \setminus \mathcal{P})$, le dénominateur ne s'annulant pas sur $\mathbb{C} \setminus \mathcal{P}$. De plus, on a alors

$$\mathrm{D}_z \frac{A(z)}{B(z)} = \frac{B(z) \cdot \mathrm{D}_z A(z) - A(z) \cdot \mathrm{D}_z B(z)}{(B(z))^2}, \quad \forall z \in \mathbb{C} \setminus \mathcal{P},$$

et ainsi $\mathrm{D}_z R$ apparaît comme étant une fraction rationnelle. Mais sous cette forme, elle peut ne pas être propre: bien sûr, le degré du numérateur est strictement inférieur à celui du dénominateur mais le numérateur et le dénominateur peuvent avoir un zéro commun. Cela ne peut cependant avoir lieu que si $B(z)$ et $\mathrm{D}_z B(z)$ ont un zéro commun, c'est-à-dire si et seulement si $B(z)$ admet un zéro b de multiplicité $\beta \geq 2$. Dans ce cas, il existe un polynôme $B'(z)$ vérifiant $B'(b) \neq 0$ et tel que $B(z) = (z-b)^\beta B'(z)$ pour tout $z \in \mathbb{C}$; on obtient alors

$$\mathrm{D}_z \frac{A(z)}{B(z)} = \frac{-\beta A(z)B'(z) + (z-b)(B'(z)\mathrm{D}_z A(z) - A(z)\mathrm{D}_z B'(z))}{(z-b)^{\beta+1}(B'(z))^2}$$

pour tout $z \in \mathbb{C} \setminus \mathcal{P}$. On en déduit de suite la dernière partie de l'énoncé.

d) De fait, si B' est le polynôme vérifiant $B(z) = (z-b)^\beta B'(z)$ pour tout $z \in \mathbb{C}$, la formule de Leibniz donne

$$\mathrm{D}_z^\beta B(z) = \beta! B'(z) + \sum_{k=0}^{\beta-1} C_\beta^k \mathrm{D}_z^k (z-b)^\beta \cdot \mathrm{D}_z^{\beta-k} B'(z).$$

La conclusion est alors immédiate.∎

Remarque. On en déduit directement de nombreuses propriétés des fractions rationnelles.

Proposition 6.3.1.2 *Si $R = A/B$ est une fraction rationnelle et si p est le degré de A et q celui de B, on a*

$$\lim_{z\to\infty} R(z) = \lim_{z\to\infty} \frac{c_p}{d_q} z^{p-q} = \left\{ \begin{array}{cl} \infty & si\ p > q \\ c_p/d_q & si\ p = q \\ 0 & si\ p < q \end{array} \right\}$$

où c_p est le coefficient de z^p dans A et d_q celui de z^q dans B.

Preuve. De fait, on a successivement

$$\lim_{z\to\infty} \frac{A(z)}{B(z)} = \lim_{z\to\infty} \frac{A(z)}{z^p} \frac{z^q}{B(z)} z^{p-q} = \frac{c_p}{d_q} \lim_{z\to\infty} z^{p-q}.∎$$

6.3.2 Décomposition d'une fraction rationnelle propre

Théorème 6.3.2.1 (fondamental) *Soit A/B une fraction rationnelle propre. Si on a $B = B_1 \cdots B_J$, les $B_1, \ldots, B_J$ étant des polynômes de degré ≥ 1 qui, deux à deux n'ont pas de zéro commun, alors il existe des polynômes $A_1, \ldots, A_J$ uniques tels que*

$$\frac{A}{B} = \frac{A_1}{B_1} + \cdots + \frac{A_J}{B_J}$$

les A_j/B_j, étant des fractions rationnelles propres.

De plus, pour tout $j \leq J$, A_j ne dépend que de A, B et B_j.

Enfin, si les coefficients de A et de B sont tous réels et si on a $B_j = \overline{B_k}$, alors on a $A_j = \overline{A_k}$. En particulier, dans ce cas, A_j a tous ses coefficients réels s'il en est de même pour B_j.

Preuve. Etablissons d'abord que 1 est un PGCD des différents polynômes $B_1 \ldots [B_j] \ldots B_J$ pour $j = 1, \ldots, J$ où le facteur entre crochets est omis. De fait, pour tout $j \leq J$, si a est un zéro de $B_1 \ldots [B_j] \ldots B_J$, c'est un zéro d'un et d'un seul des facteurs, B_k par exemple, et alors a n'est pas zéro de $B_1 \ldots [B_k] \ldots B_J$, ce qui suffit.

Vu le lemme d'Euler, il existe alors des polynômes $A'_1, \ldots, A'_J$ tels que

$$1 = \sum_{j=1}^{J} A'_j \cdot B_1 \cdots [B_j] \cdots B_J. \quad (*)$$

Par division par B puis multiplication par A, on obtient de suite l'égalité

$$\frac{A}{B} = \sum_{j=1}^{J} \frac{AA'_j}{B_j}.$$

Si le degré de AA'_j est supérieur ou égal à celui de B_j, effectuons la division: pour tout $j \leq J$, il existe des polynômes Q_j et A_j tels que $AA'_j = Q_jB_j + A_j$, le degré de A_j étant strictement inférieur à celui de B_j. Au total, nous obtenons

$$\frac{A}{B} = \sum_{j=1}^{J} Q_j + \sum_{j=1}^{J} \frac{A_j}{B_j}.$$

D'une part, on a $\sum_{j=1}^{J} Q_j = 0$ car sinon, après multiplication par B, il vient

$$A = \sum_{j=1}^{J} Q_jB + \sum_{j=1}^{J} A_jB_1 \cdots [B_j] \cdots B_J,$$

et le degré de A devrait être supérieur ou égal à celui de B.

D'autre part, chaque fraction rationnelle A_j/B_j est propre. De fait, le degré de A_j est strictement inférieur à celui de B_j, par construction. De plus, A_j et B_j n'ont pas de zéro commun: de fait, un tel zéro a devrait être un zéro de A'_j vu la relation $AA'_j = Q_jB_j + A_j$ et alors 1 serait divisible par $z - a$, vu la relation (*).

De la sorte

$$\frac{A}{B} = \sum_{j=1}^{J} \frac{A_j}{B_j}$$

est une décomposition du type indiqué.

Etablissons son unicité et même le fait que A_j ne dépend que de A, B et B_j. De fait, si B s'écrit aussi $B = B_j \cdot B'_1 \cdots B'_K$ où les B_j, $B'_1, \ldots, B'_K$ sont des polynômes de degré ≥ 1 qui, deux à deux, n'ont pas de zéro commun et si les polynômes A_j^* et $A'_1, \ldots, A'_K$ sont tels que

$$\frac{A}{B} = \frac{A_j^*}{B_j} + \sum_{k=1}^{K} \frac{A'_k}{B'_k}$$

où A_j^*/B_j et les A'_k/B'_k sont des fractions rationnelles propres, on obtient de suite

$$(A_j - A_j^*)B_1 \cdots [B_j] \cdots B_J B'_1 \cdots B'_K = B_j Q$$

où Q est un polynôme. Cela étant, $A_j - A_j^*$ est égal à 0 car il s'agit d'un polynôme de degré strictement inférieur à celui de B_j alors qu'il doit être divisible par B_j.

Enfin, supposons que tous les coefficients de A et de B soient réels et que B_j soit égal à $\overline{B_k}$ (le cas $j = k$ étant exclu). De la décomposition

$$\frac{A}{B} = \frac{A_1}{B_1} + \cdots + \frac{A_j}{B_j} + \cdots + \frac{A_k}{B_k} + \cdots + \frac{A_J}{B_J},$$

on tire de suite, en conjugant les deux membres et en revenant à z,

$$\frac{A}{B} = \frac{\overline{A_1}}{\overline{B_1}} + \cdots + \frac{\overline{A_j}}{B_k} + \cdots + \frac{\overline{A_k}}{B_j} + \cdots + \frac{\overline{A_J}}{\overline{B_J}},$$

c'est-à-dire $\overline{A_k} = A_j$, vu l'unicité de la décomposition.

Si tous les coefficients de A, de B et de B_j sont réels, de

$$\frac{A}{B} = \frac{A_1}{B_1} + \cdots + \frac{A_j}{B_j} + \cdots + \frac{A_J}{B_J},$$

on tire de suite, en conjuguant les deux membres et en revenant à z,

$$\frac{A}{B} = \frac{\overline{A_1}}{\overline{B_1}} + \cdots + \frac{\overline{A_j}}{B_j} + \cdots + \frac{\overline{A_J}}{\overline{B_J}},$$

c'est-à-dire $A_j = \overline{A_j}$, vu l'unicité de la décomposition.∎

6.3.3 Cas pratiques de décomposition

En pratique, on applique le théorème fondamental de décomposition des fractions rationnelles propres pour des factorisations de B bien particulières.

a) On peut prendre pour $B_1, \ldots, B_J$ les polynômes $(z - b)^\beta$ où b est un zéro de multiplicité β dans B. On dit qu'on *décompose A/B en fractions rationnelles simples.*

Dans ce cas, le théorème fondamental donne le résultat suivant: si b est un zéro de multiplicité β dans B, il lui correspond un polynôme unique $A_{\beta-1}$ de degré $\beta - 1$ au plus dans la décomposition de A/B en fractions rationnelles propres. De plus, si tous les coefficients de A et de B sont réels et si b est réel, alors les coefficients du polynôme $A_{\beta-1}$ sont réels.

Cependant, on préfère préciser ce résultat de la manière suivante.

En recourant à l'identité de Taylor en b relative au polynôme $A_{\beta-1}$, on obtient de suite

$$\frac{A_{\beta-1}(z)}{(z-b)^\beta} = \sum_{q=0}^{\beta-1} \frac{1}{q!} \frac{[\mathrm{D}_z^q A_{\beta-1}]_b}{(z-b)^{\beta-q}}.$$

Par conséquent, *à tout zéro b de multiplicité β du dénominateur de la fraction rationnelle propre A/B, correspondent les termes univoquement définis*

$$\frac{c_1}{z-b} + \cdots + \frac{c_\beta}{(z-b)^\beta}, \quad (c_1, \ldots, c_\beta \in \mathbb{C}),$$

dans la décomposition de A/B en fractions rationnelles simples.

De plus, si les coefficients de A et de B sont réels et si b est réel, les c_1, ..., c_β sont des nombres réels.

b) Si tous les coefficients de B sont réels, on peut prendre pour B_1, ..., B_J des polynômes d'un des types:

i) $(z-b)^\beta$ si $b \in \mathbb{R}$ est un zéro de multiplicité β de B;

ii) $(z-b)^\beta(z-\overline{b})^\beta$ si $b \in \mathbb{C} \setminus \mathbb{R}$ est un zéro de multiplicité β de B car on sait alors que $\overline{b}$ est aussi un zéro de multiplicité β de B.

Si $b \in \mathbb{R}$ est un zéro de multiplicité β de B, on applique a).

Si $b \in \mathbb{C} \setminus \mathbb{R}$ est un zéro de multiplicité β de B, on remarque que l'on a

$$(z-b)^\beta(z-\overline{b})^\beta = (z^2 - 2z \cdot \Re b + |b|^2)^\beta$$

et dès lors, le théorème fondamental de décomposition affirme qu'il lui correspond un polynôme unique $A_{2\beta-1}$ de degré $2\beta - 1$ au plus dans la décomposition de A/B en fractions rationnelles propres. De plus, si tous les coefficients de A et de B sont réels, alors tous les coefficients de $A_{2\beta-1}$ sont réels.

Cependant on préfère préciser ce résultat de la manière suivante.

Posons

$$Q(z) = (z-b)(z-\overline{b}).$$

En effectuant successivement les divisions suivantes, nous faisons apparaître des polynômes A_α de degré inférieur ou égal à α et des polynômes L_j de degré strictement inférieur à 2

$$\begin{array}{rcl} A_{2\beta-1} & = & Q^{\beta-1}L_1 + A_{2\beta-3} \\ A_{2\beta-3} & = & Q^{\beta-2}L_2 + A_{2\beta-5} \\ & \cdots & \\ A_3 & = & QL_{\beta-1} + A_1 \;=\; QL_{\beta-1} + L_\beta. \end{array}$$

En additionnant ces relations, nous obtenons

$$A_{2\beta-1} = L_1Q^{\beta-1} + L_2Q^{\beta-2} + \cdots + L_{\beta-1}Q + L_\beta.$$

Par conséquent, *à tout zéro $b \in \mathbb{C} \setminus \mathbb{R}$ de multiplicité β du dénominateur à coefficients réels de la fraction rationnelle propre A/B, correspondent les termes univoquement définis*

$$\frac{L_1}{Q} + \cdots + \frac{L_\beta}{Q^\beta},$$

les $L_1, \ldots, L_\beta$ étant des polynômes de degré 1 au plus, dans la décomposition de A/B en fractions rationnelles propres.

De plus, si les coefficients de A et de B sont réels, ceux des $L_1, \ldots, L_\beta$ le sont aussi.

6.3.4 Calcul effectif d'une décomposition

Soit A/B une fraction rationnelle propre et soit $B = B_1 \cdots B_J$ une factorisation de B en polynômes de degrés supérieurs ou égaux à 1 qui, deux à deux, n'ont pas de zéro commun.

Pour obtenir les polynômes $A_1, \ldots, A_J$ de la décomposition unique

$$\frac{A}{B} = \frac{A_1}{B_1} + \cdots + \frac{A_J}{B_J},$$

on écrit la forme générale du développement en recourant à des coefficients indéterminés pour les $A_1, \ldots, A_J$. Ensuite on multiplie les deux membres par B et on obtient ainsi une identité polynomiale. En identifiant les coefficients des termes de même puissance, on obtient un système de q équations linéaires à q inconnues si q est le degré de B.

Il est cependant possible de diminuer sensiblement la difficulté de résolution de ce système dans les deux cas considérés au paragraphe précédent.

Si b est un zéro de multiplicité β du dénominateur de la fraction rationnelle propre A/B, les coefficients $c_1, \ldots, c_\beta$ de la décomposition

$$\frac{A(z)}{B(z)} = \frac{c_1}{z-b} + \cdots + \frac{c_\beta}{(z-b)^\beta} + R(z)$$

sont solutions du système triangulaire d'équations linéaires

$$\begin{cases} A(b) & = & N(b) \\ [\mathrm{D}_z A]_b & = & [\mathrm{D}_z N]_b \\ & \cdots & \\ [\mathrm{D}_z^{\beta-1} A]_b & = & [\mathrm{D}_z^{\beta-1} N]_b \end{cases}$$

de dimension β *où* N *est le polynôme égal au second membre multiplié par* B.

De même, *si* $b \in \mathbb{C}\backslash\mathbb{R}$ *et* $\overline{b}$ *sont des zéros de même multiplicité* β *du dénominateur de la fraction rationnelle propre* A/B*, les coefficients des polynômes* $L_1, \ldots, L_\beta$ *de degrés inférieurs ou égaux à* 1 *de la décomposition*

$$\frac{A(z)}{B(z)} = \frac{L_1}{Q} + \cdots + \frac{L_\beta}{Q^\beta} + R(z)$$

sont solutions du système linéaire

$$\begin{cases} A(b) & = & N(b) & \qquad A(\overline{b}) & = & N(\overline{b}) \\ [\mathrm{D}_z A]_b & = & [\mathrm{D}_z N]_b & \qquad [\mathrm{D}_z A]_{\overline{b}} & = & [\mathrm{D}_z N]_{\overline{b}} \\ & \cdots & & & \cdots & \\ [\mathrm{D}_z^{\beta-1} A]_b & = & [\mathrm{D}_z^{\beta-1} N]_b & \qquad [\mathrm{D}_z^{\beta-1} A]_{\overline{b}} & = & [\mathrm{D}_z^{\beta-1} N]_{\overline{b}} \end{cases}$$

de dimension 2β *où* N *est le polynôme égal au second membre multiplié par* B.

Justifions ces règles pratiques.

Dans le premier cas, on sait qu'il existe un polynôme $A_{\beta-1}$ de degré $\beta - 1$ au plus tel que

$$\frac{A(z)}{B(z)} = \frac{A_{\beta-1}(z)}{(z-b)^\beta} + R(z)$$

et ainsi l'identité $A(z) = N(z)$ a la forme

$$A(z) = A_{\beta-1}(z)Q(z) + (z-b)^\beta Q'(z)$$

où Q et Q' sont deux polynômes, b n'étant pas zéro de Q. Par conséquent,

$$N(b),\ [\mathrm{D}_z N]_b,\ \ldots,\ [\mathrm{D}_z^{\beta-1} N]_b$$

ne font intervenir que les coefficients de $A_{\beta-1}$, ce qui établit que les $c_1, \ldots, c_\beta$ sont bien solutions du système indiqué. De plus, il s'agit d'un système triangulaire. En effet, de

$$\frac{A_{\beta-1}(z)}{(z-b)^\beta} = \frac{c_1}{z-b} + \cdots + \frac{c_\beta}{(z-b)^\beta},$$

on tire que l'identité $A(z) = N(z)$ s'écrit

$$A(z) = \left(c_\beta + c_{\beta-1}(z-b) + \cdots + c_1(z-b)^{\beta-1}\right) Q(z) + (z-b)^\beta Q'(z),$$

ce qui suffit.

Dans le second cas, on procède de même.

Exemple. *Décomposer la fraction $(z-1)^{-2}(z^3+1)^{-2}$ en fractions rationnelles propres.*

Suggestion. La décomposition a la forme

$$\frac{a}{z-1}+\frac{b}{(z-1)^2}+\frac{cz+d}{z^2+1}+\frac{ez+f}{(z^2+1)^2}$$

où a, b, c, d, e et f sont des nombres réels. Si on procède directement à l'identification des coefficients, on obtient

$$1=(a(z-1)+b)\,(z^2+1)^2+\big((cz+d)(z^2+1)+ez+f\big)\,(z-1)^2,$$

d'où on tire le système

$$\left\{\begin{array}{rrrrrrll} a & & +c & & & & =0 & \text{(termes en } z^5)\\ -a & +b & -2c & +d & & & =0 & \text{(termes en } z^4)\\ 2a & & +2c & -2d & +e & & =0 & \text{(termes en } z^3)\\ -2a & +2b & -2c & +2d & -2e & +f & =0 & \text{(termes en } z^2)\\ a & & +c & -2d & +e & -2f & =0 & \text{(termes en } z)\\ -a & +b & & +d & & +f & =1 & \text{(termes indépendants).} \end{array}\right.$$

Sa résolution est lente et fastidieuse si on la compare avec la méthode suivante qui utilise la simplification annoncée.

Pour obtenir a et b, il n'est même pas nécessaire de calculer $N(z)$, mais seulement $N'(z)=(a(z-1)+b)(z^2+1)^2$, puis de résoudre

$$\left\{\begin{array}{l} 1=A(1)=N'(1)=4b\\ 0=[\mathrm{D}_zA]_1=[\mathrm{D}_zN']_1=4a+8b, \end{array}\right.$$

d'où on tire $b=1/4$ et $a=-1/2$.

Pour obtenir c, d, e et f, il n'est pas nécessaire non plus de calculer $N(z)$, mais seulement $N''(z)=((cz+d)(z^2+1)+ez+f)(z-1)^2$, puis de résoudre

$$\left\{\begin{array}{lll} 1 & = & A(i)\ =\ N''(i)\ =\ (ei+f)(i-1)^2\\ 1 & = & A(-i)\ =\ N''(-i)\ =\ (-ei+f)(-i-1)^2\\ 0 & = & [\mathrm{D}_zA]_i=[\mathrm{D}_zN'']_i\\ & = & 2(ei+f)(i-1)+e(i-1)^2+2i(ci+d)(i-1)^2\\ 0 & = & [\mathrm{D}_zA]_{-i}\ =\ [\mathrm{D}_zN'']_{-i}\\ & = & 2(-ei+f)(-i-1)+e(-i-1)^2+2i(-ci+d)(-i-1)^2. \end{array}\right.$$

Cela étant, remarquons bien que les deux premières équations ne font intervenir que e et f. Mais on a même plus car la première équation s'écrit $1=2e-2if$ et procure aussitôt $e=1/2$ et $f=0$ car e et f sont réels: il était inutile d'écrire la deuxième équation. Enfin la troisième équation s'écrit $0=-1-2i+4ci+4d$ et donne donc $c=1/2$ et $d=1/4$.

D'où la conclusion.

6.4 Fonction exponentielle

6.4.1 Fonction exponentielle de z

On introduit généralement l'étude de cette fonction en recherchant la forme que devrait avoir une fonction f dérivable sur $\mathbb{C}$ et telle que $\mathrm{D}_z f = f$. Remarquons alors que la suite des polynômes

$$P_p(z) = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \cdots + \frac{z^p}{p!}$$

donne un espoir de solution car on a

$$\mathrm{D}_z P_p(z) = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \cdots + \frac{z^{p-1}}{(p-1)!}.$$

On est donc amené à étudier la série $\sum_{m=0}^{\infty} z^m/m!$.

Théorème 6.4.1.1 *La série $\sum_{m=0}^{\infty} z^m/m!$ converge absolument en tout point $z \in \mathbb{C}$.*

Preuve. C'est une conséquence directe du critère du quotient appliqué à la série des modules car, pour tout $z \in \mathbb{C} \setminus \{0\}$, la suite

$$\frac{|z|^{m+1}}{(m+1)!} \cdot \frac{m!}{|z|^m} = \frac{|z|}{m+1}$$

tend vers 0 si $m \to \infty$.∎

Définition. La fonction

$$\exp : \mathbb{C} \to \mathbb{C} \quad z \mapsto \sum_{m=0}^{\infty} \frac{z^m}{m!}$$

est appelée *fonction exponentielle.* En guise de $\exp(z)$, on trouve aussi la notation e^z. On a donc

$$\mathrm{e}^z = \exp(z) = \sum_{m=0}^{\infty} \frac{z^m}{m!}, \quad \forall z \in \mathbb{C}.$$

Remarque. Ce nombre $\mathrm{e}^z = \exp(z)$ est donc la limite de la série $\sum_{m=0}^{\infty} z^m/m!$ et non pas un nombre e (non encore défini) élevé à une puissance complexe (opération non encore définie).

Voici les principales propriétés de la fonction e^z.

a) $\mathrm{e}^0 = 1$.

b) **Nombre** e**.** Le nombre e est la limite de la série $\sum_{m=0}^{\infty} 1/m!$.

On a donc la propriété suivante: $\mathrm{e}^1 = \exp(1) = \mathrm{e}$.

c) *Pour tout* $z \in \mathbb{C}$*, on a* $\mathrm{e}^{\overline{z}} = \overline{\mathrm{e}^z}$*.*

En particulier, la restriction de la fonction exponentielle à $\mathbb{R}$ *est une fonction réelle.*

De fait, pour tout entier M, on a

$$\overline{\sum_{m=0}^{M} \frac{z^m}{m!}} = \sum_{m=0}^{M} \frac{\overline{z}^m}{m!}, \quad \forall z \in \mathbb{C},$$

d'où la conclusion par passage à la limite.

d) **Propriété fondamentale.** *Pour tous* z*,* $z' \in \mathbb{C}$*, on a* $\mathrm{e}^{z+z'} = \mathrm{e}^z \mathrm{e}^{z'}$*.*

De fait, on a successivement

$$\begin{aligned}
\left|\mathrm{e}^z \mathrm{e}^{z'} - \mathrm{e}^{z+z'}\right| &\underset{(*)}{=} \lim_{M\to\infty} \left|\sum_{m=0}^{M} \frac{z^m}{m!} \cdot \sum_{l=0}^{M} \frac{z'^l}{l!} - \sum_{m=0}^{M} \frac{(z+z')^m}{m!}\right| \\
&\underset{(**)}{=} \lim_{M\to\infty} \left|\sum_{m=0}^{M} \frac{z^m}{m!} \cdot \sum_{l=0}^{M} \frac{z'^l}{l!} - \sum_{m=0}^{M} \sum_{l=0}^{m} \frac{z^{m-l}}{(m-l)!} \frac{z'^l}{l!}\right| \\
&\underset{(***)}{\leq} \lim_{M\to\infty} \left(\mathrm{e}^{|z|} \cdot \sum_{m\geq M/2} \frac{|z'|^m}{m!} + \mathrm{e}^{|z'|} \cdot \sum_{m\geq M/2} \frac{|z|^m}{m!}\right) = 0.
\end{aligned}$$

En (*), nous avons recouru à la définition de la fonction exponentielle; en (**), nous avons utilisé la formule du binôme de Newton; en (***), nous remarquons que

$$\begin{aligned}
&\left|\sum_{m=0}^{M} \frac{z^m}{m!} \sum_{l=0}^{M} \frac{z'^l}{l!} - \sum_{m=0}^{M} \sum_{l=0}^{m} \frac{z^{m-l}}{(m-l)!} \frac{z'^l}{l!}\right| = \left|\sum_{m=1}^{M} \sum_{M-m<l\leq M} \frac{z^m}{m!} \frac{z'^l}{l!}\right| \\
&\leq \sum_{m=0}^{M} \frac{|z|^m}{m!} \sum_{M/2\leq l\leq M} \frac{|z'|^l}{l!} + \sum_{l=0}^{M} \frac{|z'|^l}{l!} \sum_{M/2\leq m\leq M} \frac{|z|^m}{m!}.
\end{aligned}$$

Les conséquences de cette propriété fondamentale sont nombreuses et fort importantes.

e) *La fonction* exp *appartient à* $\mathrm{C}_\infty(\mathbb{R}^2)$ *et*

$$\mathrm{D}_x \mathrm{e}^z = -i\mathrm{D}_y \mathrm{e}^z = \mathrm{D}_z \mathrm{e}^z = \mathrm{e}^z.$$

Etablissons d'abord que la fonction e^z est continue sur $\mathbb{C} = \mathbb{R}^2$. De fait, pour tout $z \in \mathbb{C}$, il vient

$$\lim_{h\to 0} \left|\mathrm{e}^{z+h} - \mathrm{e}^z\right| = \left|\mathrm{e}^z\right| \lim_{h\to 0} \left|\mathrm{e}^h - 1\right|$$

alors que

$$\left|\mathrm{e}^h - 1\right| = \left|\sum_{m=1}^{\infty} \frac{h^m}{m!}\right| \leq |h| \sum_{m=1}^{\infty} \frac{|h|^{m-1}}{(m-1)!} \leq |h|\, \mathrm{e}^{|h|},$$

où la dernière majorante est inférieure ou égale à $|h|\,\mathrm{e}$ si h est tel que $|h| \leq 1$.

Pour conclure, il suffit de prouver que la fonction exp est dérivable et telle que $\mathrm{D}_x \mathrm{e}^z = -i\mathrm{D}_y \mathrm{e}^z = \mathrm{e}^z$. Or, on a même

$$\begin{aligned} \lim_{h\to 0,\, h\in\mathbb{C}\setminus\{0\}} \left|\frac{\mathrm{e}^{z+h} - \mathrm{e}^z}{h} - \mathrm{e}^z\right| &= \left|\mathrm{e}^z\right| \lim_{h\to 0,\, h\in\mathbb{C}\setminus\{0\}} \left|\frac{\mathrm{e}^h - 1 - h}{h}\right| \\ &\leq \left|\mathrm{e}^z\right| \lim_{h\to 0,\, h\in\mathbb{C}\setminus\{0\}} |h|\, \mathrm{e}^{|h|}. \end{aligned}$$

f) *Pour tout* $z \in \mathbb{C}$, *on a* $\mathrm{e}^z \mathrm{e}^{-z} = 1$.

C'est immédiat, vu les propriétés d) et a).

g) *On a* $\mathrm{e}^z \neq 0$ *pour tout* $z \in \mathbb{C}$, ce qu'on énonce souvent sous la forme suivante: *la fonction* e^z *ne s'annule jamais.*

C'est immédiat, vu la propriété f).

h) *Pour tout* $z \in \mathbb{C}$, *on a* $\mathrm{e}^z = \mathrm{e}^{\Re z} \mathrm{e}^{i\Im z}$.

Cette dernière propriété, conséquence directe de d), va conditionner toute la suite de l'étude de la fonction e^z. En effet, pour connaître ses propriétés, il suffit de la considérer comme étant le produit de deux fonctions d'une variable réelle, à savoir e^x et e^{iy}.

Calcul approché de e^z**.** *Pour tout* $z \in \mathbb{C}$ *et tout* $M \in \mathbb{N}_0$ *vérifiant* $M \geq 2(|z| - 1)$, *les sommes partielles de la série définissant* e^z *approchent* e^z *selon*

$$|r_M(z)| = \left|\mathrm{e}^z - \sum_{m=0}^{M} \frac{z^m}{m!}\right| \leq 2\frac{|z|^{M+1}}{(M+1)!}.$$

De fait, pour tout $m \geq M+1$, il vient successivement

$$\frac{|z|^m}{m!} = \frac{|z|^{M+1}}{(M+1)!} \cdot \frac{|z|^{m-(M+1)}}{(M+2)\cdots m} \leq \frac{|z|^{M+1}}{(M+1)!} \left(\frac{|z|}{(M+2)}\right)^{m-(M+1)}.$$

D'où la conclusion car il vient alors

$$|r_M(z)| \leq \sum_{m=M+1}^{\infty} \frac{|z|^m}{m!} \leq \frac{|z|^{M+1}}{(M+1)!} \sum_{m=0}^{\infty} 2^{-m}.$$

En particulier, pour $z = 1$, il vient

$$\sum_{m=0}^{M} \frac{1}{m!} \leq \mathrm{e} \leq \sum_{m=0}^{M} \frac{1}{m!} + \frac{2}{(M+1)!}$$

pour tout $M \in \mathbb{N}_0$. Pour $M = 20$, par exemple, on obtient une approximation de e comportant 19 décimales exactes.

En fait, e est un nombre $* \rightarrow$ irrationnel et même transcendant (cf. p. 178) $\leftarrow *$ dont une valeur approximative est donnée par $2,718$.

6.4.2 Fonction exponentielle de x

Définition. La *fonction exponentielle de* $x \in \mathbb{R}$ est la restriction de la fonction exponentielle $\exp(z) = \mathrm{e}^z$ à $\mathbb{R}$.

Il ne s'agit pas à proprement parler d'une nouvelle fonction; nous avons déjà les propriétés suivantes.

a) *La fonction exponentielle est définie sur* $\mathbb{R}$ *par*

$$\exp(x) = \mathrm{e}^x = \sum_{m=0}^{\infty} \frac{x^m}{m!}.$$

b) *Pour tout* $x \in \mathbb{R}$, e^x *est réel et diffère de* 0.

c) $\mathrm{e}^0 = 1$.

d) $\mathrm{e}^1 = \mathrm{e}$.

e) $\mathrm{e}^x \in \mathrm{C}_\infty(\mathbb{R})$ *et* $\mathrm{De}^x = \mathrm{e}^x$.

f) *Pour tous* $x, x' \in \mathbb{R}$, *on a* $\mathrm{e}^{x+x'} = \mathrm{e}^x \mathrm{e}^{x'}$.

En particulier, *pour tout* $x \in \mathbb{R}$, *on a* $\mathrm{e}^x \mathrm{e}^{-x} = 1$.

Cette fonction e^x jouit en outre des propriétés suivantes.

g) *Pour tout* $x \in \mathbb{R}$, *on a* $\mathrm{e}^x > 0$.

De fait, pour tout $x \in \mathbb{R}$, il vient

$$\mathrm{e}^x = \mathrm{e}^{x/2} \cdot \mathrm{e}^{x/2} = (\mathrm{e}^{x/2})^2 > 0.$$

h) e^x *est une fonction strictement croissante.*

En particulier, *deux points* x, x' *de* $\mathbb{R}$ *sont égaux si et seulement si on a* $\mathrm{e}^x = \mathrm{e}^{x'}$.

De fait, sa dérivée est une fonction positive sans zéro.

i) e^x *est une fonction convexe sur* $\mathbb{R}$.

De fait, sa dérivée seconde est une fonction positive sans zéro.

Etudions enfin le comportement de la fonction e^x aux extrémités de son intervalle de définition.

j) *On a*

$$\lim_{x\to+\infty} \mathrm{e}^x = +\infty \textit{ et } \lim_{x\to-\infty} \mathrm{e}^x = 0.$$

Pour tout $x > 0$, il vient

$$\mathrm{e}^x = \sum_{m=0}^{\infty} \frac{x^m}{m!} > x,$$

ce qui suffit pour obtenir la première limite. La deuxième s'en déduit de suite: on a successivement

$$\lim_{x\to-\infty} \mathrm{e}^x = \lim_{x\to+\infty} \mathrm{e}^{-x} = \lim_{x\to+\infty} \frac{1}{\mathrm{e}^x} = 0.$$

Cette dernière propriété peut être sensiblement améliorée.

l) *Pour tout $p \in \mathbb{N}$, on a*

$$\lim_{x\to+\infty} \frac{\mathrm{e}^x}{x^p} = +\infty \textit{ et } \lim_{x\to-\infty} x^p\mathrm{e}^x = 0.$$

On exprime bien souvent ces formules en disant que *l'exponentielle domine toute puissance entière antagoniste en $+\infty$ et en $-\infty$.*

Pour tout $x > 0$ et tout $p \in \mathbb{N}$, il vient

$$\mathrm{e}^x = \sum_{m=0}^{\infty} \frac{x^m}{m!} \geq \frac{x^{p+1}}{(p+1)!},$$

ce qui suffit pour obtenir la première limite. La deuxième s'en déduit aussitôt vu que

$$\lim_{x\to-\infty} x^p\mathrm{e}^x = \lim_{x\to+\infty} (-x)^p\mathrm{e}^{-x} = (-1)^p \lim_{x\to+\infty} \frac{x^p}{\mathrm{e}^x} = 0.$$

Exercice. *Etablir que le nombre* e *est irrationnel.*

Suggestion. Si le nombre e est rationnel, il s'écrit $\mathrm{e} = p/q$ avec $p,\ q \in \mathbb{N}_0$. Posons $n = \sup\{3, q\}$. Comme la fonction exponentielle e^x est à valeurs réelles et appartient à $\mathrm{C}_\infty(\mathbb{R})$, la formule de Taylor limitée à l'ordre n peut être appliquée entre 0 et 1, et procure un accroissement $h \in]0,1[$ tel que

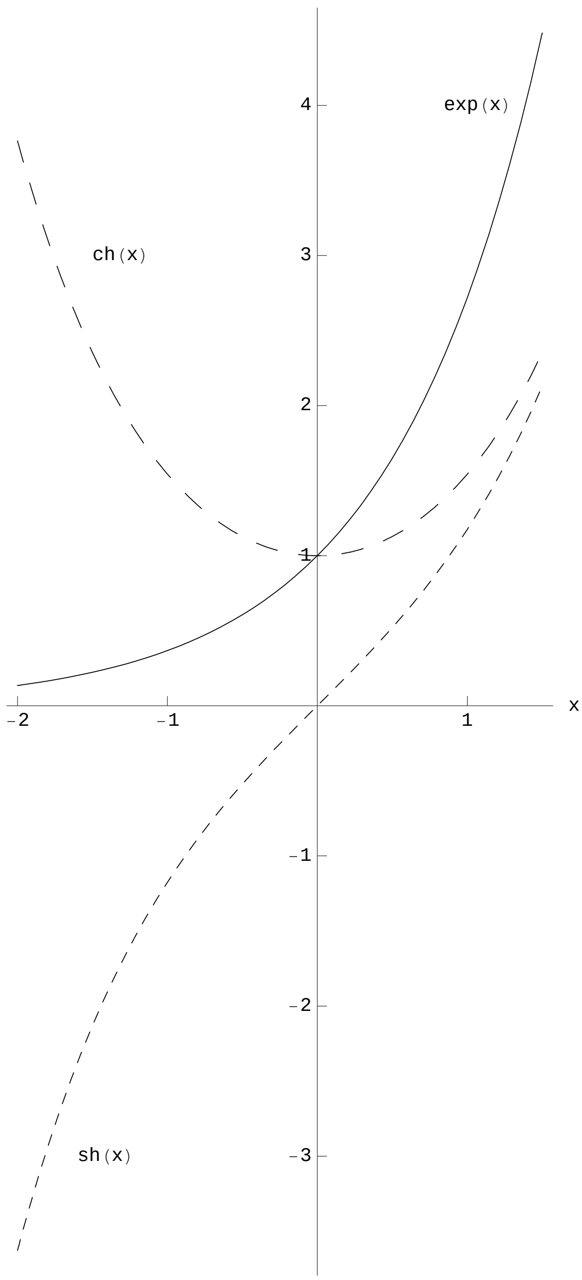
$$\frac{p}{q} = 1 + 1 + \frac{1}{2!} + \ldots + \frac{1}{n!} + \frac{\mathrm{e}^h}{(n+1)!}. \quad (*)$$

Comme on a

$$1 < \mathrm{e}^h < \mathrm{e} = 1 + 1 + \sum_{n=2}^{\infty} \frac{1}{n!} < 2 + \sum_{n=1}^{\infty} \frac{1}{2^n} = 3,$$

on obtient directement une contradiction en multipliant les deux membres de l'égalité $(*)$ par $n!$.

$* \rightarrow$ Au prix d'une démonstration très ardue, on peut aussi établir que *le nombre* e *est transcendant*, c'est-à-dire qu'il ne peut être zéro d'un polynôme à coefficients entiers sur $\mathbb{R}$. $\leftarrow *$



6.5 Fonctions hyperboliques

Définition. Une fonction réelle f définie sur $\mathbb{R}$ est

a) *paire* si on a $f(x) = f(-x)$ pour tout $x \in \mathbb{R}$,

b) *impaire* si on a $f(x) = -f(-x)$ pour tout $x \in \mathbb{R}$.

Etant donné une fonction réelle f définie sur $\mathbb{R}$, on vérifie aussitôt que

a) $g : \mathbb{R} \to \mathbb{R} \quad x \mapsto \frac{1}{2}(f(x) + f(-x))$ est une fonction paire; elle est appelée *la partie paire de* f,

b) $h : \mathbb{R} \to \mathbb{R} \quad x \mapsto \frac{1}{2}(f(x) - f(-x))$ est une fonction impaire; elle est appelée *la partie impaire de* f.

Comme on a bien sûr $f = g + h$, toute fonction réelle sur $\mathbb{R}$ est la somme de ses parties paire et impaire. L'application de cette idée à la fonction exponentielle est fort enrichissante.

Définitions. a) Le *cosinus hyperbolique*, noté ch, est la partie paire de la fonction exponentielle e^x.

b) Le *sinus hyperbolique*, noté sh, est la partie impaire de la fonction exponentielle e^x.

Les formules suivantes sont donc valables pour tout $x \in \mathbb{R}$:

$$\begin{aligned} \mathrm{ch}(x) &= \tfrac{1}{2}(e^x + e^{-x}) \qquad & \mathrm{sh}(x) &= \tfrac{1}{2}(e^x - e^{-x}) \\ e^x &= \mathrm{ch}(x) + \mathrm{sh}(x) \qquad & e^{-x} &= \mathrm{ch}(x) - \mathrm{sh}(x). \end{aligned}$$

Comme les séries définissant e^x et e^{-x} sont absolument convergentes, on tire directement les développements en série suivants

$$\mathrm{ch}(x) = \sum_{m=0}^{\infty} \frac{x^{2m}}{(2m)!}, \qquad \mathrm{sh}(x) = \sum_{m=0}^{\infty} \frac{x^{2m+1}}{(2m+1)!}$$

valables pour tout $x \in \mathbb{R}$, ces séries étant absolument convergentes.

Les propriétés de ces fonctions ch et sh découlent directement de celles de la fonction exponentielle.

a) *Les fonctions* ch *et* sh *sont définies sur* $\mathbb{R}$*, sont réelles et sont respectivement paire et impaire.*

b) *On a* $\mathrm{ch}(0) = 1$ *et* $\mathrm{ch}(x) > 1$ *pour tout* $x \in]-\infty, 0[\cup]0, +\infty[$.

On a également $\mathrm{sh}(0) = 0$, $\mathrm{sh}(x) < 0$ *pour tout* $x \in]-\infty, 0[$ *et* $\mathrm{sh}(x) > 0$ *pour tout* $x \in]0, +\infty[$.

c) *On a*

$$\lim_{x\to\pm\infty} \mathrm{ch}(x) = +\infty \quad \textit{et} \quad \lim_{x\to\pm\infty} \mathrm{sh}(x) = \pm\infty.$$

d) *On a*

$$\begin{array}{l} \mathrm{ch} \in \mathrm{C}_\infty(\mathbb{R}) \quad \textit{et} \quad \mathrm{Dch} = \mathrm{sh}, \\ \mathrm{sh} \in \mathrm{C}_\infty(\mathbb{R}) \quad \textit{et} \quad \mathrm{Dsh} = \mathrm{ch}. \end{array}$$

e) *La fonction* ch *est strictement décroissante sur* $]-\infty, 0[$ *et strictement croissante sur* $]0, +\infty[$.

La fonction sh *est strictement croissante sur* $\mathbb{R}$.

f) *La fonction* ch *est convexe sur* $\mathbb{R}$.

La fonction sh *est concave sur* $]-\infty, 0[$ *et convexe sur* $]0, +\infty[$.

A partir des fonctions ch et sh, on développe le *calcul hyperbolique*, portant sur les quatre fonctions principales suivantes, dites *fonctions hyperboliques*

a) le cosinus hyperbolique: ch;

b) le sinus hyperbolique: sh;

c) la *tangente hyperbolique:* c'est la fonction réelle th définie sur $\mathbb{R}$ par

$$\mathrm{th}(x) = \frac{\mathrm{sh}(x)}{\mathrm{ch}(x)}, \quad \forall x \in \mathbb{R}.$$

Il s'agit donc d'une fonction impaire qui appartient à $\mathrm{C}_\infty(\mathbb{R})$.

d) la *cotangente hyperbolique:* c'est la fonction réelle coth définie sur $\mathbb{R} \setminus \{0\}$ par

$$\coth(x) = \frac{\mathrm{ch}(x)}{\mathrm{sh}(x)}, \quad \forall x \in \mathbb{R} \setminus \{0\}.$$

Il s'agit donc d'une fonction impaire qui appartient à $\mathrm{C}_\infty(\mathbb{R} \setminus \{0\})$.

Voici les identités fondamentales du calcul hyperbolique.

a) **Formules d'addition et de soustraction.** *Pour tous* $x, y \in \mathbb{R}$,

$$\begin{aligned} \mathrm{ch}(x \pm y) &= \mathrm{ch}(x)\,\mathrm{ch}(y) \pm \mathrm{sh}(x)\,\mathrm{sh}(y), \\ \mathrm{sh}(x \pm y) &= \mathrm{sh}(x)\,\mathrm{ch}(y) \pm \mathrm{ch}(x)\,\mathrm{sh}(y), \\ \mathrm{th}(x \pm y) &= \frac{\mathrm{th}(x) \pm \mathrm{th}(y)}{1 \pm \mathrm{th}(x)\,\mathrm{th}(y)}. \end{aligned}$$

En ce qui concerne les fonctions ch et sh, les démonstrations sont semblables. Il suffit, par exemple, de noter que, d'une part,

$$\mathrm{ch}(x + y) = \frac{\mathrm{e}^{x+y} + \mathrm{e}^{-x-y}}{2}$$

et, d'autre part,

$$\begin{aligned}\mathrm{ch}(x)\mathrm{ch}(y)+\mathrm{sh}(x)\mathrm{sh}(y) &= \frac{\mathrm{e}^x+\mathrm{e}^{-x}}{2}\cdot\frac{\mathrm{e}^y+\mathrm{e}^{-y}}{2}+\frac{\mathrm{e}^x-\mathrm{e}^{-x}}{2}\cdot\frac{\mathrm{e}^y-\mathrm{e}^{-y}}{2}\\ &= \frac{1}{4}(\mathrm{e}^{x+y}+\mathrm{e}^{x-y}+\mathrm{e}^{-x+y}+\mathrm{e}^{-x-y}\\ &\qquad +\mathrm{e}^{x+y}-\mathrm{e}^{x-y}-\mathrm{e}^{-x+y}+\mathrm{e}^{-x-y}).\end{aligned}$$

Pour les formules relatives à th, on a par exemple

$$\mathrm{th}(x+y)=\frac{\mathrm{sh}(x+y)}{\mathrm{ch}(x+y)}=\frac{\mathrm{sh}(x)\,\mathrm{ch}(y)+\mathrm{ch}(x)\,\mathrm{sh}(y)}{\mathrm{ch}(x)\,\mathrm{ch}(y)+\mathrm{sh}(x)\,\mathrm{sh}(y)}=\frac{\mathrm{th}(x)+\mathrm{th}(y)}{1+\mathrm{th}(x)\,\mathrm{th}(y)}.$$

b) En prenant la formule exprimant $\mathrm{ch}(x-y)$ en $y=x$, on obtient

$$1=\mathrm{ch}^2(x)-\mathrm{sh}^2(x),\quad \forall x\in\mathbb{R},$$

d'où on tire de suite

$$\begin{aligned}1 &= \mathrm{th}^2(x)+\frac{1}{\mathrm{ch}^2(x)},\quad \forall x\in\mathbb{R},\\ 1 &= \mathrm{coth}^2(x)-\frac{1}{\mathrm{sh}^2(x)},\quad \forall x\in\mathbb{R}\setminus\{0\}.\end{aligned}$$

Ces trois formules permettent alors d'exprimer chacune des fonctions ch, sh, th et coth au moyen d'une expression simple qui ne fait intervenir qu'une des autres fonctions.

Ainsi, par exemple, on a

$$\mathrm{ch}(x)=\sqrt{1+\mathrm{sh}^2(x)},\quad \forall x\in\mathbb{R},$$

et

$$\mathrm{sh}(x)=\begin{cases}\sqrt{\mathrm{ch}^2(x)-1}, & \forall x\geq 0,\\ -\sqrt{\mathrm{ch}^2(x)-1}, & \forall x\leq 0.\end{cases}$$

c) En exprimant les formules d'addition en $y=x$, il vient

$$\begin{aligned}\mathrm{ch}(2x) &= \mathrm{ch}^2(x)+\mathrm{sh}^2(x),\quad \forall x\in\mathbb{R},\\ \mathrm{sh}(2x) &= 2\,\mathrm{sh}(x)\,\mathrm{ch}(x),\quad \forall x\in\mathbb{R},\\ \mathrm{th}(2x) &= \frac{2\,\mathrm{th}(x)}{1+\mathrm{th}^2(x)},\quad \forall x\in\mathbb{R}.\end{aligned}$$

d) En particulier, en combinant les formules b) et c), on obtient

$$\begin{aligned}
\operatorname{ch}(2x)+1 &= 2\operatorname{ch}^2(x), \quad \forall x \in \mathbb{R},\\
\operatorname{ch}(2x)-1 &= 2\operatorname{sh}^2(x), \quad \forall x \in \mathbb{R}.
\end{aligned}$$

e) Les formules suivantes sont valables pour tous x, $y \in \mathbb{R}$:

$$\begin{aligned}
\operatorname{ch}(x+y)+\operatorname{ch}(x-y) &= 2\operatorname{ch}(x)\operatorname{ch}(y),\\
\operatorname{sh}(x+y)+\operatorname{sh}(x-y) &= 2\operatorname{sh}(x)\operatorname{ch}(y),\\
\operatorname{ch}(x+y)-\operatorname{ch}(x-y) &= 2\operatorname{sh}(x)\operatorname{sh}(y);
\end{aligned}$$

ce sont des conséquences directes de a).

f) Il s'ensuit que, pour tous x, $y \in \mathbb{R}$, on a

$$\begin{aligned}
\operatorname{ch}(x)+\operatorname{ch}(y) &= 2\operatorname{ch}\frac{x+y}{2}\operatorname{ch}\frac{x-y}{2}\\
\operatorname{ch}(x)-\operatorname{ch}(y) &= 2\operatorname{sh}\frac{x+y}{2}\operatorname{sh}\frac{x-y}{2}\\
\operatorname{sh}(x)+\operatorname{sh}(y) &= 2\operatorname{sh}\frac{x+y}{2}\operatorname{ch}\frac{x-y}{2},\\
\operatorname{sh}(x)-\operatorname{sh}(y) &= 2\operatorname{sh}\frac{x-y}{2}\operatorname{ch}\frac{x+y}{2}.
\end{aligned}$$

g) Enfin, on peut, pour tout entier $m \in \mathbb{N}_0$,

i) exprimer $\operatorname{ch}(mx)$ et $\operatorname{sh}(mx)$ au moyen des puissances entières de degré $\leq m$ de $\operatorname{ch}(x)$ et $\operatorname{sh}(x)$. Il suffit pour cela de sommer et de soustraire les formules

$$(\operatorname{ch}(x) \pm \operatorname{sh}(x))^m = \mathrm{e}^{\pm mx} = \operatorname{ch}(mx) \pm \operatorname{sh}(mx),$$

ii) exprimer $\operatorname{ch}^m(x)$ et $\operatorname{sh}^m(x)$ au moyen des fonctions $\operatorname{ch}(px)$ et $\operatorname{sh}(px)$ pour les valeurs entières de p telles que $0 \leq p \leq m$. Il suffit pour cela d'écrire

$$\operatorname{ch}^m(x) = 2^{-m}(\mathrm{e}^x + \mathrm{e}^{-x})^m \text{ et } \operatorname{sh}^m(x) = 2^{-m}(\mathrm{e}^x - \mathrm{e}^{-x})^m,$$

puis de développer les seconds membres au moyen de la formule du binôme de Newton et enfin de grouper les termes deux à deux à partir des extrêmes.

Traitons quelques exemples.

Dans le cas où m est égal à 2, il vient

$$\operatorname{ch}^2(x) \pm 2\operatorname{ch}(x)\operatorname{sh}(x) + \operatorname{sh}^2(x) = \operatorname{ch}(2x) \pm \operatorname{sh}(2x)$$

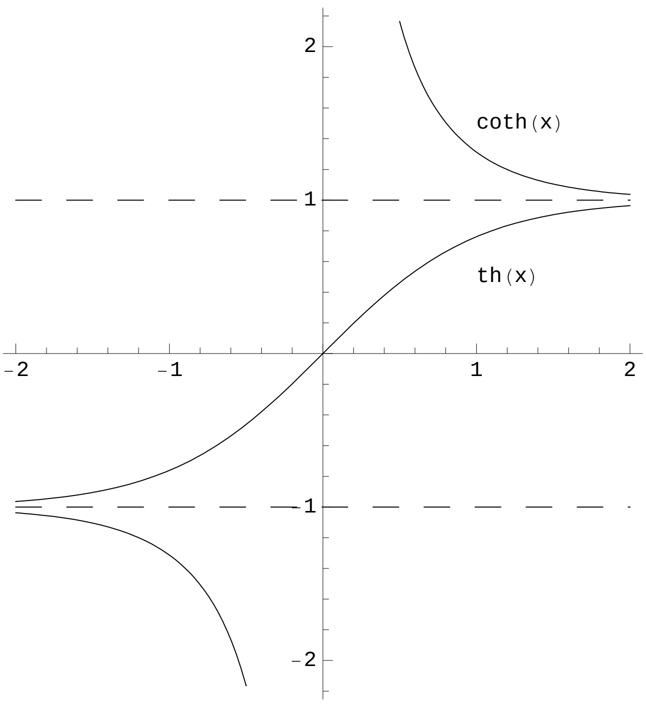
et on retrouve les formules c) en prenant la demi-somme ou la demi-différence de ces relations.

De même, de

$$\begin{array}{rcccl}\mathrm{ch}^2(x) & = & \frac{1}{4}(\mathrm{e}^x + \mathrm{e}^{-x})^2 & = & \frac{1}{4}(\mathrm{e}^{2x} + 2 + \mathrm{e}^{-2x}) \\ \mathrm{sh}^2(x) & = & \frac{1}{4}(\mathrm{e}^x - \mathrm{e}^{-x})^2 & = & \frac{1}{4}(\mathrm{e}^{2x} - 2 + \mathrm{e}^{-2x}),\end{array}$$

on tire de suite les formules d).

Revenons à présent aux fonctions th et coth. En fait, nous en avons déjà obtenu de nombreuses propriétés. En outre, les propriétés suivantes sont immédiates et permettent de mieux réaliser leur graphique.



a) $\mathrm{th}(0) = 0$.

b) $\mathrm{Dth} = 1/\mathrm{ch}^2$ *sur* $\mathbb{R}$ *et* $\mathrm{Dcoth} = -1/\mathrm{sh}^2$ *sur* $\mathbb{R} \setminus \{0\}$.

c) *La fonction* th *est strictement croissante.*

La fonction coth *est strictement décroissante sur* $]-\infty, 0[$ *et sur* $]0, +\infty[$ (ce qui ne signifie pas que coth soit une fonction strictement décroissante sur son ensemble de définition!).

d) *La fonction* th *est convexe sur* $]-\infty, 0[$ *et concave sur* $]0, +\infty[$.

La fonction coth *est concave sur* $]-\infty, 0[$ *et convexe sur* $]0, +\infty[$.

e) *On a*

$$
\begin{aligned}
\lim_{x\to\pm\infty} \operatorname{th}(x) &= \lim_{x\to\pm\infty} \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{\mathrm{e}^x + \mathrm{e}^{-x}} = \left\{ \begin{array}{c} 1^- \\ (-1)^+ \end{array} \right\}, \\
\lim_{x\to\pm\infty} \coth(x) &= \lim_{x\to\pm\infty} \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{\mathrm{e}^x - \mathrm{e}^{-x}} = \left\{ \begin{array}{c} 1^+ \\ (-1)^- \end{array} \right\}, \\
\lim_{x\to 0^\pm} \coth(x) &= \lim_{x\to 0^\pm} \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{\mathrm{e}^x - \mathrm{e}^{-x}} = \left\{ \begin{array}{c} +\infty \\ -\infty \end{array} \right\}.
\end{aligned}
$$

6.6 Fonctions circulaires

6.6.1 Fonction e^{ix}

Le nombre e^{ix} est défini pour tout $x \in \mathbb{R}$ par

$$\mathrm{e}^{ix} = \sum_{m=0}^{\infty} \frac{(ix)^m}{m!} = 1 + i\frac{x}{1!} - \frac{x^2}{2!} - i\frac{x^3}{3!} + \dots$$

Dès lors, e^{ix} est une fonction (à valeurs complexes) définie sur $\mathbb{R}$.

Pour l'étudier, on examine séparément les propriétés de sa partie réelle et de sa partie imaginaire. On introduit donc les fonctions suivantes:

a) la fonction cos (on dit *cosinus*) définie sur $\mathbb{R}$ par

$$\cos(x) = \Re \mathrm{e}^{ix} = \tfrac{1}{2}(\mathrm{e}^{ix} + \mathrm{e}^{-ix}), \quad \forall x \in \mathbb{R}.$$

b) la fonction sin (on dit *sinus*) définie sur $\mathbb{R}$ par

$$\sin(x) = \Im \mathrm{e}^{ix} = \tfrac{1}{2i}(\mathrm{e}^{ix} - \mathrm{e}^{-ix}), \quad \forall x \in \mathbb{R}.$$

Ce sont des fonctions réelles sur $\mathbb{R}$; elles sont appelées *fonctions circulaires.*

Bien sûr, pour tout $x \in \mathbb{R}$, on a évidemment les formules suivantes

$$\mathrm{e}^{ix} = \cos(x) + i\sin(x) \quad \text{et} \quad \mathrm{e}^{-ix} = \cos(x) - i\sin(x),$$

appelées *formules d'Euler.*

Comme les séries qui définissent e^{ix} et e^{-ix} sont absolument convergentes, on déduit de suite les développements en série suivants

$$
\begin{aligned}
\cos(x) &= \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m}}{(2m)!} \\
\sin(x) &= \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m+1}}{(2m+1)!},
\end{aligned}
$$

valables pour tout $x \in \mathbb{R}$; ces séries étant absolument convergentes.

Remarques. a) Les fonctions cos et sin sont donc purement et simplement des fonctions élémentaires de l'analyse mathématique. Pour le moment, leur définition ne permet aucun raccord avec des notions géométriques telles que le "cercle trigonométrique". Nous ferons plus loin la liaison avec ces interprétations géométriques.

b) Insistons également sur le fait que *toutes les formules établies ici ne sont valables que si x est exprimé en radians.*

6.6.2 Fonctions circulaires

Les propriétés des fonctions cos et sin se déduisent de suite de celles de la fonction exponentielle.

a) cos *et* sin *sont des fonctions réelles sur* $\mathbb{R}$.

b) $\cos(0) = 1$ *et* $\sin(0) = 0$.

Il suffit de noter que $e^{i0} = e^0 = 1$.

c) cos *est une fonction paire;* sin *est une fonction impaire.*

d) *On a*

$$\begin{aligned} &\cos \in C_\infty(\mathbb{R}) \quad \textit{et} \quad \mathrm{D}\cos = -\sin, \\ &\sin \in C_\infty(\mathbb{R}) \quad \textit{et} \quad \mathrm{D}\sin = \cos\,. \end{aligned}$$

Il suffit de noter qu'on a $e^{ix} \in C_\infty(\mathbb{R})$ et $\mathrm{D}e^{ix} = ie^{ix}$.

e) *On a*

$$\left|e^{ix}\right| = 1, \quad \forall x \in \mathbb{R},$$

donc

$$\cos^2(x) + \sin^2(x) = 1, \quad \forall x \in \mathbb{R}.$$

En particulier, pour tout $x \in \mathbb{R}$, il vient

$$-1 \leq \cos(x) \leq 1 \quad \textit{et} \quad -1 \leq \sin(x) \leq 1$$

et les nombres $\cos(x)$ *et* $\sin(x)$ *ne sont jamais simultanément nuls.*

De fait, pour tout $x \in \mathbb{R}$, il vient

$$\cos^2(x) + \sin^2(x) = \left|e^{ix}\right|^2 = e^{ix}e^{-ix} = e^0 = 1.$$

Remarquons que les fonctions cos et sin étant continues sur $\mathbb{R}$,

$$\{\, x \in \mathbb{R} : \cos(x) = 0\} \quad \text{et} \quad \{\, x \in \mathbb{R} : \sin(x) = 0\}$$

sont des parties fermées de $\mathbb{R}$. Cela étant, à partir des fonctions cos et sin, on développe le *calcul trigonométrique,* portant sur les quatre fonctions principales suivantes, dites *fonctions circulaires* ou *trigonométriques:*

a) le cosinus: cos
b) le sinus: sin
c) la *tangente:*

$$\operatorname{tg} : \Omega = \mathbb{R} \setminus \{\, x : \cos(x) = 0\} \to \mathbb{R} \quad x \mapsto \frac{\sin(x)}{\cos(x)}.$$

Il s'agit donc d'une fonction impaire appartenant à $\mathrm{C}_\infty(\Omega)$.
d) la *cotangente:*

$$\operatorname{cotg} : \Omega' = \mathbb{R} \setminus \{\, x : \sin(x) = 0\} \to \mathbb{R} \quad x \mapsto \cos(x)/\sin(x).$$

Il s'agit donc d'une fonction impaire appartenant à $\mathrm{C}_\infty(\Omega')$.

Voici les identités fondamentales du calcul trigonométrique.

a) **Formules d'addition et de soustraction.** *Pour tous x, $y \in \mathbb{R}$,*

$$\begin{array}{rcl} \cos(x \pm y) & = & \cos(x)\cos(y) \mp \sin(x)\sin(y), \\ \sin(x \pm y) & = & \sin(x)\cos(y) \pm \cos(x)\sin(y), \end{array}$$

et, pour autant que $\operatorname{tg}(x \pm y)$, $\operatorname{tg}(x)$ *et* $\operatorname{tg}(y)$ *soient définis,*

$$\operatorname{tg}(x \pm y) = \frac{\operatorname{tg}(x) \pm \operatorname{tg}(y)}{1 \mp \operatorname{tg}(x)\operatorname{tg}(y)}.$$

D'une part, de $\mathrm{e}^{i(x \pm y)} = \mathrm{e}^{ix}\mathrm{e}^{\pm iy}$, on tire de suite

$$\begin{array}{rcl} \cos(x \pm y) + i\sin(x \pm y) & = & (\cos(x) + i\sin(x))(\cos(y) \pm i\sin(y)) \\ & = & (\cos(x)\cos(y) \mp \sin(x)\sin(y)) \\ & & \quad + i(\sin(x)\cos(y) \pm \cos(x)\sin(y)), \end{array}$$

c'est-à-dire les formules relatives aux fonctions cos et sin, en recourant aux parties réelle et imaginaire respectivement.

D'autre part, il vient

$$\operatorname{tg}(x \pm y) = \frac{\sin(x \pm y)}{\cos(x \pm y)} = \frac{\sin(x)\cos(y) \pm \cos(x)\sin(y)}{\cos(x)\cos(y) \mp \sin(x)\sin(y)} = \frac{\operatorname{tg}(x) \pm \operatorname{tg}(y)}{1 \mp \operatorname{tg}(x)\operatorname{tg}(y)}.$$

b) En prenant en $y = x$ la formule obtenue en a) pour décomposer $\cos(x - y)$, on obtient la formule $1 = \cos^2(x) + \sin^2(x)$ pour tout $x \in \mathbb{R}$; on en déduit aussitôt les résultats suivants

$$1 = \frac{1}{\cos^2(x)} - \operatorname{tg}^2(x), \quad \forall x \in \Omega,$$

et

$$1 = \frac{1}{\sin^2(x)} - \operatorname{cotg}^2(x), \quad \forall x \in \Omega'.$$

Remarque. Ces trois formules permettrons plus tard d'exprimer chacune des fonctions cos, sin, tg et cotg au moyen d'une expression simple qui ne fait intervenir qu'une des autres fonctions. Pour le moment, par exemple, pour tout nombre réel x, nous avons $\cos(x) = (1 - \sin^2(x))^{1/2}$ ou $\cos(x) = -(1 - \sin^2(x))^{1/2}$ et rien ne permet de choisir entre ces deux valeurs, sauf le recours aux valeurs explicites, ce qu'on désire éviter.

c) En exprimant les formules d'addition en $y = x$, il vient

$$\begin{aligned} \cos(2x) &= \cos^2(x) - \sin^2(x), \quad \forall x \in \mathbb{R}, \\ \sin(2x) &= 2\sin(x)\cos(x), \quad \forall x \in \mathbb{R}, \end{aligned}$$

et pour autant que x et $2x$ appartiennent à Ω,

$$\mathrm{tg}(2x) = \frac{2\,\mathrm{tg}(x)}{1 - \mathrm{tg}^2(x)}.$$

d) En particulier, en combinant les formules b) et c), on obtient

$$\begin{aligned} 1 + \cos(2x) &= 2\cos^2(x), \quad \forall x \in \mathbb{R}, \\ 1 - \cos(2x) &= 2\sin^2(x), \quad \forall x \in \mathbb{R}. \end{aligned}$$

e) Les formules suivantes sont claires à partir de a)

$$\begin{aligned} \cos(x+y) + \cos(x-y) &= 2\cos(x)\cos(y), \\ \sin(x+y) + \sin(x-y) &= 2\sin(x)\cos(y), \\ \cos(x+y) - \cos(x-y) &= -2\sin(x)\sin(y), \end{aligned}$$

pour tous x, $y \in \mathbb{R}$.

f) Il s'ensuit que, pour tous x, $y \in \mathbb{R}$, on a

$$\begin{aligned} \cos(x) + \cos(y) &= 2\cos\frac{x+y}{2}\cos\frac{x-y}{2}, \\ \cos(x) - \cos(y) &= -2\sin\frac{x+y}{2}\sin\frac{x-y}{2}, \\ \sin(x) + \sin(y) &= 2\sin\frac{x+y}{2}\cos\frac{x-y}{2}, \\ \sin(x) - \sin(y) &= 2\sin\frac{x-y}{2}\cos\frac{x+y}{2}. \end{aligned}$$

g) Enfin, on peut, pour tout entier $m \in \mathbb{N}_0$ et tout $x \in \mathbb{R}$,
i) exprimer $\cos(mx)$ et $\sin(mx)$ au moyen des puissances entières de degré $\leq m$ de $\cos(x)$ et $\sin(x)$. Il suffit pour cela de prendre la partie réelle et la partie imaginaire de la formule suivante, connue sous le nom de *relation de Moivre*

$$(\cos(x) + i\sin(x))^m = \mathrm{e}^{imx} = \cos(mx) + i\sin(mx).$$

ii) exprimer $\cos^m(x)$ et $\sin^m(x)$ au moyen des fonctions $\cos(px)$ et $\sin(px)$ pour les valeurs entières de p telles que $0 \leq p \leq m$. Il suffit pour cela d'écrire

$$\cos^m(x) = 2^{-m}(\mathrm{e}^{ix} + \mathrm{e}^{-ix})^m \quad \text{et} \quad \sin^m(x) = (2i)^{-m}(\mathrm{e}^{ix} - \mathrm{e}^{-ix})^m,$$

puis de développer les seconds membres au moyen de la formule du binôme de Newton et enfin de grouper les termes deux à deux à partir des extrêmes.

Traitons quelques exemples.

Dans le cas où m est égal à 2, il vient

$$\cos^2(x) - \sin^2(x) + i\,2\sin(x)\cos(x) = \cos(2x) + i\sin(2x)$$

et on retrouve les formules c). De même, de

$$\cos^2(x) = \tfrac{1}{4}(\mathrm{e}^{ix} + \mathrm{e}^{-ix})^2 = \tfrac{1}{4}(\mathrm{e}^{2ix} + 2 + \mathrm{e}^{-2ix})$$

et

$$\sin^2(x) = -\tfrac{1}{4}(\mathrm{e}^{ix} - \mathrm{e}^{-ix})^2 = -\tfrac{1}{4}(\mathrm{e}^{2ix} - 2 + \mathrm{e}^{-2ix}),$$

on tire de suite les formules d).

Remarque. Nous ne sommes pas encore à même de donner la forme du graphique des fonctions circulaires ni même de caractériser les ouverts Ω et Ω' où les fonctions tg et cotg sont définies. Ce problème est résolu aux deux paragraphes suivants.

6.6.3 Le nombre π

Lemme 6.6.3.1 *Pour tout $x \in \,]0,2[$, on a $\sin(x) > 0$.*

Preuve. Pour tout $x > 0$, il vient

$$\frac{\sin(x)}{x} = \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m}}{(2m+1)!}$$

où le second membre est une série alternée dont le terme général tend vers 0 car cette série converge. En outre, pour tout $x \in \,]0,2[$ et tout entier $m \geq 0$, on a

$$\frac{x^{2m+2}}{(2m+3)!} \leq \frac{x^{2m}}{(2m+1)!} \Longleftrightarrow x^2 \leq (2m+3)(2m+2).$$

De là, pour tout $x \in \,]0,2[$, le critère d'Abel procure la majoration

$$\frac{\sin(x)}{x} \geq 1 - \frac{x^2}{3!} \geq 1 - \frac{2^2}{3!} = \frac{1}{3},$$

ce qui suffit.∎

Lemme 6.6.3.2 *On a* $\cos(2) < 0$.

Preuve. On a

$$\cos(2) = \sum_{m=0}^{\infty} (-1)^m \frac{2^{2m}}{(2m)!}$$

et cette série alternée dont le terme général tend vers 0 est telle que

$$\frac{2^{2m+2}}{(2m+2)!} \leq \frac{2^{2m}}{(2m)!}, \quad \forall m \in \mathbb{N}_0.$$

Dès lors, le critère d'Abel donne

$$\cos(2) \leq 1 - \frac{2^2}{2!} + \frac{2^4}{4!} = 1 - 2 + \frac{2}{3} = -\frac{1}{3}. \blacksquare$$

Proposition 6.6.3.3 *Il existe un zéro et un seul de la fonction* cos *dans* $]0, 2[$.

Preuve. La fonction cos est continue sur $\mathbb{R}$, vaut 1 en 0 et est strictement négative en 2. Vu le théorème des valeurs intermédiaires, elle admet donc au moins un zéro dans $]0, 2[$. De plus, comme on a $\mathrm{D}\cos = -\sin$, le premier lemme signale que la fonction cos est strictement décroissante sur $]0, 2[$. Au total, cos admet un zéro et un seul dans $]0, 2[$. $\blacksquare$

Définition. Le nombre π est défini comme étant égal à deux fois le zéro unique de $\cos(x)$ dans $]0, 2[$.

On a donc déjà les propriétés suivantes.

Théorème 6.6.3.4 a) *La fonction* cos *est réelle sur* $\mathbb{R}$*, appartient à* $\mathrm{C}_\infty(\mathbb{R})$ *et vérifie* $\cos(0) = 1$ *et* $\cos(\pi/2) = 0$*. De plus, elle est strictement décroissante et concave sur* $[0, \pi/2]$.

b) *La fonction* sin *est réelle sur* $\mathbb{R}$*, appartient à* $\mathrm{C}_\infty(\mathbb{R})$ *et vérifie* $\sin(0) = 0$ *et* $\sin(\pi/2) = 1$*. De plus, elle est strictement croissante et concave sur* $[0, \pi/2]$*. Enfin on a* $\sin(\pi) = 0$ *et* $\sin(x) > 0$ *pour tout* $x \in]0, \pi[$.

Preuve. Tout est conséquence directe de ce qui précède si on note que la formule $\sin(x) = 2\sin(x/2)\cos(x/2)$ procure aussitôt $\sin(\pi) = 0$ et $\sin(x) > 0$ pour tout $x \in]0, \pi[$. $\blacksquare$

Corollaire 6.6.3.5 *On a*

$$\mathrm{e}^0 = 1, \quad \mathrm{e}^{i\pi/2} = i, \quad \mathrm{e}^{i\pi} = -1, \quad \mathrm{e}^{3i\pi/2} = -i, \quad \mathrm{e}^{2i\pi} = 1,$$

donc

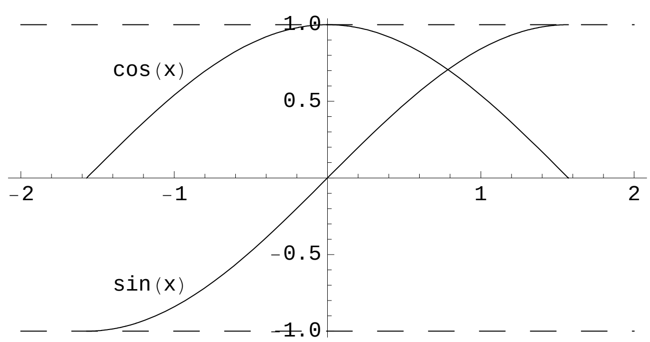
$$\cos(0) = 1, \quad \cos(\pi/2) = 0, \quad \cos(\pi) = -1, \quad \cos(3\pi/2) = 0, \quad \cos(2\pi) = 1$$

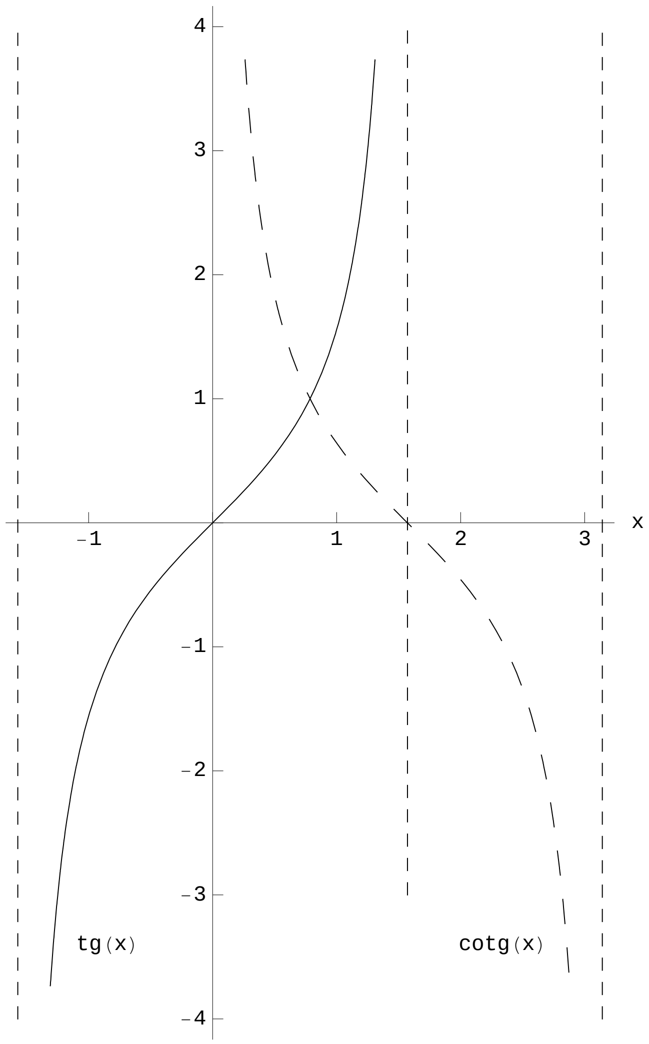
et

$$\sin(0) = 0, \quad \sin(\pi/2) = 1, \quad \sin(\pi) = 0, \quad \sin(3\pi/2) = -1, \quad \sin(2\pi) = 0. \blacksquare$$

Calcul de π. Nous ne sommes pas en mesure à présent de donner une méthode élégante permettant de calculer la valeur de ce nombre π. De telles méthodes seront données en deuxième candidature, au chapitre relatif à la convergence uniforme. $* \rightarrow$ Signalons cependant déjà que π est un nombre irrationnel dont une valeur approchée est donnée par $3,1416$. $\leftarrow *$

Remarque. Nous pouvons dès maintenant donner la forme des graphiques des fonctions cos et sin sur l'intervalle $]0, \pi/2[$ et par conséquent celui des fonctions tg et cotg sur ce même intervalle. Comme cos est une fonction paire et comme sin, tg et cotg sont des fonctions impaires, nous les connaissons en fait déjà sur l'intervalle $]-\pi/2, \pi/2[$ pour sin, cos et tg et sur $]-\pi/2, 0[\cup]0, \pi/2[$ pour cotg.





Notre but consiste maintenant à établir que ces renseignements suffisent pour connaître parfaitement les fonctions cos, sin, tg et cotg et sont même surabondants: en fait, il suffit de connaître une des deux fonctions cos ou sin sur $[0, \pi/2[$ pour déterminer $\cos(x)$ et $\sin(x)$, donc $\mathrm{tg}(x)$ et $\mathrm{cotg}(x)$, en un point quelconque de $\mathbb{R}$. C'est l'objet du paragraphe suivant.

6.6.4 Formules du calcul trigonométrique

Formules de réduction. *Pour tout $m \in \mathbb{Z}$ et tout $x \in \mathbb{R}$,*

$$\left\{ \begin{array}{c} \cos \\ \sin \end{array} \right\} (m\frac{\pi}{2} + x) = \left\{ \begin{array}{c} \Re \\ \Im \end{array} \right\} (\mathrm{e}^{im\pi/2}\mathrm{e}^{ix}) = \left\{ \begin{array}{c} \Re \\ \Im \end{array} \right\} i^m(\cos(x) + i\sin(x)).$$

Or on a

$$\begin{array}{ll} i^m = 1 & \text{si } m \text{ est un multiple de 4,} \\ i^m = i & \text{si } m-1 \text{ est un multiple de 4,} \\ i^m = -1 & \text{si } m-2 \text{ est un multiple de 4,} \\ i^m = -i & \text{si } m-3 \text{ est un multiple de 4.} \end{array}$$

Dès lors, la connaissance des valeurs de cos et de sin sur $[0, \pi/2[$ donne la connaissance des fonctions cos et sin sur $\mathbb{R}$.

En particulier, on a les formules suivantes.

a) **Formules des compléments.** *Pour tout* $x \in \mathbb{R}$,

$$\cos(\frac{\pi}{2} - x) = \sin(x) \quad \text{et} \quad \sin(\frac{\pi}{2} - x) = \cos(x).$$

En particulier, la connaissance des valeurs de cos (resp. sin) sur $[0, \pi/2]$ détermine celle de sin (resp. cos) sur $[0, \pi/2]$.

b) **Formule des suppléments.** *Pour tout* $x \in \mathbb{R}$,

$$\cos(\pi - x) = -\cos(x) \quad \text{et} \quad \sin(\pi - x) = \sin(x).$$

c) **Relations de périodicité.** *Pour tout* $k \in \mathbb{Z}$ *et tout* $x \in \mathbb{R}$,

$$\cos(2k\pi + x) = \cos(x) \quad \text{et} \quad \sin(2k\pi + x) = \sin(x).$$

Ces dernières formules peuvent s'énoncer de la manière suivante: *les fonctions* cos *et* sin *sont périodiques de période* 2π, si on adopte la définition suivante.

Définition. Soit un nombre $T > 0$. Une fonction f définie sur $\mathbb{R}$ est *périodique* de *période* T si, pour tout $x \in \mathbb{R}$ et tout $k \in \mathbb{Z}$, on a $f(x + kT) = f(x)$.

Les conséquences de ces formules sont importantes.

Théorème 6.6.4.1 *Les nombres $k\pi$ avec $k \in \mathbb{Z}$ sont les seuls zéros de la fonction* sin. *L'ouvert de définition de la fonction* cotg *est donc égal à* $\mathbb{R} \setminus \{ k\pi : k \in \mathbb{Z}\}$.

Preuve. De fait, pour tout $x \in \mathbb{R}$, il existe $k \in \mathbb{Z}$ tel que $x - k\pi \in [0, \pi[$. De là, $\sin(x) = 0$ implique $\sin(x - k\pi) = 0$, donc $x = k\pi$.∎

Théorème 6.6.4.2 *Les nombres $(2k+1)\pi/2$ avec $k \in \mathbb{Z}$ sont les seuls zéros de la fonction* cos.

L'ouvert de définition de la fonction tg *est donc égal à* $\mathbb{R} \setminus \{ (2k+1)\pi/2 : k \in \mathbb{Z}\}$.∎

Proposition 6.6.4.3 *Si $T \in \mathbb{R}$ est tel que* $\cos(x+T) = \cos(x)$ *pour tout $x \in \mathbb{R}$, alors il existe $k \in \mathbb{Z}$ tel que $T = 2k\pi$.*

De même, si $T \in \mathbb{R}$ est tel que $\sin(x+T) = \sin(x)$ *pour tout $x \in \mathbb{R}$, alors il existe $k \in \mathbb{Z}$ tel que $T = 2k\pi$.*

Preuve. C'est une conséquence directe des formules

$$\begin{aligned} \cos(x+T) - \cos(x) &= -2\sin(T/2)\sin(x+T/2) = 0 \\ \sin(x+T) - \sin(x) &= 2\sin(T/2)\cos(x+T/2) = 0 \end{aligned}$$

valables pour tous x, $T \in \mathbb{R}$.∎

Remarque. Si la fonction f définie sur $\mathbb{R}$ est périodique de période T, il est bien évident que f est aussi périodique de période kT pour tout $k \in \mathbb{N}_0$. Par contre, il n'est pas du tout certain qu'il existe un nombre $T' \in]0, T[$ tel que f soit encore périodique de période T'. La proposition précédente affirme en fait que 2π est une période des fonctions cos et sin et qu'il n'existe pas de nombre $T \in]0, 2\pi[$ tel que cos (resp. sin) soit encore périodique de période T.

Proposition 6.6.4.4 *Si les nombres réels x et T sont tels que*

$$T > 0, \ \cos(x+T) = \cos(x) \quad \textit{et} \quad \sin(x+T) = \sin(x),$$

alors il existe $k \in \mathbb{N}_0$ tel que $T = 2k\pi$.

Preuve. De fait, on a

$$\cos(x+T) - \cos(x) = -2\sin(T/2)\sin(x+T/2) = 0$$

et

$$\sin(x+T) - \sin(x) = 2\sin(T/2)\cos(x+T/2) = 0$$

alors que les nombres $\sin(x+T/2)$ et $\cos(x+T/2)$ ne peuvent être simultanément nuls.∎

Cela étant, nous pouvons compléter notre connaissance des fonctions tg et cotg de la manière suivante.

a) La fonction tg appartient à $\mathrm{C}_\infty(\Omega)$ avec $\Omega = \mathbb{R} \setminus \{\, \pi/2 + k\pi : k \in \mathbb{Z}\}$. Elle vérifie la relation de périodicité $\mathrm{tg}(x+k\pi) = \mathrm{tg}(x)$ pour tout $x \in \Omega$. Il suffit donc de connaître la fonction tg sur $]-\pi/2, \pi/2[$ pour la connaître sur Ω. Or on a

$$\mathrm{tg}(0) = 0, \ \lim_{x \to (-\pi/2)^+} \mathrm{tg}(x) = -\infty, \ \lim_{x \to (\pi/2)^-} \mathrm{tg}(x) = +\infty, \ \mathrm{Dtg} = \frac{1}{\cos^2}.$$

Il s'agit donc d'une fonction strictement croissante sur $]-\pi/2, \pi/2[$ qui est concave sur $]-\pi/2, 0[$ et convexe sur $]0, \pi/2[$.

b) La fonction cotg appartient à $\mathrm{C}_\infty(\Omega')$ avec $\Omega' = \mathbb{R} \setminus \{\, k\pi : k \in \mathbb{Z}\}$. Son étude se déduit aussitôt de celle de la fonction tg car on a $\mathrm{cotg}(x) = \mathrm{tg}(\pi/2 - x)$ pour tout $x \in \Omega'$.

6.6.5 Retour à e^z

Pour tout $z \in \mathbb{C}$, nous avons

$$\mathrm{e}^z = \mathrm{e}^{\Re z + i\Im z} = \mathrm{e}^{\Re z}\mathrm{e}^{i\Im z} = \mathrm{e}^{\Re z}(\cos(\Im z) + i\sin(\Im z))$$

donc

$$\Re \mathrm{e}^z = \mathrm{e}^{\Re z}\cos(\Im z),\ \Im \mathrm{e}^z = \mathrm{e}^{\Re z}\sin(\Im z) \quad \text{et} \quad |\mathrm{e}^z| = \mathrm{e}^{\Re z}.$$

Proposition 6.6.5.1 *Les nombres complexes z et z' sont tels que $\mathrm{e}^z = \mathrm{e}^{z'}$ si et seulement s'il existe $k \in \mathbb{Z}$ tel que $z = z' + 2ik\pi$.*

Preuve. La condition est nécessaire. Des égalités

$$\mathrm{e}^{\Re z} = |\mathrm{e}^z| = \left|\mathrm{e}^{z'}\right| = \mathrm{e}^{\Re z'}$$

et du fait que e^x est une fonction strictement croissante, on tire déjà $\Re z = \Re z'$. Cela étant, de

$$\mathrm{e}^{\Re z}\cos(\Im z) = \Re \mathrm{e}^z = \Re \mathrm{e}^{z'} = \mathrm{e}^{\Re z'}\cos(\Im z')$$

et

$$\mathrm{e}^{\Re z}\sin(\Im z) = \Im \mathrm{e}^z = \Im \mathrm{e}^{z'} = \mathrm{e}^{\Re z'}\sin(\Im z'),$$

on tire ensuite

$$\cos(\Im z) = \cos(\Im z') \quad \text{et} \quad \sin(\Im z) = \sin(\Im z'),$$

d'où la conclusion par le dernier résultat du paragraphe précédent.

La condition est bien sûr suffisante, vu ce qui précède.∎

Le résultat suivant est très important.

Théorème 6.6.5.2 *Pour tout $z \in \mathbb{C}$ tel que $|z| = 1$, l'équation $\mathrm{e}^{i\theta} = z$ admet une solution et une seule dans $[0, 2\pi[$.*

Preuve. Les intervalles $[0, \pi/2[$, $[\pi/2, \pi[$, $[\pi, 3\pi/2[$ et $[3\pi/2, 2\pi[$ constituent une partition de $[0, 2\pi[$. Vu ce qui précède, dans ces intervalles, les fonctions cos et sin prennent respectivement des valeurs

$$> 0, \leq 0, < 0, \geq 0 \quad \text{et} \quad \geq 0, > 0, \leq 0, < 0.$$

Dès lors, selon les signes de $\Re z$ et de $\Im z$, on est immédiatement ramené à chercher une valeur de θ dans un et un seul de ces intervalles. Dans cet intervalle, la valeur d'un des nombres $\cos(x)$ et $\sin(x)$ détermine celle de l'autre et les fonctions $\cos(x)$ et $\sin(x)$ sont strictement monotones. Dès lors, il existe une solution et une seule.∎

Plus généralement, on voit de suite que *l'équation proposée admet une solution et une seule dans tout intervalle du type* $[a, a+2\pi[$ *avec* $a \in \mathbb{R}$.

On en déduit la représentation trigonométrique des nombres complexes.

Théorème 6.6.5.3 *Tout nombre complexe* z *différent de* 0 *peut s'écrire sous la forme* $z = \rho \mathrm{e}^{i\theta}$ *avec* $\rho > 0$ *et* $\theta \in \mathbb{R}$. *Dans cette écriture, on a* $\rho = |z|$ *et* θ *est défini à un multiple de* 2π *près par les équations*

$$\cos(\theta) = \frac{\Re z}{\rho} \quad \textit{et} \quad \sin(\theta) = \frac{\Im z}{\rho}.$$

Preuve. En effet, $z = \rho \mathrm{e}^{i\theta}$ a lieu si et seulement si on a $|z| = \rho$ et $\mathrm{e}^{i\theta} = z/\rho$ et il est clair que $|z/\rho| = 1$.∎

Définition. Bien sûr, ρ est le module de z; θ est appelé *argument* de z et n'est donc défini qu'à l'addition d'un multiple de 2π près.

La représentation trigonométrique des nombres complexes se prête particulièrement bien aux opérations suivantes sur les nombres complexes non nuls

a) *le conjugué de* $z = \rho \mathrm{e}^{i\theta}$ *est égal à* $\rho \mathrm{e}^{-i\theta}$.

b) *le produit de* $z = \rho \mathrm{e}^{i\theta}$ *par* $z' = \rho' \mathrm{e}^{i\theta'}$ *est égal à* $\rho\rho' \mathrm{e}^{i(\theta+\theta')}$.

c) *le quotient de* $z = \rho \mathrm{e}^{i\theta}$ *par* $z' = \rho' \mathrm{e}^{i\theta'}$ *est égal à* $(\rho/\rho')\mathrm{e}^{i(\theta-\theta')}$.

d) *étant donné* $m \in \mathbb{N}_0$, *les racines* m*-èmes de* $z = \rho \mathrm{e}^{i\theta}$ *sont données par*

$$\rho^{1/m} \mathrm{e}^{i(\theta+2k\pi)/m} \quad \textit{avec } k \in \{0, \ldots, m-1\}.$$

De fait, ces nombres sont des solutions de l'équation $z'^m = z$ et ce sont les seules, vu ce qui précède ou vu la théorie des zéros des polynômes.∎

6.6.6 Retour aux séries

Le critère d'Abel admet le corollaire suivant.

Proposition 6.6.6.1 *Si la suite* $a_m \in \mathbb{R}$ *décroît vers* 0 *et si* $\theta \in \mathbb{R}$ *n'est pas un multiple de* 2π, *alors les séries*

$$\sum_{m=1}^{\infty} a_m \cos(m\theta), \quad \sum_{m=1}^{\infty} a_m \sin(m\theta) \quad \textit{et} \quad \sum_{m=1}^{\infty} a_m \mathrm{e}^{im\theta}$$

convergent et, pour tout $M \in \mathbb{N}_0$, *on a*

$$\left| \sum_{m=M}^{\infty} a_m \cos(m\theta) \right|, \quad \left| \sum_{m=M}^{\infty} a_m \sin(m\theta) \right|, \quad \left| \sum_{m=M}^{\infty} a_m \mathrm{e}^{im\theta} \right| \leq \frac{a_M}{|\sin(\theta/2)|}.$$

Preuve. De fait, pour tous p, $q \in \mathbb{N}_0$ tels que $p \leq q$, on a

$$\begin{aligned}\left|\sum_{m=p}^{q}\left\{\begin{array}{c}\cos\\ \sin\end{array}\right\}(m\theta)\right| &= \left|\left\{\begin{array}{c}\Re\\ \Im\end{array}\right\}(\mathrm{e}^{ip\theta}+\ldots+\mathrm{e}^{iq\theta})\right|\\ &\leq \left|\mathrm{e}^{ip\theta}+\ldots+\mathrm{e}^{iq\theta}\right| \leq \frac{\left|\mathrm{e}^{ip\theta}-\mathrm{e}^{i(q+1)\theta}\right|}{\left|1-\mathrm{e}^{i\theta}\right|}\\ &\leq \frac{\left|\mathrm{e}^{i(p-1/2)\theta}-\mathrm{e}^{i(q+1/2)\theta}\right|}{\left|\mathrm{e}^{-i\theta/2}-\mathrm{e}^{i\theta/2}\right|} \leq \frac{1}{\left|\sin(\theta/2)\right|}.\blacksquare\end{aligned}$$

Ce résultat permet d'étudier la convergence des séries qui dépendent d'un paramètre complexe. Reprenons, par exemple, le dernier exercice du paragraphe 2.5.8.

Exercice. Etudier la convergence de la série numérique dont le terme général est donné par

$$z_m = \frac{(c-2)^m}{2^m c^m \sqrt{2m+1}},$$

pour toutes les valeurs du paramètre complexe c.

Suggestion. Pour assurer un sens au terme général, nous devons éliminer la valeur $c = 0$ du paramètre. De plus, il est clair que, pour $c = 2$, la série converge absolument. Cela étant, limitons-nous aux éléments de $\mathbb{C} \setminus \{0, 2\}$.

Vu la forme de $|z_m|$, recourrons au critère du quotient. Comme on a

$$\frac{|z_{m+1}|}{|z_m|} = \frac{|c-2|}{2\,|c|}\frac{\sqrt{2m+1}}{\sqrt{2m+3}} \longrightarrow \left(\frac{|c-2|}{2\,|c|}\right)^{-},$$

nous avons déjà les résultats suivants:
$|c-2|/(2\,|c|) < 1$: la série converge,
$|c-2|/(2\,|c|) > 1$: la série diverge car son terme général ne tend pas vers 0.

Il reste à envisager les cas où $|c-2| = 2\,|c|$. Si $c-2$ est égal à $2c$, la série s'écrit

$$\sum_{m=1}^{\infty}\frac{1}{\sqrt{2m+1}}$$

et diverge, vu une application aisée du critère de Riemann. Si ce n'est pas le cas, il existe $\theta \in]0, 2\pi[$ tel que $c - 2 = 2c\,\mathrm{e}^{i\theta}$ et dès lors, la série s'écrit

$$\sum_{m=1}^{\infty}\frac{\mathrm{e}^{im\theta}}{\sqrt{2m+1}}.$$

Comme la suite $(2m+1)^{-1/2}$ décroît vers 0, une application directe de la proposition précédente établit que cette série converge, donc est semi-convergente car sa série des modules diverge.

Comme $|c-2| =$ (resp. $<$; $>$) $2\,|c|$ avec $c = r+is \in \mathbb{C}$ a lieu si et seulement si la même relation a lieu entre les carrés des deux membres, c'est-à-dire si et seulement si on a $16/9 =$ (resp. $<$; $>$) $(r+2/3)^2+(s-0)^2$, on obtient de suite une description graphique de ces résultats.

Si D est le disque fermé de centre $(-2/3, 0)$ et de rayon $4/3$,
a) la série converge absolument pour tout $c \in \mathbb{C} \setminus D$,
b) la série diverge pour tout $c \in D^\circ \setminus \{0\}$,
c) la série diverge pour $c = -2$,
d) la série est semi-convergente pour tout $c \in D^\bullet \setminus \{-2\}$,
e) la série n'est pas définie pour $c = 0$.

6.6.7 Fonctions hyperboliques ou circulaires de z

On peut aussi introduire les fonctions suivantes sur $\mathbb{C}$

$$\left\{\begin{array}{l} \cos(z) = \dfrac{e^{iz}+e^{-iz}}{2} \\ \sin(z) = \dfrac{e^{iz}-e^{-iz}}{2i} \end{array}\right. \quad \text{et} \quad \left\{\begin{array}{l} \mathrm{ch}(z) = \dfrac{e^{z}+e^{-z}}{2} \\ \mathrm{sh}(z) = \dfrac{e^{z}-e^{-z}}{2} \end{array}\right.$$

L'étude de ces fonctions de $z \in \mathbb{C}$ montre qu'elles ne sont pas indépendantes et permet de comprendre la raison profonde de l'analogie entre les calculs "trigonométrique" et "hyperbolique".

Cependant elles ne sont pas d'un usage courant et leurs propriétés sont immédiates. Aussi nous n'allons pas aborder leur étude ici.

6.7 Fonctions élémentaires inverses

6.7.1 Logarithme népérien

Définition. Comme la fonction réelle e^x appartient à $C_\infty(\mathbb{R})$ et est à valeurs strictement positives, strictement croissante et telle que

$$e^0 = 1, \quad e^1 = e, \quad De^x = e^x, \quad \lim_{x\to+\infty} e^x = +\infty \ \text{ et } \ \lim_{x\to-\infty} e^x = 0^+,$$

elle admet une fonction inverse appelée *logarithme népérien* et notée ln.

Les propriétés de cette fonction se déduisent directement de celles de e^x.

Théorème 6.7.1.1 *La fonction* $\ln(x)$
a) *est définie sur* $]0,+\infty[$,
b) *est réelle,*
c) *est telle que* $\ln(1)=0$ *et* $\ln(\mathrm{e})=1$. *Plus généralement, on a*

$$\begin{aligned}\ln(\mathrm{e}^x) &= x, \quad \forall x \in \mathbb{R},\\ \mathrm{e}^{\ln(x)} &= x, \quad \forall x > 0,\end{aligned}$$

d) *est strictement croissante,*
e) *appartient à* $\mathrm{C}_\infty(]0,+\infty[)$ *et vérifie*

$$\mathrm{D}\ln(x) = \frac{1}{x},$$

f) *est telle que*

$$\lim_{x\to 0^+} \ln(x) = -\infty \quad \textit{et} \quad \lim_{x\to+\infty} \ln(x) = +\infty,$$

g) *est concave,*
h) *est telle que*

$$\ln(xy) = \ln(x) + \ln(y), \quad \forall x, y \in \left]0,+\infty\right[,$$

i) *est telle que*

$$\lim_{x\to 0^+} x\ln^m(x) = 0 \quad \textit{et} \quad \lim_{x\to+\infty} \frac{\ln^m(x)}{x} = 0^+$$

pour tout $m \in \mathbb{N}_0$.

Preuve. Les propriétés a) à f) découlent aussitôt du théorème de la fonction inverse.

g) De fait, on a $\mathrm{D}^2\ln(x) = -x^{-2}$.

h) De fait, on a

$$\mathrm{e}^{\ln(xy)} = xy = \mathrm{e}^{\ln(x)}\mathrm{e}^{\ln(y)} = \mathrm{e}^{\ln(x)+\ln(y)}, \quad \forall x, y \in \left]0,+\infty\right[.$$

i) De fait, on a

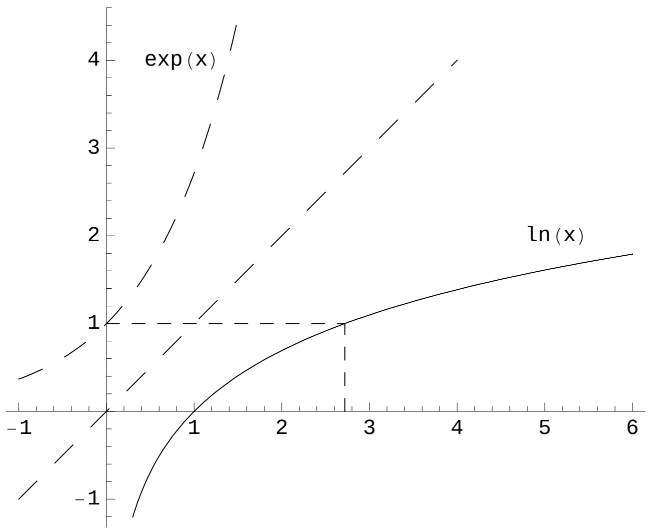
$$\lim_{x\to+\infty} \frac{\ln^m(x)}{x} = \lim_{x\to+\infty} \ln^m(x)\cdot \mathrm{e}^{-\ln(x)} = \lim_{y\to+\infty} y^m\mathrm{e}^{-y} = 0$$

et

$$\lim_{x\to 0^+} x\ln^m(x) = (-1)^m \lim_{y\to+\infty} \frac{\ln^m(y)}{y} = 0. \blacksquare$$

On exprime la propriété i) de la manière suivante: *la fonction ln*(x) *est dominée par toute puissance antagoniste de* x.

Le graphique de la fonction ln se présente donc sous la forme suivante où le graphique de la fonction e est représenté en pointillés.



Voici une application intéressante de la fonction ln.

Comparaison de moyennes remarquables. Pour tout $x > 0$, on a bien sûr

$$\frac{\mathrm{sh}(x)}{\sqrt{\mathrm{sh}^2(x)+\mathrm{ch}^2(x)}} < \mathrm{th}(x) < x < \mathrm{sh}(x) < \frac{1}{2}\mathrm{sh}(2x)$$

donc

$$\frac{\sqrt{\mathrm{sh}^2(x)+\mathrm{ch}^2(x)}}{\mathrm{sh}(x)} > \coth(x) > \frac{1}{x} > \frac{1}{\mathrm{sh}(x)} > \frac{2}{\mathrm{sh}(2x)}.$$

Dès lors pour tous a, $b \in]0,+\infty[$ tels que $a < b$, il vient

$$\frac{\sqrt{2(b^2+a^2)}}{b-a} > \frac{b+a}{b-a} > \frac{2}{\ln(b)-\ln(a)} > \frac{2\sqrt{ab}}{b-a} > \frac{4}{b/a-a/b}$$

en considérant les inégalités précédentes en $x = \ln(b/a)^{1/2}$. Finalement, en multipliant par $(b-a)/2$, il vient

$$\frac{2}{1/a+1/b} < \sqrt{ab} < \frac{b-a}{\ln(b)-\ln(a)} < \frac{a+b}{2} < \sqrt{\frac{a^2+b^2}{2}},$$

c'est-à-dire la comparaison entre les moyennes harmonique, géométrique, logarithmique, arithmétique et quadratique.

Quelques inégalités remarquables. a) *Pour tout* $n \in \mathbb{N}_0$, *tous* $r_1, \ldots, r_n > 0$ *tels que* $\sum_{k=1}^n r_k = 1$ *et tous* $x_1, \ldots, x_n > 0$, *on a*

$$x_1^{r_1} \ldots x_n^{r_n} \leq r_1 x_1 + \cdots + r_n x_n.$$

De fait, la fonction ln étant concave sur $]0, +\infty[$, il vient

$$\ln(r_1 x_1 + \ldots + r_n x_n) \geq r_1 \ln(x_1) + \cdots + r_n \ln(x_n),$$

d'où la conclusion en recourant à la fonction exponentielle.

b) **Inégalité de Hölder.** *Pour tous* p, $q > 1$ *tels que* $1/p + 1/q = 1$ *et tous* x, $y \in \mathbb{R}^n$, *on a*

$$|< x, y >| \leq \sum_{k=1}^n |x_k y_k| \leq \left(\sum_{k=1}^n |x_k|^p \right)^{1/p} \cdot \left(\sum_{k=1}^n |y_k|^q \right)^{1/q}.$$

Posons $X = \sum_{k=1}^n |x_k|^p$ et $Y = \sum_{k=1}^n |y_k|^q$. Si on a $X = 0$ ou $Y = 0$, c'est trivial. Sinon, pour tout $k \in \{1, \ldots, n\}$ tel que $x_k y_k \neq 0$, l'inégalité précédente donne

$$\left(\frac{|x_k|^p}{X} \right)^{1/p} \cdot \left(\frac{|y_k|^q}{Y} \right)^{1/q} \leq \frac{1}{p} \frac{|x_k|^p}{X} + \frac{1}{q} \frac{|y_k|^q}{Y},$$

formule valable aussi si $x_k y_k = 0$. Après sommation sur k, il vient

$$\sum_{k=1}^n \frac{|x_k|\,|y_k|}{X^{1/p} Y^{1/q}} \leq \frac{1}{p} + \frac{1}{q} = 1.$$

(Pour $p = q = 2$, on retrouve l'inégalité de Cauchy-Schwarz.)

c) **Inégalité de Minkowski.** *Pour tout* $p \geq 1$ *et tous* $x, y \in \mathbb{R}^n$, *on a*

$$\left(\sum_{k=1}^n |x_k + y_k|^p \right)^{1/p} \leq \left(\sum_{k=1}^n |x_k|^p \right)^{1/p} + \left(\sum_{k=1}^n |y_k|^p \right)^{1/p}.$$

Nous avons déjà établi cette inégalité pour $p = 1$.

Si $p > 1$, introduisons d'abord $q > 1$ défini par $1/p + 1/q = 1$ et posons $Z = \sum_{k=1}^n |x_k + y_k|^p$. Le cas $Z = 0$ est trivial. Supposons $Z > 0$. Il vient successivement, en recourant à l'inégalité de Hölder,

$$\begin{aligned} Z &= \sum_{k=1}^{n} |x_k + y_k| \, |x_k + y_k|^{p-1} \\ &\leq \sum_{k=1}^{n} |x_k| \, |x_k + y_k|^{p-1} + \sum_{k=1}^{n} |y_k| \, |x_k + y_k|^{p-1} \\ &\leq \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p} \cdot \left(\sum_{k=1}^{n} |x_k + y_k|^{(p-1)q} \right)^{1/q} \\ &\quad + \left(\sum_{k=1}^{n} |y_k|^p \right)^{1/p} \cdot \left(\sum_{k=1}^{n} |x_k + y_k|^{(p-1)q} \right)^{1/q} \\ &\leq \left(\left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p} + \left(\sum_{k=1}^{n} |y_k|^p \right)^{1/p} \right) \cdot Z^{1/q}. \end{aligned}$$

d) **Inégalité de Jensen**. *Pour tous p, q tels que $0 < p < q$ et tout $x \in \mathbb{R}^n$, on a*

$$\left(\sum_{k=1}^{n} |x_k|^q \right)^{1/q} \leq \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p}.$$

Posons $X^p = \sum_{k=1}^{n} |x_k|^p$. Si $X = 0$, c'est trivial. Sinon il vient $|x_k| / X \leq 1$ pour tout $k \in \{1, \ldots, n\}$, donc

$$\sum_{k=1}^{n} \left(\frac{|x_k|}{X} \right)^q \leq \sum_{k=1}^{n} \left(\frac{|x_k|}{X} \right)^p = 1.$$

6.7.2 Fonction X^z

Définition. Pour tout $X \in \,]0, +\infty[$ et tout $z \in \mathbb{C}$, on pose

$$X^z = \mathrm{e}^{z \ln(X)}.$$

Dans ce symbole X^z, X porte le nom d'*argument* et z celui d'*exposant*.

Dans ces conditions, nous avons

$$\begin{aligned} X^z &= \mathrm{e}^{z \ln(X)} = \mathrm{e}^{\Re z \cdot \ln(X) + i \Im z \cdot \ln(X)} \\ &= \mathrm{e}^{\Re z \cdot \ln(X)} (\cos(\Im z \cdot \ln(X)) + i \sin(\Im z \cdot \ln(X))). \end{aligned}$$

Remarque. La notation X^z a déjà été introduite auparavant dans d'autres circonstances. Nous allons montrer à présent que la définition de X^z recouvre ces différents cas.

a) Si $z = m$ appartient à $\mathbb{N}_0$, on retrouve la puissance entière de X vu la formule $\mathrm{e}^{\ln(X)} = X$ et la propriété fondamentale de l'exponentielle.

b) Si z est égal à $1/m$ avec $m \in \mathbb{N}_0$, on retrouve la racine m-ème de $X > 0$.

c) En combinant a) et b), on retrouve les puissances fractionnaires rationnelles de $X > 0$.

d) Si on a $X = \mathrm{e}$, on retrouve l'exponentielle de $z \in \mathbb{C}$.

Définitions. Le symbole X^z permet de définir deux fonctions d'une seule variable:

a) la *puissance d'exposant complexe* X^{z_0} si X varie dans $]0, +\infty[$, $z_0 \in \mathbb{C}$ étant fixé;

b) l'*exponentielle d'argument positif* X_0^z si z varie dans $\mathbb{C}, X_0 \in\,]0, +\infty[$ étant fixé.

Ces deux fonctions s'étudient simultanément. Cependant, dans la suite, nous n'utiliserons guère que X^{z_0} car X_0^z apparaît plutôt comme une fonction de fonction $X_0^z = \mathrm{e}^{z \ln(X_0)}$.

Etudions tout d'abord les propriétés générales de X^z.

a) *X^z est une fonction définie sur $]0, +\infty[\times \mathbb{C}$.*

b) *On a $(X^z)^- = X^{\overline{z}}$ et $|X^z| = X^{\Re z}$ pour tout $(x, z) \in\,]0, +\infty[\times \mathbb{C}$.*

De fait, on a successivement

$$\begin{array}{rcccl} \overline{X^z} & = & \mathrm{e}^{\overline{z \ln(X)}} & = \mathrm{e}^{\overline{z} \ln(X)} & = X^{\overline{z}}, \\ |X^z| & = & \left|\mathrm{e}^{\Re z \cdot \ln(X)} \mathrm{e}^{i \Im z \cdot \ln(X)}\right| & & = X^{\Re z}. \blacksquare \end{array}$$

c) *Pour tout $X > 0$ et tous $z,\, z' \in \mathbb{C}$, on a $X^z X^{z'} = X^{z+z'}$ donc $X^{-z} = 1/X^z$.*

De fait, il vient successivement

$$X^z X^{z'} = \mathrm{e}^{z \ln(X)} \mathrm{e}^{z' \ln(X)} = \mathrm{e}^{(z+z') \ln(X)} = X^{z+z'}. \blacksquare$$

d) *Pour tous $X,\, Y > 0$ et tout $z \in \mathbb{C}$, on a $X^z Y^z = (XY)^z$.*

De fait, il vient successivement

$$X^z Y^z = \mathrm{e}^{z \ln(X)} \mathrm{e}^{z \ln(Y)} = \mathrm{e}^{z(\ln(X) + \ln(Y))} = (XY)^z. \blacksquare$$

e) Devant un symbole du type $(X^z)^{z'}$, on doit être très prudent car il n'est défini que dans les deux cas suivants.

i) *Pour tout $X > 0$, tout $r \in \mathbb{R}$ et tout $z' \in \mathbb{C}$, X^r est un nombre strictement positif tel que $(X^r)^{z'} = X^{rz'}$.*

En particulier, pour tout $X > 0$ et tout $r \in \mathbb{R}_0$, on a $(X^{1/r})^r = X$.
ii) *Pour tout $X > 0$, tout $z \in \mathbb{C}$ et tout $n \in \mathbb{N}$, X^z est un nombre complexe tel que $(X^z)^n = X^{zn}$.*

i) De $X^r = \mathrm{e}^{r\ln(X)}$, on tire que X^r est strictement positif et dès lors, il vient

$$(X^r)^{z'} = \mathrm{e}^{z'\ln(\mathrm{e}^{r\ln(X)})} = \mathrm{e}^{rz'\ln(X)} = X^{rz'}.$$

ii) De fait, dans ce cas, X^z est un nombre complexe tel que

$$(X^z)^n = X^z \cdots X^z = \mathrm{e}^{z\ln(X)} \cdots \mathrm{e}^{z\ln(X)} = \mathrm{e}^{nz\ln(X)} = X^{nz}.$$

Remarque. La formule "$(X^z)^{z'} = X^{zz'}$ pour tout $X > 0$ et tous $z, z' \in \mathbb{C}$" est fausse. De fait, elle entraînerait les égalités

$$1 = 1^\alpha = (\mathrm{e}^{2i\pi})^\alpha \text{ “=” } \mathrm{e}^{2i\alpha\pi} = \cos(2\alpha\pi) + i\sin(2\alpha\pi)$$

pour tout $\alpha \in \mathbb{R}$!

f) *La fonction X^z appartient à $\mathrm{C}_\infty(]0, +\infty[\times \mathbb{R}^2)$ et est telle que*

$$\mathrm{D}_x X^z = -i\mathrm{D}_y X^z = \mathrm{D}_z X^z = X^z \ln(X) \quad \text{et} \quad \mathrm{D}_X X^z = zX^{z-1}.$$

En particulier,
i) *pour tout $X_0 \in]0, +\infty[$ fixé, la fonction X_0^z appartient à $\mathrm{C}_\infty(\mathbb{R}^2)$ et est telle que $\mathrm{D}_x X_0^z = -i\mathrm{D}_y X_0^z = X_0^z \ln(X_0)$.*
ii) *pour tout $z_0 \in \mathbb{C}$ fixé, la fonction X^{z_0} appartient à $\mathrm{C}_\infty(]0, +\infty[)$ et est telle que $\mathrm{D}_X X^{z_0} = z_0 X^{z_0-1}$.*

C'est immédiat à partir du théorème de dérivation des fonctions composées.

6.7.3 Puissance d'exposant x et exponentielle d'argument X

Passons au cas particulier suivant:

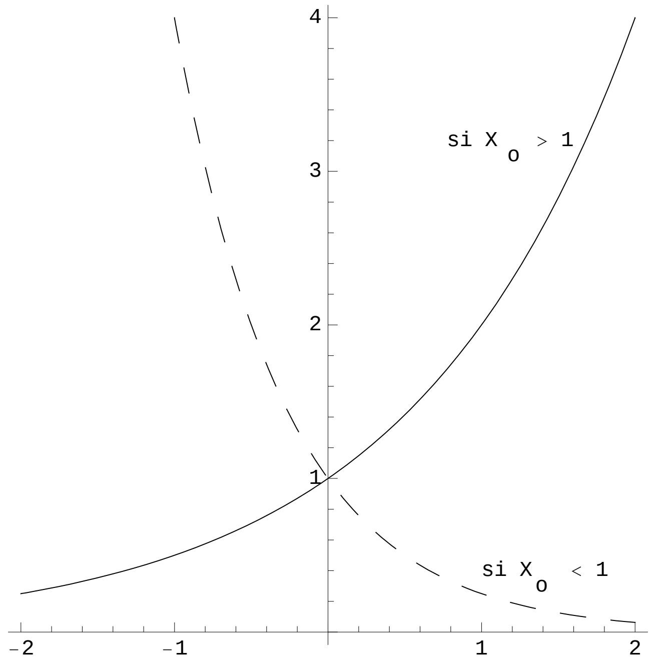
$$X^x = \mathrm{e}^{x\ln(X)}, \quad \forall X \in]0, +\infty[, \forall x \in \mathbb{R}.$$

Bien sûr, il s'agit d'une fonction appartenant à $\mathrm{C}_\infty(]0, +\infty[\times \mathbb{R})$ qui jouit des propriétés suivantes: *pour tous $X, Y > 0$ et tous $x, x' \in \mathbb{R}$, on a*
a) $X^x > 0$.
b) $X^x X^{x'} = X^{x+x'}$, $X^x Y^x = (XY)^x$ *et* $(X^x)^{x'} = X^{xx'}$.
c) $\ln(X^x) = x\ln(X)$.
d) $X^{\ln(x)} = x^{\ln(X)}$ *si, en outre, on a $x > 0$.* De fait, il vient alors

$$X^{\ln(x)} = \mathrm{e}^{\ln(x)\cdot\ln(X)} = \mathrm{e}^{\ln(X)\cdot\ln(x)} = x^{\ln(X)}. \blacksquare$$

Fonction X_0^x

Si on fixe $X_0 > 0$, on obtient une fonction X_0^x définie sur $\mathbb{R}$ qui jouit des propriétés suivantes.



Pour tout $X_0 > 0$ fixé, X_0^x est une fonction

a) *strictement positive sur $\mathbb{R}$;*

b) *appartenant à $\mathrm{C}_\infty(\mathbb{R})$ et telle que $\mathrm{D}_x X_0^x = X_0^x \ln(X_0)$;*

c) *strictement croissante si $X_0 > 1$, strictement décroissante si $0 < X_0 < 1$;*

d) *convexe;*

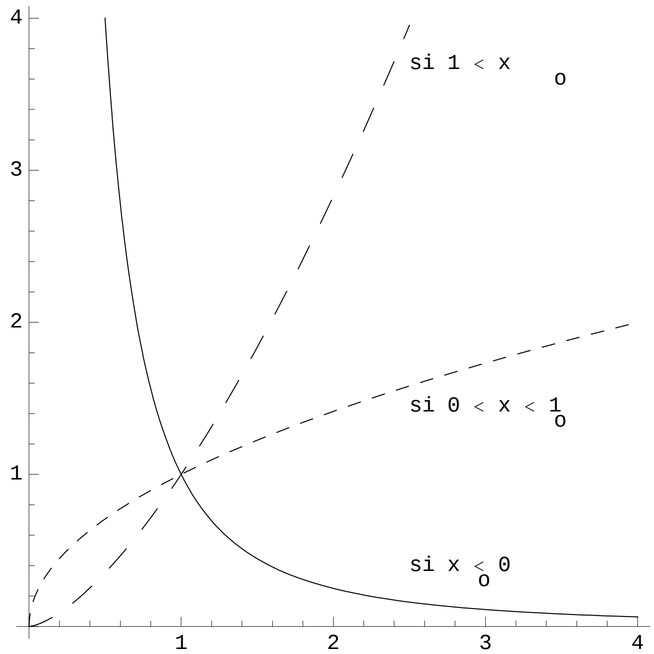
e) *telle que*

$$\lim_{x\to+\infty} X_0^x = \begin{cases} +\infty & si \quad X_0 > 1, \\ 0^+ & si \quad 0 < X_0 < 1, \end{cases}$$

$$\lim_{x\to-\infty} X_0^x = \begin{cases} 0^+ & si \quad X_0 > 1, \\ +\infty & si \quad 0 < X_0 < 1. \end{cases}$$

Fonction X^{x_0}

Si on fixe $x_0 \in \mathbb{R}$, on obtient une fonction X^{x_0} définie sur $]0, +\infty[$ qui jouit des propriétés suivantes.



Pour tout $x_0 \in \mathbb{R}$ fixé, X^{x_0} est une fonction

a) *strictement positive sur $]0, +\infty[$;*

b) *appartenant à $C_\infty(]0, +\infty[)$ et telle que $D_X X^{x_0} = x_0 X^{x_0 - 1}$;*

c) *strictement croissante si $x_0 > 0$, strictement décroissante si $x_0 < 0$;*

d) *convexe si on a $x_0 \in]-\infty, 0[\cup]1, +\infty[$ et concave sinon;*

e) *telle que*

$$\lim_{X \to +\infty} X^{x_0} = \begin{cases} +\infty & \text{si} \quad x_0 > 0, \\ 0^+ & \text{si} \quad x_0 < 0, \end{cases}$$

$$\lim_{X \to 0^+} X^{x_0} = \begin{cases} 0^+ & \text{si} \quad x_0 > 0, \\ +\infty & \text{si} \quad x_0 < 0. \end{cases}$$ ∎

Définition. Dans le cas $x_0 > 0$, on convient de prolonger la fonction X^{x_0} par la valeur 0 en $X = 0$ et de continuer à noter X^{x_0} cette fonction qui appartient alors en outre à $C_0([0, +\infty[)$.

Le tableau suivant donne la valeur de la limite

$$\lim_{x \to \xi,\, X \to \Xi} X^x = \lim_{x \to \xi,\, X \to \Xi} e^{x \ln(X)}$$

où ξ appartient à $\mathbb{R}$ ou désigne $+\infty$ ou $-\infty$ et où Ξ appartient à $]0, +\infty[$ ou désigne 0^+ ou $+\infty$. Elles sont toutes immédiates.

	$-\infty$	$x<0$	0	$0<x<1$	1	$1<x$	$+\infty$
0	$+\infty$	$+\infty$		0	0	0	0
$0<X<1$	$+\infty$	X^x	1	X^x	X	X^x	0
1		1	1	1	1	1	
$1<X$	0	X^x	1	X^x	X	X^x	$+\infty$
$+\infty$	0	0		$+\infty$	$+\infty$	$+\infty$	$+\infty$

6.7.4 Logarithme de base $X \in \left]0,1\right[$ ou $X \in \left]1,+\infty\right[$

Définition. Si X appartient à $]0,1[$ ou à $]1,+\infty[$, on appelle *logarithme de base* X la fonction $\log_X$ définie sur $]0,+\infty[$ par

$$\log_X(x) = \frac{\ln(x)}{\ln(X)}, \quad \forall x \in \left]0,+\infty\right[.$$

Le logarithme népérien apparaît alors comme étant le logarithme de base e.

Le logarithme de base X n'est donc pas une nouvelle fonction élémentaire: ses propriétés se déduisent immédiatement de celles de la fonction $\ln(x)$.

La formule suivante lie de suite les logarithmes de bases X et Y

$$\log_X(x) = \frac{\ln(x)}{\ln(X)} = \frac{\ln(Y)}{\ln(X)} \cdot \frac{\ln(x)}{\ln(Y)} = \log_X(Y) \cdot \log_Y(x).$$

L'introduction de cette notion de logarithme de base X est justifiée par le fait que $\log_X(x)$ *est l'exposant qu'il faut donner à* X *pour obtenir* x vu que

$$X^{\log_X(x)} = \mathrm{e}^{\frac{\ln(x)}{\ln(X)}\ln(X)} = \mathrm{e}^{\ln(x)} = x.$$

6.7.5 Fonctions hyperboliques inverses

La fonction arcsh

Rappelons les propriétés suivantes de la fonction sh:

a) $\mathrm{sh} \in \mathrm{C}_\infty(\mathbb{R})$ et vérifie $\mathrm{Dsh} = \mathrm{ch}$,

b) sh est une fonction réelle et strictement croissante,

c) $\mathrm{sh}(0) = 0$,

d) $\lim_{x\to-\infty} \mathrm{sh}(x) = -\infty$ et $\lim_{x\to+\infty} \mathrm{sh}(x) = +\infty$.

Définition. Dès lors, la fonction sh admet une fonction inverse, appelée *arc sinus hyperbolique* et notée arcsh, qui jouit des propriétés suivantes.

Théorème 6.7.5.1 *La fonction* arcsh
a) *est définie, réelle et strictement croissante sur* $\mathbb{R}$,
b) *vérifie les égalités*

$$\operatorname{arcsh}(\operatorname{sh}(x)) = x \quad et \quad \operatorname{sh}(\operatorname{arcsh}(x)) = x, \quad \forall x \in \mathbb{R},$$

c) $\operatorname{arcsh}(0) = 0$,
d) *appartient à* $\mathrm{C}_\infty(\mathbb{R})$ *et on a*

$$\operatorname{Darcsh}(x) = \frac{1}{\sqrt{1+x^2}},$$

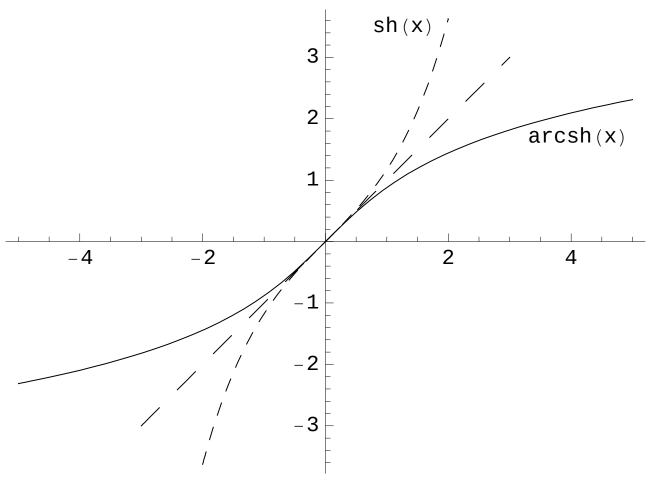
e) *est convexe sur* $]-\infty, 0[$ *et concave sur* $]0, +\infty[$,
f) *donne lieu aux relations*

$$\lim_{x\to-\infty} \operatorname{arcsh}(x) = -\infty \quad et \quad \lim_{x\to+\infty} \operatorname{arcsh}(x) = +\infty.$$

Preuve. Tout est conséquence directe du théorème de la fonction inverse et ne pose aucun problème sauf la dérivée pour laquelle on note que

$$\operatorname{Darcsh}(x) = \frac{1}{[\operatorname{Dsh}(y)]_{y=\operatorname{arcsh}(x)}} = \frac{1}{\operatorname{ch}(\operatorname{arcsh}(x))}, \quad \forall x \in \mathbb{R}.$$

En utilisant l'identité $\operatorname{ch}^2 - \operatorname{sh}^2 = 1$, on obtient de suite $\operatorname{ch}^2(\operatorname{arcsh}(x)) = 1 + x^2$ pour tout $x \in \mathbb{R}$, d'où la valeur de la dérivée car $\operatorname{ch}(y)$ est un nombre supérieur ou égal à 1 pour tout $y \in \mathbb{R}$.∎



Remarque. En fait, la fonction arcsh n'est pas à proprement parler une fonction élémentaire: de

$$\operatorname{arcsh}(x) = X \Leftrightarrow \operatorname{sh}(X) = x \Leftrightarrow \mathrm{e}^X - \mathrm{e}^{-X} = 2x,$$

on tire en effet la relation

$$\operatorname{arcsh}(x) = X \Leftrightarrow \mathrm{e}^{2X} - 2x\mathrm{e}^X - 1 = 0.$$

Or le second membre est une équation du second degré en l'inconnue e^X et admet pour solution les valeurs $x + \sqrt{x^2+1}$ et $x - \sqrt{x^2+1}$. Nous avons donc

$$\operatorname{arcsh}(x) = \ln(x + \sqrt{x^2+1}), \quad \forall x \in \mathbb{R},$$

car e^X est strictement positif pour tout $X \in \mathbb{R}$.

La fonction arcch

Rappelons les propriétés suivantes de la fonction ch:

a) $\operatorname{ch} \in \mathrm{C}_0([0,+\infty[) \cap \mathrm{C}_\infty(]0,+\infty[)$ et vérifie $\mathrm{D}\operatorname{ch} = \operatorname{sh}$,

b) ch est une fonction réelle et strictement croissante sur $[0,+\infty[$,

c) $\operatorname{ch}(0) = 1$,

d) $\lim_{x\to+\infty} \operatorname{ch}(x) = +\infty$.

Définition. Dès lors, la fonction ch admet une fonction inverse sur $[0,+\infty[$, appelée *arc cosinus hyperbolique* et notée arcch, qui jouit des propriétés suivantes.

Théorème 6.7.5.2 *La fonction* arcch

a) *est définie, réelle et strictement croissante sur* $[1,+\infty[$,

b) *vérifie les égalités*

$$\begin{aligned} \operatorname{arcch}(\operatorname{ch}(x)) &= x, \quad \forall x \in [0,+\infty[\,, \\ \operatorname{ch}(\operatorname{arcch}(x)) &= x, \quad \forall x \in [1,+\infty[\,, \end{aligned}$$

c) *appartient à* $\mathrm{C}_0([1,+\infty[)$ *et à* $\mathrm{C}_\infty(]1,+\infty[)$*; de plus, on a*

$$\mathrm{D}\operatorname{arcch}(x) = \frac{1}{\sqrt{x^2-1}}.$$

d) *est concave,*

e) *donne lieu aux relations*

$$\operatorname{arcch}(1) = 0 \quad \textit{et} \quad \lim_{x\to+\infty} \operatorname{arcch}(x) = +\infty.$$

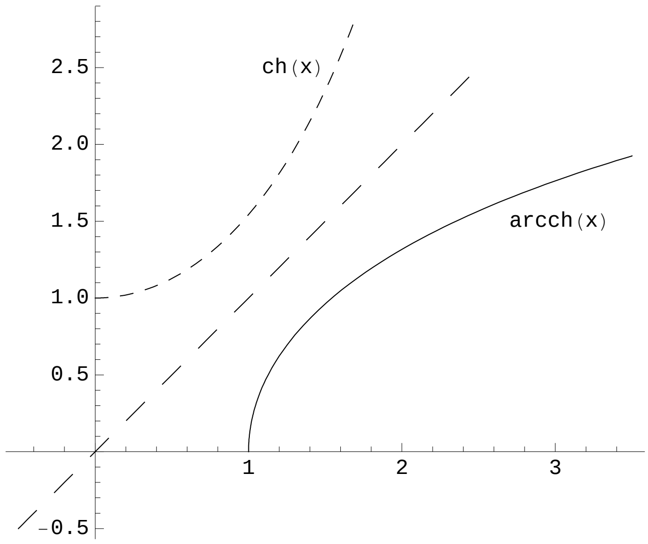
Preuve. Seule l'obtention de la dérivée pose un petit problème. Vu le théorème de la fonction inverse, il vient

$$\mathrm{Darcch}(x) = \frac{1}{\mathrm{sh}(\mathrm{arcch}(x))}, \quad \forall x \in \,]1, +\infty[\,.$$

Pour conclure, il suffit alors de noter que

$$\mathrm{sh}^2(\mathrm{arcch}(x)) = x^2 - 1, \quad \forall x \in \,]1, +\infty[\,,$$

et que la dérivée est une fonction positive car arcch est une fonction dérivable et strictement croissante sur l'intervalle $]1, +\infty[$.∎



Remarque. En procédant comme pour la fonction arcsh, on obtient de suite

$$\mathrm{arcch}(x) = X \Leftrightarrow \mathrm{e}^{2X} - 2x\mathrm{e}^X + 1 = 0,$$

d'où on tire

$$\mathrm{arcch}(x) = \ln(x + \sqrt{x^2 - 1}), \quad \forall x \in [1, +\infty[\,,$$

et la fonction arcch n'est pas à proprement parler une nouvelle fonction élémentaire. Pour le choix de la solution de l'équation du second degré, il suffit de noter que

$$\mathrm{arcch}(x) > 0, \quad \forall x > 1$$

et

$$x + \sqrt{x^2 - 1} > 1, \quad \forall x > 1; \quad x - \sqrt{x^2 - 1} < 1, \quad \forall x > 1.$$

La fonction arcth

En procédant de même à partir de la fonction th, on introduit la fonction suivante.

Définition. La fonction th admet une fonction inverse, appelée *arc tangente hyperbolique* et notée arcth, qui jouit des propriétés suivantes.

Théorème 6.7.5.3 *La fonction* arcth
a) *est définie, réelle et strictement croissante sur* $]-1, 1[$,
b) *vérifie les égalités*

$$\begin{aligned} \operatorname{arcth}(\operatorname{th}(x)) &= x, \quad \forall x \in \mathbb{R}, \\ \operatorname{th}(\operatorname{arcth}(x)) &= x, \quad \forall x \in \,]-1, 1[\,, \end{aligned}$$

c) $\operatorname{arcth}(0) = 0$,
d) *appartient à* $\mathrm{C}_\infty(]-1, 1[)$ *et on a* $\operatorname{Darcth}(x) = 1/(1 - x^2)$,
e) *est concave sur* $]-1, 0[$ *et convexe sur* $]0, 1[$,
f) *est telle que*

$$\lim_{x \to (-1)^+} \operatorname{arcth}(x) = -\infty \quad \textit{et} \quad \lim_{x \to 1^-} \operatorname{arcth}(x) = +\infty.$$

Preuve. Tout est immédiat à partir du théorème de la fonction inverse et de la formule $\operatorname{ch}^2(x) = 1/(1 - \operatorname{th}^2(x))$.∎

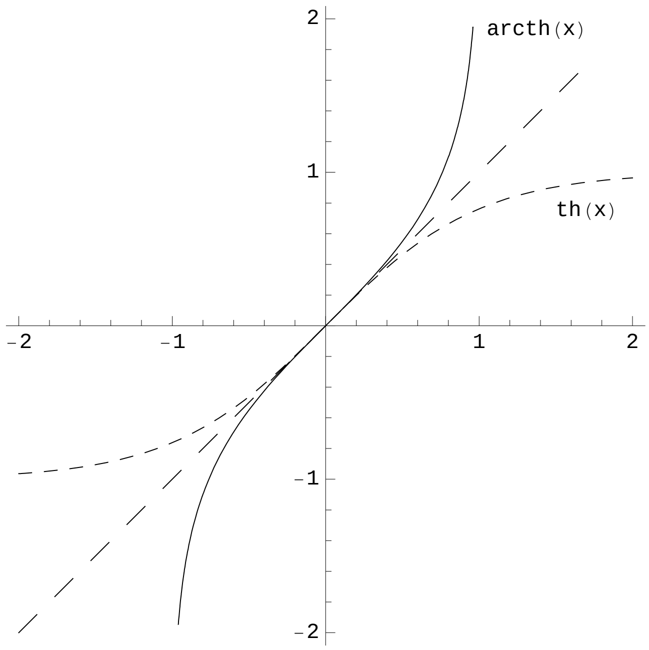
Remarque. En procédant comme pour la fonction arcsh, on obtient de suite

$$\operatorname{arcth}(x) = X \Leftrightarrow \mathrm{e}^{2X} = \frac{1 + x}{1 - x},$$

donc

$$\operatorname{arcth}(x) = \frac{1}{2} \ln(\frac{1 + x}{1 - x}), \quad \forall x \in \,]-1, 1[\,;$$

la fonction arcth n'est donc pas à proprement parler une nouvelle fonction élémentaire.



La fonction arccoth

En procédant de même à partir de la restriction de la fonction coth à l'intervalle $]0, +\infty[$, on introduit la fonction suivante.

Définition. La restriction de la fonction coth à $]0, +\infty[$ admet une fonction inverse, appelée *arc cotangente hyperbolique* et notée arccoth, qui jouit des propriétés suivantes.

Théorème 6.7.5.4 *La fonction* arccoth

a) *est définie, réelle et strictement décroissante sur* $]1, +\infty[$,

b) *vérifie les égalités*

$$\begin{aligned} \operatorname{arccoth}(\coth(x)) &= x, \quad \forall x \in \,]0, +\infty[\,, \\ \coth(\operatorname{arccoth}(x)) &= x, \quad \forall x \in \,]1, +\infty[\,, \end{aligned}$$

c) *appartient à* $\mathrm{C}_\infty(]1, +\infty[)$ *et on a* $\mathrm{D}\operatorname{arccoth}(x) = 1/(1 - x^2)$,

d) *est convexe,*

e) *est telle que*

$$\lim_{x \to 1^+} \operatorname{arccoth}(x) = +\infty \quad \textit{et} \quad \lim_{x \to +\infty} \operatorname{arccoth}(x) = 0. \blacksquare$$

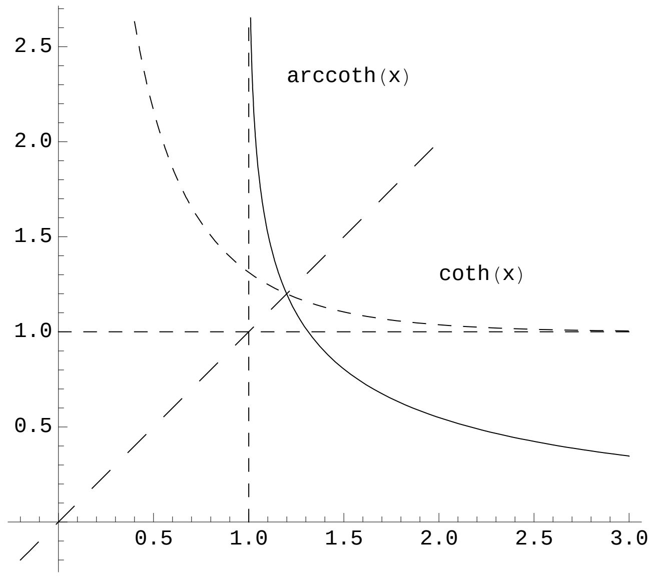
Remarque. En procédant comme pour la fonction arcsh, on obtient de suite

$$\operatorname{arccoth}(x) = X \Leftrightarrow \mathrm{e}^{2X} = \frac{x+1}{x-1},$$

donc

$$\operatorname{arccoth}(x) = \frac{1}{2}\ln(\frac{x+1}{x-1}), \quad \forall x \in \,]1, +\infty[\,;$$

la fonction arccoth n'est donc pas à proprement parler une nouvelle fonction élémentaire.



6.7.6 Fonctions trigonométriques inverses

La fonction arcsin

Définition. La restriction de la fonction sin à $[-\pi/2, \pi/2]$ admet une fonction inverse sur $[-1, 1]$, appelée *arc sinus* et notée arcsin, qui jouit des propriétés suivantes.

Théorème 6.7.6.1 *La fonction* arcsin

a) *est définie, réelle et strictement croissante sur* $[-1, 1]$,

b) *vérifie les égalités*

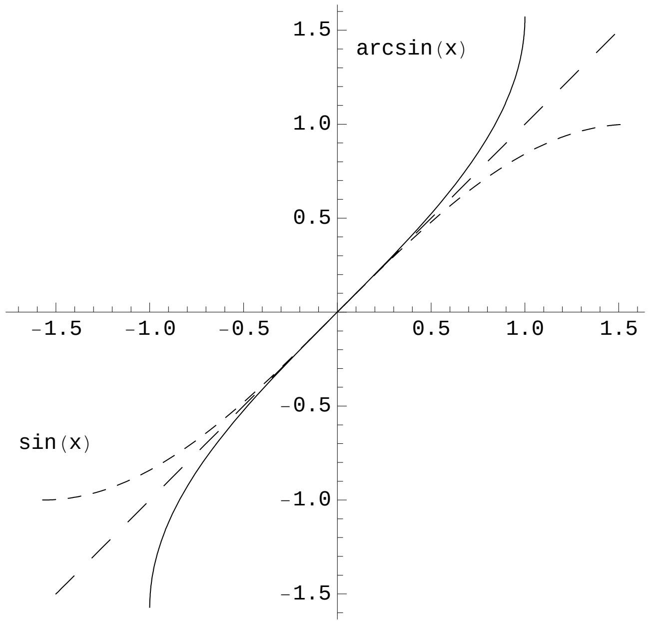
$$\begin{aligned} \arcsin(\sin(x)) &= x, \quad \forall x \in [-\pi/2, \pi/2]\,, \\ \sin(\arcsin(x)) &= x, \quad \forall x \in [-1, 1]\,. \end{aligned}$$

En particulier, on a $\arcsin(-1) = -\pi/2$ *et* $\arcsin(1) = \pi/2$.

c) *appartient à* $C_0([-1,1])$ *et à* $C_\infty(]-1,1[)$. *De plus,*

$$\mathrm{D}\arcsin(x) = \frac{1}{\sqrt{1-x^2}}.$$

d) *est concave sur* $[-1,0]$ *et convexe sur* $[0,1]$.∎



La fonction arccos

Vu la formule des compléments $\cos(x) = \sin(\pi/2 - x)$, on est amené à introduire la définition suivante.

Définition. La fonction arccos, appelée *arc cosinus*, est définie sur $[-1,1]$ par

$$\arccos(x) = \frac{\pi}{2} - \arcsin(x), \quad \forall x \in [-1,1].$$

Il ne s'agit donc pas d'une nouvelle fonction élémentaire. Signalons quand même ses propriétés.

Théorème 6.7.6.2 *La fonction* arccos

a) *est définie, réelle et strictement décroissante sur* $[-1,1]$,

b) *vérifie les égalités*

$$\begin{aligned} \arccos(\cos(x)) &= x, \quad \forall x \in [0,\pi], \\ \cos(\arccos(x)) &= x, \quad \forall x \in [-1,1]. \end{aligned}$$

En particulier, on a $\arccos(-1) = \pi$ *et* $\arccos(1) = 0$.

c) *appartient à* $C_0([-1,1])$ *et à* $C_\infty(]-1,1[)$ *et*

$$\mathrm{Darccos}(x) = -\frac{1}{\sqrt{1-x^2}}.$$

d) *est convexe sur* $[-1,0]$ *et concave sur* $[0,1]$.∎

La fonction arctg

Définition. La restriction de la fonction tg à $]-\pi/2, \pi/2[$ admet une fonction inverse sur $\mathbb{R}$, appelée *arc tangente* et notée arctg, qui jouit des propriétés suivantes.

Théorème 6.7.6.3 *La fonction* arctg

a) *est définie, réelle et strictement croissante sur* $\mathbb{R}$,

b) *vérifie les égalités*

$$\begin{array}{rcl} \mathrm{arctg}(\mathrm{tg}(x)) & = & x, \quad \forall x \in \,]-\pi/2, \pi/2[\,, \\ \mathrm{tg}(\mathrm{arctg}(x)) & = & x, \quad \forall x \in \,]-\infty, +\infty[\,. \end{array}$$

En particulier, on a $\mathrm{arctg}(0) = 0$.

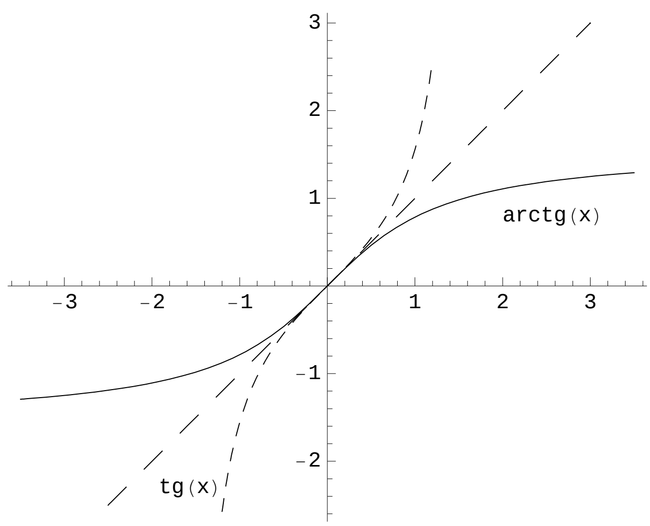
c) *appartient à* $C_\infty(\mathbb{R})$ *et*

$$\mathrm{Darctg}(x) = \frac{1}{1+x^2}.$$

d) *est convexe sur* $]-\infty, 0[$ *et concave sur* $]0, +\infty[$.

e) *donne lieu aux relations*

$$\lim_{x\to-\infty} \mathrm{arctg}(x) = (-\pi/2)^+ \quad \textit{et} \quad \lim_{x\to+\infty} \mathrm{arctg}(x) = (\pi/2)^-.$$∎



La fonction arccotg

Vu les formules des compléments, on a la formule $\mathrm{cotg}(x) = \mathrm{tg}(\pi/2 - x)$. Dès lors, on est amené à introduire la définition suivante.

Définition. La fonction arccotg, appelée *arc cotangente* est définie sur $\mathbb{R}$ par

$$\mathrm{arccotg}(x) = \frac{\pi}{2} - \mathrm{arctg}(x), \quad \forall x \in \mathbb{R}.$$

Théorème 6.7.6.4 *La fonction* arccotg

a) *est définie, réelle et strictement décroissante sur* $\mathbb{R}$,

b) *vérifie les égalités*

$$\begin{aligned} \mathrm{arccotg}(\mathrm{cotg}(x)) &= x, \quad \forall x \in \left]0, \pi\right[, \\ \mathrm{cotg}(\mathrm{arccotg}(x)) &= x, \quad \forall x \in \mathbb{R}. \end{aligned}$$

c) *appartient à* $\mathrm{C}_\infty(\mathbb{R})$ *et*

$$\mathrm{Darccotg}(x) = -\frac{1}{1+x^2}.$$

d) *est concave sur* $\left]-\infty, 0\right[$ *et convexe sur* $\left]0, +\infty\right[$.

e) *donne lieu aux relations*

$$\lim_{x\to-\infty} \mathrm{arccotg}(x) = \pi^- \quad et \quad \lim_{x\to+\infty} \mathrm{arccotg}(x) = 0^+. \blacksquare$$

6.7.7 Bases de la technique de la dérivation

La technique de la dérivation est basée sur les données suivantes.

a) Les renseignements repris au tableau de la page 217 et que nous venons de développer doivent être parfaitement connus.

b) Les théorèmes de dérivation des fonctions composées et des fonctions de fonction sont souvent utilisés sous la formulation simplifiée suivante: si J appartient à $\mathbb{N}_0$ ou est égal à ∞, si $f_1, \ldots, f_J$ appartiennent à $\mathrm{C}_p(\Omega)$ et sont réels et si f appartient à $\mathrm{C}_p(\omega)$ avec $\omega \supset \{ (f_1(x), \ldots, f_J(x)) : x \in \Omega\}$, alors $f(f_1, \ldots, f_J)$ appartient à $\mathrm{C}_p(\Omega)$ et

$$[\mathrm{D}_k f(f_1, \ldots, f_J)]_x = \sum_{j=1}^{J} [\mathrm{D}_j f]_{(f_1(x), \ldots, f_J(x))} \cdot [\mathrm{D}_k f_j]_x.$$

Remarquons d'ailleurs que ces théorèmes de dérivation assurent que la dérivée d'une fonction composée obtenue à partir de fonctions élémentaires uniquement est encore une fonction du même type.

Symbole	Dérivabilité	Dérivée
x^m	$\in \mathrm{C}_\infty(\mathbb{R})$	mx^{m-1}
x^{-m}	$\in \mathrm{C}_\infty(\mathbb{R} \setminus \{0\})$	$-mx^{-m-1}$
x^c	$\in \mathrm{C}_\infty(]0, +\infty[)$	cx^{c-1}
r^x	$\in \mathrm{C}_\infty(\mathbb{R})$	$r^x \ln(r)$
$\cos(x)$	$\in \mathrm{C}_\infty(\mathbb{R})$	$-\sin(x)$
$\sin(x)$	$\in \mathrm{C}_\infty(\mathbb{R})$	$\cos(x)$
$\mathrm{tg}(x)$	$\in \mathrm{C}_\infty(\mathbb{R} \setminus \{\,(2k+1)\pi/2 : k \in \mathbb{Z}\})$	$1/\cos^2(x)$
$\mathrm{cotg}(x)$	$\in \mathrm{C}_\infty(\mathbb{R} \setminus \{\,k\pi : k \in \mathbb{Z}\})$	$-1/\sin^2(x)$
$\mathrm{ch}(x)$	$\in \mathrm{C}_\infty(\mathbb{R})$	$\mathrm{sh}(x)$
$\mathrm{sh}(x)$	$\in \mathrm{C}_\infty(\mathbb{R})$	$\mathrm{ch}(x)$
$\mathrm{th}(x)$	$\in \mathrm{C}_\infty(\mathbb{R})$	$1/\mathrm{ch}^2(x)$
$\mathrm{coth}(x)$	$\in \mathrm{C}_\infty(\mathbb{R} \setminus \{0\})$	$-1/\mathrm{sh}^2(x)$
$\ln(x)$	$\in \mathrm{C}_\infty(]0, +\infty[)$	$1/x$
$\arcsin(x)$	$\in \mathrm{C}_\infty(]-1, 1[)$	$1/\sqrt{1-x^2}$
$\arccos(x)$	$\in \mathrm{C}_\infty(]-1, 1[)$	$-1/\sqrt{1-x^2}$
$\mathrm{arctg}(x)$	$\in \mathrm{C}_\infty(\mathbb{R})$	$1/(1+x^2)$
$\mathrm{arccotg}(x)$	$\in \mathrm{C}_\infty(\mathbb{R})$	$-1/(1+x^2)$
$\mathrm{arcch}(x)$	$\in \mathrm{C}_\infty(]1, +\infty[)$	$1/\sqrt{x^2-1}$
$\mathrm{arcsh}(x)$	$\in \mathrm{C}_\infty(\mathbb{R})$	$1/\sqrt{1+x^2}$
$\mathrm{arcth}(x)$	$\in \mathrm{C}_\infty(]-1, 1[)$	$1/(1-x^2)$
$\mathrm{arccoth}(x)$	$\in \mathrm{C}_\infty(]1, +\infty[)$	$1/(1-x^2)$

Remarques: $m \in \mathbb{N}_0,\ r > 0,\ c \in \mathbb{C}$,
si $r > 0,\ x^r \in \mathrm{C}_0([0, +\infty[)$ avec $0^r = 0$,
$\arcsin(x) \in \mathrm{C}_0([-1, 1])$ avec $\arcsin(\pm 1) = \pm\pi/2$,
$\arccos(x) = \pi/2 - \arcsin(x)$,
$\mathrm{arccotg}(x) = \pi/2 - \mathrm{arctg}(x)$,
$\mathrm{arcch}(x) = \ln(x + \sqrt{x^2-1}) \in \mathrm{C}_0([1, +\infty[)$ avec $\mathrm{arcch}(1) = 0$,
$\mathrm{arcsh}(x) = \ln(x + \sqrt{1+x^2})$,
$\mathrm{arcth}(x) = \frac{1}{2}\ln((1+x)/(1-x))$,
$\mathrm{arccoth}(x) = \frac{1}{2}\ln((x+1)/(x-1))$.

6.7.8 Egalité de fonctions

Nous avons déjà vu des exemples concrets du fait que certaines fonctions peuvent s'écrire sous deux formes distinctes, très différentes a priori. Il en est ainsi par exemple pour

$$\begin{aligned}
\operatorname{arcch}(x) &= \ln(x+\sqrt{x^2-1}), \quad \forall x \in [1,+\infty[\,, \\
\operatorname{arcsh}(x) &= \ln(x+\sqrt{1+x^2}), \quad \forall x \in \mathbb{R}, \\
\operatorname{arcth}(x) &= \frac{1}{2}\ln\frac{1+x}{1-x}, \quad \forall x \in \,]-1,1[\,.
\end{aligned}$$

Pour établir ces égalités, on aurait pu aussi recourir au théorème de l'ouvert connexe. Ainsi, par exemple, nous savons que arcsh appartient à $\mathrm{C}_\infty(\mathbb{R})$ et est tel que $\operatorname{arcsh}(0) = 0$ et $\mathrm{D}\operatorname{arcsh}(x) = (1+x^2)^{-1/2}$. De plus, on vérifie aisément que $\ln(x+(1+x^2)^{1/2})$ appartient à $\mathrm{C}_\infty(\mathbb{R})$ et est tel que $\ln(0+(1+0^2)^{1/2}) = 0$ et $\mathrm{D}\ln(x+(1+x^2)^{1/2}) = (1+x^2)^{-1/2}$. D'où l'égalité de ces deux fonctions.

Le résultat suivant, qui peut être considéré comme étant un exercice, a des conséquences importantes.

Exercice. *Etablir que*

$$\begin{aligned}
\operatorname{arctg}(x)+\operatorname{arctg}(1/x) &= \pi/2, \quad \forall x > 0, \\
\operatorname{arctg}(x)+\operatorname{arctg}(1/x) &= -\pi/2, \quad \forall x < 0.
\end{aligned}$$

Suggestion. Comme on a $\operatorname{arctg} \in \mathrm{C}_\infty(\mathbb{R})$ et $1/x \in \mathrm{C}_\infty(\mathbb{R}\setminus\{0\})$, le théorème de dérivation des fonctions de fonction signale aussitôt que $\operatorname{arctg}(1/x)$ appartient à $\mathrm{C}_\infty(\mathbb{R}\setminus\{0\})$. Dès lors, il vient

$$\operatorname{arctg}(x)+\operatorname{arctg}(1/x) \in \mathrm{C}_\infty(]-\infty,0[\,\cup\,]0,+\infty[).$$

Calculons à présent la dérivée de cette fonction. Dans les ouverts connexes $]-\infty,0[$ et $]0,+\infty[$, on a

$$\mathrm{D}\operatorname{arctg}(x)+\operatorname{arctg}(1/x) = \frac{1}{1+x^2}+\frac{1}{1+(1/x)^2}\cdot\frac{-1}{x^2} = 0.$$

Il s'ensuit que la fonction $\operatorname{arctg}(x)+\operatorname{arctg}(1/x)$ est constante sur $]-\infty,0[$ et sur $]0,+\infty[$ (mais il peut fort bien s'agir de constantes différentes C_1 et C_2).

Déterminons C_1. On peut, par exemple, calculer C_1 au moyen de la valeur de la fonction $\operatorname{arctg}(x)+\operatorname{arctg}(1/x)$ en un point particulier de $]-\infty,0[$. Ainsi, en $x=-1$, on obtient

$$C_1 = \operatorname{arctg}(-1)+\operatorname{arctg}(-1) = -\pi/2.$$

On peut aussi calculer C_1 par

$$C_1 = \lim_{x \to 0^-} (\operatorname{arctg}(x) + \operatorname{arctg}(1/x)) = 0 - \pi/2 = -\pi/2$$

ou par

$$C_1 = \lim_{x \to -\infty} (\operatorname{arctg}(x) + \operatorname{arctg}(1/x)) = -\pi/2 + 0 = -\pi/2.$$

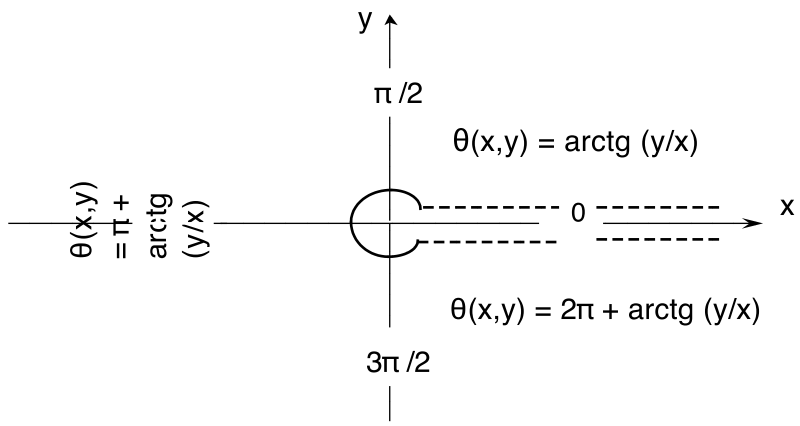
On obtient C_2 de la même manière. Il vient ainsi

$$C_2 = \operatorname{arctg}(1) + \operatorname{arctg}(1) = \pi/2$$

et nous voyons donc que C_1 diffère de C_2. En particulier, la fonction étudiée n'admet pas de prolongement continu sur $\mathbb{R}$.

6.7.9 Etude de la fonction $\arg(z)$ sur $\mathbb{C} \setminus \{0\}$

Théorème 6.7.9.1 *La solution unique* $\arg(z) = \theta(x, y) \in [0, 2\pi[$ *de l'équation* $e^{i\theta(x,y)} = z/|z|$ *avec* $z = x + iy \in \mathbb{C} \setminus \{0\}$ *est donnée par le graphique ci-dessous:*



De plus, on a

$$\theta(x, y) \in \mathrm{C}_\infty(\mathbb{R}^2 \setminus \{ (x, 0) : x \geq 0\}),$$

ainsi que

$$[\mathrm{D}_1\theta]_{(x,y)} = -\frac{y}{x^2 + y^2} \quad \textit{et} \quad [\mathrm{D}_2\theta]_{(x,y)} = \frac{x}{x^2 + y^2}.$$

Preuve. a) Vérifions tout d'abord que la solution proposée est correcte.

Il est clair qu'on doit avoir $\theta(0, y) = \pi/2$ pour tout $y > 0$ et $\theta(0, y) = 3\pi/2$ pour tout $y < 0$. De plus de

$$e^{i\theta(x,y)} = \frac{x + iy}{\sqrt{x^2 + y^2}} \Longleftrightarrow \left\{ \begin{array}{l} \cos(\theta(x, y)) = \frac{x}{\sqrt{x^2+y^2}} \\ \sin(\theta(x, y)) = \frac{y}{\sqrt{x^2+y^2}} \end{array} \right\},$$

on tire aussitôt l'égalité $\mathrm{tg}(\theta(x,y)) = y/x$ pour tout $x \neq 0$. Dans ces conditions, pour $x \neq 0$, il vient

$$\theta(x,y) = \mathrm{arctg}(y/x) + k\pi \quad \text{avec} \quad k \in \{0,1,2\}.$$

Les formules $\cos(\theta + k\pi) = (-1)^k \cos(\theta)$ et $\sin(\theta + k\pi) = (-1)^k \sin(\theta)$ valables pour tout $k \in \mathbb{Z}$ et tout $\theta \in \mathbb{R}$ permettent alors de conclure aussitôt.

b) Etudions la dérivabilité de cette fonction.

Comme on a

$$\mathrm{arctg}(y/x) \in \mathrm{C}_\infty(\{\,(x,y) : x \neq 0, y \in \mathbb{R}\}),$$

ainsi que

$$\begin{aligned} \mathrm{D}_x \mathrm{arctg}(y/x) &= -\frac{y}{x^2+y^2} \quad \text{sur} \quad \{\,(x,y) : x \neq 0, y \in \mathbb{R}\}\,, \\ \mathrm{D}_y \mathrm{arctg}(y/x) &= \frac{x}{x^2+y^2} \quad \text{sur} \quad \{\,(x,y) : x \neq 0, y \in \mathbb{R}\}\,, \end{aligned}$$

il suffit de noter les deux points suivants.

Considérons d'abord un point $(0,y)$ avec $y > 0$. D'une part, on a

$$\lim_{h \to 0^+,\, h \in \mathbb{R}_0} \frac{\mathrm{arctg}(y/h) - \pi/2}{h} = \lim_{h \to 0^+,\, h \in \mathbb{R}_0} \frac{-\mathrm{arctg}(h/y)}{h} = -\frac{1}{y}$$

et

$$\lim_{h \to 0^-,\, h \in \mathbb{R}_0} \frac{\pi + \mathrm{arctg}(y/h) - \pi/2}{h} = \lim_{h \to 0^-,\, h \in \mathbb{R}_0} \frac{-\mathrm{arctg}(h/y)}{h} = -\frac{1}{y}$$

donc la fonction $\theta(x,y)$ est dérivable par rapport à x en $(0,y)$ et

$$[\mathrm{D}_x \theta]_{(0,y)} = -\frac{1}{y} = \left[-\frac{y}{x^2+y^2}\right]_{(0,y)}, \quad \forall y > 0.$$

D'autre part, on a bien sûr

$$[\mathrm{D}_y \theta]_{(0,y)} = 0 = \left[\frac{x}{x^2+y^2}\right]_{(0,y)}, \quad \forall y > 0.$$

Le cas d'un point $(0,y)$ avec $y < 0$ se traîte de même.

D'où la conclusion.∎

$* \to$ On en déduit de suite le résultat suivant.

Théorème 6.7.9.2 *Pour tout nombre complexe non nul $z = x + iy$, la solution unique $\theta'(x, y) \in [a, a + 2\pi[$ de l'équation $\mathrm{e}^{i\theta(x,y)} = z/|z|$ est donnée au moyen de la fonction θ de l'énoncé précédent par*

$$\theta'(x, y) = \theta(x\cos(a) + y\sin(a), -x\sin(a) + y\cos(a)) + a.$$

Elle appartient à

$$\mathrm{C}_\infty(\mathbb{R}^2 \setminus \{\, (r\cos(a), r\sin(a) : r \geq 0\})$$

et ses dérivées sont données par

$$\mathrm{D}_x\theta(x, y) = -\frac{y}{x^2 + y^2} \quad \textit{et} \quad \mathrm{D}_y\theta(x, y) = \frac{x}{x^2 + y^2}. \blacksquare \quad \leftarrow *$$

6.7.10 Calcul des dérivées multiples

Il n'y a que quelques cas particuliers pour lesquels on peut donner des formules ou méthodes qui permettent de calculer aisément une dérivée d'ordre quelconque d'une fonction composée obtenue à partir de fonctions élémentaires.

A. Voici tout d'abord quelques formules usuelles.

a) Pour tout $c \in \mathbb{C}$ et tous entiers $m \in \mathbb{N}_0$ et $p \in \mathbb{N}$, on a

$$\mathrm{D}^p(x - c)^m = \begin{cases} \dfrac{m!}{(m-p)!}(x - c)^{m-p} & \text{si} \quad p \leq m, \\ 0 & \text{si} \quad p > m. \end{cases}$$

b) Pour tout $c \in \mathbb{C}$ et tous m, $p \in \mathbb{N}_0$, on a

$$\mathrm{D}^p\frac{1}{(x - c)^m} = (-1)^p m(m + 1) \ldots (m + p - 1)\frac{1}{(x - c)^{m+p}}.$$

c) Pour tout $c \in \mathbb{C}$ et tout $p \in \mathbb{N}$, on a

$$\mathrm{D}^p\mathrm{e}^{cx} = c^p\mathrm{e}^{cx}.$$

d) Pour tout $r > 0$ et tout $p \in \mathbb{N}$, on a

$$\mathrm{D}^p r^x = (\ln(r))^p r^x.$$

B. On en déduit aisément les dérivées des combinaisons linéaires de telles fonctions.

Cette circonstance a notamment lieu pour

a) les polynômes.

b) les fractions rationnelles. Ainsi par exemple, pour tous $a, b \in \mathbb{C}$ tels que $a \neq b$ et tout $p \in \mathbb{N}_0$, il vient

$$\begin{aligned} \mathrm{D}^p \frac{1}{(x-a)(x-b)} &= \mathrm{D}^p \left(\frac{1}{a-b} \left(\frac{1}{x-a} - \frac{1}{x-b} \right) \right) \\ &= \frac{(-1)^p p!}{a-b} \left(\frac{1}{(x-a)^{p+1}} - \frac{1}{(x-b)^{p+1}} \right). \end{aligned}$$

c) $\mathrm{D}^p\cos(ax)$ et $\mathrm{D}^p\sin(ax)$ si $a \in \mathbb{R}$ et $p \in \mathbb{N}_0$. De fait, il vient

$$\mathrm{D}^p \left\{ \begin{array}{c} \cos(ax) \\ \sin(ax) \end{array} \right\} = \left\{ \begin{array}{c} \Re \\ \Im \end{array} \right\} \mathrm{D}^p \mathrm{e}^{iax} = \left\{ \begin{array}{c} \Re \\ \Im \end{array} \right\} \left(i^p a^p (\cos(ax) + i \sin(ax)) \right).$$

d) $\mathrm{D}^p(\mathrm{e}^{ax}\cos(bx))$ et $\mathrm{D}^p(\mathrm{e}^{ax}\sin(bx))$ si $a, b \in \mathbb{R}$. En effet, on a

$$\mathrm{D}^p \left\{ \begin{array}{c} \mathrm{e}^{ax}\cos(bx) \\ \mathrm{e}^{ax}\sin(bx) \end{array} \right\} = \left\{ \begin{array}{c} \Re \\ \Im \end{array} \right\} \mathrm{D}^p \mathrm{e}^{(a+ib)x} = \left\{ \begin{array}{c} \Re \\ \Im \end{array} \right\} \left((a+ib)^p \mathrm{e}^{(a+ib)x} \right).$$

e) les polynômes en e^{ax}, $\mathrm{ch}(bx+c)$, $\mathrm{sh}(bx+c)$, $\cos(bx+c)$ ou $\sin(bx+c)$ car un tel polynôme peut s'exprimer sous la forme d'une combinaison linéaire d'expressions de la forme $\mathrm{e}^{\alpha x}\cos(\beta x)$ ou $\mathrm{e}^{\alpha x}\sin(\beta x)$. Ainsi, par exemple, on a

$$\mathrm{D}^p\cos^2(x) = \mathrm{D}^p \frac{1+\cos(2x)}{2} = \frac{1}{2}\mathrm{D}^p\cos(2x), \quad \forall p \in \mathbb{N}_0.$$

C. Il convient aussi d'essayer de recourir à la formule de Leibniz; on obtient parfois une formule de récurrence entre les dérivées successives de la fonction.

En voici quelques exemples théoriques.

a) Si R et S sont des polynômes de degrés respectifs r et s tels que $r < s$, on obtient une relation de récurrence entre les dérivées successives de $f = R/S$ en calculant $\mathrm{D}^p R = \mathrm{D}^p(fS)$ avec $p \geq s$.

b) Si R et S sont des polynômes de degrés respectifs r et s tels que $r < s$, on obtient une relation de récurrence entre les dérivées successives de $f = \mathrm{e}^{R/S}$ en notant que la *"dérivée logarithmique"* de f, c'est-à-dire l'expression $(\mathrm{D}f)/f$, est donnée par

$$\frac{\mathrm{D}f}{f} = \frac{S \cdot \mathrm{D}R - R \cdot \mathrm{D}S}{S^2} = \frac{P}{Q}$$

où P et Q sont des polynômes. On calcule ensuite $\mathrm{D}^p(Q\mathrm{D}f) = \mathrm{D}^p(fP)$ avec $p \geq 2s$.

c) Si f est une fonction inverse, on remarquera que sa dérivée première ou sa dérivée seconde donnent lieu à une relation du type $R \cdot \mathrm{D}f = S$ ou $R \cdot \mathrm{D}^2 f = S \cdot \mathrm{D}f$ où R et S sont des polynômes. On calcule ensuite $\mathrm{D}^p(R \cdot f) = \mathrm{D}^p S$ ou $\mathrm{D}^p(R \cdot \mathrm{D}^2 f) = \mathrm{D}^p(S \cdot \mathrm{D}f)$ où p est supérieur ou égal aux degrés de R et de S.

Exemple. *Obtenir une relation de récurrence entre les dérivées successives de la fonction*

$$f(x) = \frac{ax+b}{cx^2+dx+e}, \quad (a,b,c,d,e \in \mathbb{C}).$$

Dans le plus grand ouvert Ω de $\mathbb{R}$ tel que $f \in \mathrm{C}_\infty(\Omega)$, on a

$$(cx^2+dx+e)f(x) - (ax+b) = 0.$$

Dès lors, pour tout entier $p \geq 2$, la formule de Leibniz donne

$$(cx^2+dx+e)\mathrm{D}^p f(x) + p(2cx+d)\mathrm{D}^{p-1}f(x) + p(p-1)c\mathrm{D}^{p-2}f(x) = 0.$$

Pour être complet, il faut encore noter qu'on a f et que $\mathrm{D}f$ s'écrit

$$\mathrm{D}f = \frac{a(cx^2+dx+e) - (ax+b)(2cx+d)}{(cx^2+dx+e)^2}.$$

Exemple. *Obtenir une relation de récurrence entre les dérivées successives de la fonction* arcsin. *En déduire la valeur de*

$$[\mathrm{D}^p \mathrm{arcsin}]_0, \quad \forall p \in \mathbb{N}.$$

On sait que la fonction $f = \mathrm{arcsin}$ appartient à $\mathrm{C}_\infty(]-1,1[)$ et que

$$\mathrm{Darcsin}(x) = \frac{1}{\sqrt{1-x^2}}, \quad \forall x \in \left]-1,1\right[.$$

En dérivant cette identité, il vient

$$\mathrm{D}^2 f = \frac{x}{1-x^2}\mathrm{D}f, \quad \forall x \in \left]-1,1\right[,$$

donc

$$(1-x^2)\cdot \mathrm{D}^2 f - x \cdot \mathrm{D}f = 0.$$

Il s'ensuit par la formule de Leibniz que, pour tout entier $p \geq 2$,

$$(1-x^2)\cdot \mathrm{D}^{p+2}f - 2px \cdot \mathrm{D}^{p+1}f - p(p-1)\mathrm{D}^p f - x\mathrm{D}^{p+1}f - p\mathrm{D}^p f = 0,$$

c'est-à- dire

$$(1-x^2)\cdot \mathrm{D}^{p+2}f - (2p+1)x \cdot \mathrm{D}^{p+1}f - p^2\mathrm{D}^p f = 0$$

sur $]-1,1[$ pour tout entier $p \geq 2$. (* On peut simplifier ce qui suit si on remarque que cette dernière égalité est en fait valable sur $]-1,1[$ pour tout entier $p \geq 0$.*) Pour que la relation de récurrence soit complète, on doit encore noter que

$$\begin{array}{rcl} f(x) & = & \arcsin(x), \\ [\mathrm{D}f]_x & = & (1-x^2)^{-1/2}, \\ [\mathrm{D}^2 f]_x & = & x(1-x^2)^{-3/2}, \\ [\mathrm{D}^3 f]_x & = & (1-x^2)^{-3/2} + 3x^2(1-x^2)^{-5/2}, \end{array}$$

sur $]-1,1[$. En particulier, en $0 \in\,]-1,1[$, il vient

$$\left\{ \begin{array}{lcl} f(0) & = & 0, \\ [\mathrm{D}f]_0 & = & 1, \\ [\mathrm{D}^2 f]_0 & = & 0, \\ [\mathrm{D}^3 f]_0 & = & 1, \\ [\mathrm{D}^{p+2} f]_0 & = & p^2 [\mathrm{D}^p f]_0, \quad \forall p \geq 2, \end{array} \right.$$

c'est-à-dire

$$[\mathrm{D}^{2p} f]_0 = 0, \quad \forall p \geq 2,$$

et

$$[\mathrm{D}^{2p+1} f]_0 = (2p-1)^2 (2p-3)^2 \dots 3^2 = \frac{((2p)!)^2}{2^{2p}(p!)^2}, \quad \forall p \geq 2.$$

Chapitre 7

Applications de la dérivation

7.1 Calcul de limites

7.1.1 Position du problème

Nous avons déjà répondu partiellement à la question suivante.

Problème. Soient $f_1, \ldots, f_J$ des fonctions réelles, en nombre fini et définies sur une même partie de $\mathbb{R}^n$, et soit f une fonction définie sur une partie A' de $\mathbb{R}^J$ telle que

$$A' \supset \{ (f_1(x), \ldots, f_J(x)) : x \in A \} .$$

Soit en outre ξ un point $x_0 \in A^-$ ou ∞ (cette dernière éventualité n'étant envisagée que si A n'est pas borné); il existe donc au moins une suite x_m de points de A qui converge vers ξ. Supposons que chacune des fonctions $f_1, \ldots, f_J$ admette une limite (finie ou non) en ξ à savoir

$$\lim_{x \to \xi} f_j(x) = \alpha_j,$$

pour tout $j \in \{1, \ldots, J\}$. La fonction $f(f_1, \ldots, f_J)$ est alors définie sur A et on demande si la limite

$$\lim_{x \to \xi} f(f_1(x), \ldots, f_J(x))$$

existe et, si la réponse est affirmative, d'en donner la valeur.

La réponse obtenue plus tôt s'énonce comme suit: *si les limites $\alpha_1, \ldots, \alpha_J$ sont finies, si $\alpha = (\alpha_1, \ldots, \alpha_J)$ appartient à A' et si f est continu en $(\alpha_1, \ldots, \alpha_J)$, alors*

$$\lim_{x \to \xi} f(f_1(x), \ldots, f_J(x)) = f(\alpha_1, \ldots, \alpha_J). \quad (*)$$

En fait, la démonstration en était aisée et prouve que, *si $\lim_{X \to \alpha} f(X)$ a un sens pour $\alpha = (\alpha_1, \ldots, \alpha_J)$, alors*

$$\lim_{x \to \xi} f(f_1(x), \ldots, f_J(x)) = \lim_{X \to \alpha} f(X).$$

Dans le cas (*), on voit que la substitution pure et simple de α dans f procure le résultat. Sans l'hypothèse de continuité de f en α, cette substitution pure et simple est dénuée de sens.

Cependant, il existe des cas où cette substitution admet une interprétation directe. En voici quelques exemples concrets:

a) *avec* ∞*:*

$$\begin{array}{lll} c+\infty=\infty, & \forall c\in\mathbb{C}, & \text{cf.:}\quad \lim_{x\to+\infty}(\mathrm{th}(x)+x),\\ c\cdot\infty=\infty, & \forall c\in\mathbb{C}\setminus\{0\}, & \qquad\;\; \lim_{x\to+\infty}(x\cdot\mathrm{th}(x)),\\ c/\infty=0, & \forall c\in\mathbb{C}. & \qquad\;\; \lim_{x\to+\infty}(\mathrm{th}(x))/x. \end{array}$$

Une généralisation immédiate procure aussitôt les résultats suivants:

$$\lim_{x\to+\infty}(\sin(x)+x)=\infty,\quad \lim_{x\to+\infty}(2+\sin(x))\cdot x=\infty \quad\text{et}\quad \lim_{x\to+\infty}\frac{\sin(x)}{x}=0.$$

b) *avec* $\pm\infty$*:*

$$\begin{array}{l} a\pm\infty=\pm\infty,\quad \forall a\in\mathbb{R},\\ \pm\infty\pm\infty=\pm\infty,\\ a\cdot(\pm\infty)=\begin{cases}\pm\infty & \text{si}\quad a>0,\\ \mp\infty & \text{si}\quad a<0,\end{cases}\\ \dfrac{a}{+\infty}=\begin{cases}0^+ & \text{si}\quad a>0,\\ 0^- & \text{si}\quad a<0\end{cases}\quad\text{et}\quad \dfrac{a}{-\infty}=\begin{cases}0^- & \text{si}\quad a>0,\\ 0^+ & \text{si}\quad a<0.\end{cases} \end{array}$$

Cependant il existe des cas où la limite particulière cherchée existe même si $\lim_{x\to\alpha} f(x)$ n'existe pas. Ainsi $f_1(x)=x+\sin(x)$ et $f_2(x)=\sin(x)$ sont des fonctions réelles appartenant à $C_\infty(\mathbb{R})$ et convergeant vers 0 si $x\to 0$. La fonction $f(x,y)=y/x$ appartient à $C_\infty(\{(x,y): x\neq 0\})$ et n'admet pas de limite en $(0,0)$. Cependant nous allons voir que la limite $\lim_{x\to0}(x+\sin(x))/\sin(x)$ a un sens et vaut même 2.

Il s'agit de cas où la substitution de α dans $f(x)$ conduit à un symbole dénué d'interprétation; on dit qu'il s'agit d'un symbole *illusoire* ou *indéterminé.*

Ces cas sont nombreux; en voici les principaux.

Symboles illusoires de base: $0/0,\ 0\cdot\infty,\ \infty/\infty$

Ces trois cas sont équivalents car on a

$$\frac{f}{g}=f\cdot\frac{1}{g}=\frac{1/g}{1/f}.$$

On les trouve par exemple dans les limites suivantes

$$\lim_{x\to0}\frac{\sin(x)}{x},\quad \lim_{x\to+\infty}\frac{1}{x}\mathrm{e}^x,\quad\text{et}\quad \lim_{x\to+\infty}\frac{\ln(x)}{x}.$$

Symboles illusoires principaux

a) $\infty - \infty$: il se ramène aux précédents par l'opération

$$f - g = f \cdot (1 - \frac{g}{f})$$

et l'indétermination ne subsiste donc que si on a $g/f \to 1$.

b) Cas exponentiels: $0^0, 1^\infty, \infty^0$. Ils se ramènent aux symboles illusoires de base en utilisant la formule $X^x = \mathrm{e}^{x \ln(X)}$. On les trouve par exemple dans les limites suivantes

$$\lim_{x \to 0^+} (1/x - \operatorname{cotg}(x)), \quad \lim_{x \to 0^+} x^x, \quad \lim_{x \to +\infty} (1 + 1/x)^x, \quad \lim_{x \to +\infty} x^{1/x}.$$

Symboles illusoires qui se ramènent aux précédents

Ce sont des symboles tels que

$$\frac{0}{0} \cdot \frac{\infty - \infty}{0 \cdot \infty} \cdot 1^\infty.$$

L'apparition d'un symbole indéterminé indique tout simplement que la détermination de $\lim_{x \to \xi} f(f_1(x), \ldots, f_J(x))$ est délicate et exige une étude supplémentaire.

L'étude de ces limites repose sur la connaissance des propriétés des limites et sur le théorème de l'Hospital.

7.1.2 Théorème de l'Hospital

Introduisons tout d'abord quelques notations valables dans tout ce paragraphe.

Notations. Nous désignons par
r un point de $\mathbb{R}$,
ξ soit r^+, soit r^-, soit $+\infty$, soit $-\infty$,
V un ouvert de $\mathbb{R}$ pour lequel il existe

$$\begin{array}{llll}
\varepsilon' > 0 & \text{tel que} & V \supset]r, r + \varepsilon'[& \text{si on calcule} & \lim_{x \to r^+}, \\
\varepsilon' > 0 & \text{tel que} & V \supset]r - \varepsilon', r[& \text{si on calcule} & \lim_{x \to r^-}, \\
N' > 0 & \text{tel que} & V \supset]N', +\infty[& \text{si on calcule} & \lim_{x \to +\infty}, \\
N' > 0 & \text{tel que} & V \supset]-\infty, -N'[& \text{si on calcule} & \lim_{x \to -\infty}.
\end{array}$$

Théorème 7.1.2.1 *Si*

a) *f et g sont des fonctions réelles et dérivables sur V,*

b) *g et $\mathrm{D}g$ diffèrent de 0 en tout point de V,*

c) *on a le cas* $0/0$ (c'est-à-dire $\lim_{x\to\xi} f(x) = 0$ et $\lim_{x\to\xi} g(x) = 0$) *ou le cas* ∞/∞, (c'est-à-dire $\lim_{x\to\xi} g(x) = \infty$),

d) *on a $\lim_{x\to\xi} \mathrm{D}f(x)/\mathrm{D}g(x) = l$ (l pouvant être un nombre réel, $+\infty$ ou $-\infty$), alors il vient*

$$\lim_{x\to\xi} \frac{f(x)}{g(x)} = l.$$

Preuve. Nous allons établir le résultat simultanément pour les deux cas $\xi = r^-$ et $\xi = +\infty$; pour les cas $\xi = r^+$ et $\xi = -\infty$, il suffit de procéder de manière analogue.

Etablissons tout d'abord une formule auxiliaire.

Pour tous x, $y \in V$ tels que $x < y$ et $[x, y] \subset V$, la fonction de Z suivante

$$\mathrm{dtm}\begin{pmatrix} f(Z) & g(Z) & 1 \\ f(x) & g(x) & 1 \\ f(y) & g(y) & 1 \end{pmatrix} = \left\{ \begin{array}{l} f(Z)(g(x) - g(y)) - g(Z)(f(x) - f(y)) \\ \;+ (f(x)g(y) - f(y)g(x)) \end{array} \right.$$

est définie, réelle et dérivable sur V, et s'annule en x et en y. Vu le théorème de Rolle, il existe $z \in]x, y[$ où sa dérivée s'annule: on a alors

$$(g(x) - g(y))\mathrm{D}f(z) - (f(x) - f(y))\mathrm{D}g(z) = 0.$$

Pour tous x, $y \in V$ tels que $]x, y[\subset V$, il existe donc $z \in]x, y[$ tel que

$$\frac{f(x)}{g(x)} = \left(1 - \frac{g(y)}{g(x)}\right) \frac{\mathrm{D}f(z)}{\mathrm{D}g(z)} + \frac{f(y)}{g(x)}.$$

A. Cela étant, établissons le résultat dans le cas $l \in \mathbb{R}$.

La formule auxiliaire donne lieu à la formulation suivante: pour tous $x, y \in V$ tels que $]x, y[\subset V$, il existe $z \in]x, y[$ tel que

$$\left|\frac{f(x)}{g(x)} - l\right| \le \left|\frac{\mathrm{D}f(z)}{\mathrm{D}g(z)} - l\right| + \left|\frac{g(y)}{g(x)}\right| \left|\frac{\mathrm{D}f(z)}{\mathrm{D}g(z)}\right| + \left|\frac{f(y)}{g(x)}\right|.$$

a) Envisageons le cas $0/0$.

Fixons $\varepsilon > 0$. Il existe alors $x_0 \in V$ suffisamment voisin de ξ pour que

$$\left|\frac{\mathrm{D}f(z)}{\mathrm{D}g(z)} - l\right| \le \frac{\varepsilon}{2}, \quad \forall z \in]x_0, \xi[\,.$$

Ceci implique notamment la relation

$$\left|\frac{\mathrm{D}f(z)}{\mathrm{D}g(z)}\right| \le |l| + \frac{\varepsilon}{2}, \quad \forall z \in]x_0, \xi[\,.$$

Cela étant, vu l'hypothèse 0/0, pour tout $x \in \,]x_0, \xi[$, il existe $y \in \,]x, \xi[$ tel que

$$(|l| + \frac{\varepsilon}{2}) \left| \frac{g(y)}{g(x)} \right| + \left| \frac{f(y)}{g(x)} \right| \leq \frac{\varepsilon}{2}.$$

Au total, vu la formule auxiliaire, nous obtenons

$$\left| \frac{f(x)}{g(x)} - l \right| \leq \varepsilon, \quad \forall x \in \,]x_0, \xi[\,,$$

ce qui suffit.

b) Envisageons le cas ∞/∞.

Fixons $\varepsilon > 0$. Il existe alors $y_0 \in V$ suffisamment voisin de ξ tel que

$$\left| \frac{\mathrm{D}f(z)}{\mathrm{D}g(z)} - l \right| \leq \frac{\varepsilon}{2}, \quad \forall z \in \,]y_0, \xi[\,.$$

Ceci implique notamment la relation

$$\left| \frac{\mathrm{D}f(z)}{\mathrm{D}g(z)} \right| \leq |l| + \frac{\varepsilon}{2}, \quad \forall z \in \,]y_0, \xi[\,.$$

Cela étant, vu l'hypothèse ∞/∞, il existe $x_0 \in \,]y_0, \xi[$ tel que

$$(|l| + \frac{\varepsilon}{2}) \left| \frac{g(y_0)}{g(x)} \right| + \left| \frac{f(y_0)}{g(x)} \right| \leq \frac{\varepsilon}{2}, \quad \forall x \in \,]x_0, \xi[\,.$$

Au total, vu la formule auxiliaire, nous obtenons

$$\left| \frac{f(x)}{g(x)} - l \right| \leq \varepsilon, \quad \forall x \in \,]x_0, \xi[\,,$$

ce qui suffit.

B. Etablissons à présent le résultat dans le cas $l = +\infty$ ou $-\infty$.

La formule auxiliaire s'adapte à ces différentes possibilités de la manière suivante: pour tous x, $y \in V$ tels que $]x, y[\,\subset V$, il existe $z \in \,]x, y[$ tel que

i) si $l = +\infty$: $\quad \dfrac{f(x)}{g(x)} \geq \left(1 - \dfrac{g(y)}{g(x)}\right) \dfrac{\mathrm{D}f(z)}{\mathrm{D}g(z)} - \left| \dfrac{f(y)}{g(x)} \right|,$

ii) si $l = -\infty$: $\quad \dfrac{f(x)}{g(x)} \leq \left(1 - \dfrac{g(y)}{g(x)}\right) \dfrac{\mathrm{D}f(z)}{\mathrm{D}g(z)} + \left| \dfrac{f(y)}{g(x)} \right|.$

Cela étant, les cas $l = +\infty$ et $l = -\infty$ se traitent de manière analogue; aussi nous n'allons établir la propriété que dans le cas $l = +\infty$.

a) Envisageons le cas 0/0.

Fixons $N > 0$. Il existe alors $x_0 \in V$ suffisamment voisin de ξ tel que

$$\left|\frac{\mathrm{D}f(z)}{\mathrm{D}g(z)}\right| \geq 2(N+1), \quad \forall z \in \,]x_0, \xi[\,.$$

Cela étant, vu l'hypothèse 0/0, pour tout $x \in \,]x_0, \xi[$, il existe $y \in \,]x, \xi[$ tel que

$$\left|\frac{g(y)}{g(x)}\right| \leq \frac{1}{2} \text{ et } \left|\frac{f(y)}{g(x)}\right| \leq 1.$$

Au total, vu la formule auxiliaire, nous obtenons

$$\left|\frac{f(x)}{g(x)}\right| \geq \left|1 - \frac{g(y)}{g(x)}\right| \left|\frac{\mathrm{D}f(z)}{\mathrm{D}g(z)}\right| - \left|\frac{f(y)}{g(x)}\right| \geq N, \quad \forall x \in \,]x_0, \xi[\,,$$

ce qui suffit.

b) Envisageons le cas ∞/∞.

Fixons $N > 0$. Il existe alors $y_0 \in V$ suffisamment voisin de ξ tel que

$$\left|\frac{\mathrm{D}f(z)}{\mathrm{D}g(z)}\right| \geq 2(N+1), \quad \forall z \in \,]y_0, \xi[\,.$$

Cela étant, vu l'hypothèse ∞/∞, il existe $x_0 \in \,]y_0, \xi[$ tel que

$$\left|\frac{g(y_0)}{g(x)}\right| \leq \frac{1}{2} \text{ et } \left|\frac{f(y_0)}{g(x)}\right| \leq 1, \quad \forall x \in \,]x_0, \xi[\,.$$

Au total, vu la formule auxiliaire, nous obtenons

$$\left|\frac{f(x)}{g(x)}\right| \geq \left|1 - \frac{g(y_0)}{g(x)}\right| \left|\frac{\mathrm{D}f(z)}{\mathrm{D}g(z)}\right| - \left|\frac{f(y_0)}{g(x)}\right| \geq N, \quad \forall x \in \,]x_0, \xi[\,,$$

ce qui suffit. ∎

Remarques. a) La fonction g, dérivable sur un ouvert V de $\mathbb{R}$, est continue sur V. Comme g ne s'annule en aucun point de V, g garde un signe constant sur tout intervalle inclus dans V, vu le théorème des valeurs intermédiaires.

b) Si g est une fonction réelle et dérivable sur l'ouvert V de $\mathbb{R}$ et si $\mathrm{D}g$ ne s'annule en aucun point de V, le théorème de Rolle affirme que g ne peut s'annuler qu'une fois au plus sur tout intervalle ouvert inclus dans V. On peut donc ne pas vérifier l'hypothèse que g ne s'annule en aucun point de V car cette condition sera automatiquement vérifiée sur un ouvert V' inclus dans V et vérifiant l'hypothèse formulée sur V.

c) Si on a les hypothèses a), b) et c) du théorème de l'Hospital ainsi que la propriété $\lim_{x\to\xi} f(x)/g(x) = l$, l pouvant être un nombre réel, $+\infty$ ou $-\infty$, on ne peut pas en déduire l'égalité $\lim_{x\to\xi} \mathrm{D}f(x)/\mathrm{D}g(x) = l$.

Ainsi les fonctions $f(x) = x + \sin(x)$ et $g(x) = x$ vérifient les conditions a), b) et c) du théorème de l'Hospital avec $\xi = +\infty$ et $V =]0, +\infty[$ et sont telles que

$$\lim_{x\to+\infty} \frac{x+\sin(x)}{x} = 1 + \lim_{x\to+\infty} \frac{\sin(x)}{x} = 1,$$

alors que

$$\lim_{x\to+\infty} \frac{\mathrm{D}f(x)}{\mathrm{D}g(x)} = \lim_{x\to+\infty} (1+\cos(x))$$

n'existe pas.

d) Parfois le recours au théorème de l'Hospital ne permet pas de conclure.

Ainsi les fonctions $f(x) = x$ et $g(x) = (1+x^2)^{1/2}$ vérifient les conditions a), b) et c) du théorème de l'Hospital avec $\xi = +\infty$ et $V =]0, +\infty[$, et on a

$$\lim_{x\to+\infty} \frac{\mathrm{D}f(x)}{\mathrm{D}g(x)} = \lim_{x\to+\infty} \frac{\sqrt{1+x^2}}{x},$$

puis

$$\lim_{x\to+\infty} \frac{\mathrm{D}\sqrt{1+x^2}}{\mathrm{D}x} = \lim_{x\to+\infty} \frac{x}{\sqrt{1+x^2}}.$$

Cependant un recours aisé aux propriétés des limites donne

$$\lim_{x\to+\infty} \frac{x}{\sqrt{1+x^2}} = \lim_{x\to+\infty} \frac{1}{\sqrt{x^{-2}+1}} = 1^-.$$

Le même phénomène se présente dans le calcul de

$$\lim_{x\to+\infty} ((ax+b)/(cx+d))^{1/2}$$

avec a, b, c, $d > 0$.

e) Si l appartient à $\mathbb{R}$, il convient de remarquer également que les hypothèses a), b) et c) du théorème de l'Hospital et

$$\lim_{x\to\xi} \frac{\mathrm{D}f(x)}{\mathrm{D}g(x)} = l^+ \quad (\text{resp. } l^-)$$

ne permettent pas de conclure que

$$\lim_{x\to\xi} \frac{f(x)}{g(x)} = l^+ \quad (\text{resp. } l^-)\ .$$

Ainsi on a

$$\lim_{x\to+\infty} \frac{x}{\sqrt{1+x^2}} = 1^- \quad \text{et} \quad \lim_{x\to+\infty} \frac{\sqrt{1+x^2}}{x} = 1^+.$$

Le tableau suivant contient quelques exemples directs d'application du théorème de l'Hospital. Ils donnent le comportement de certaines fonctions simples à une extrémité de leur intervalle de définition. Elles servent bien souvent de limites de référence qui permettent de simplifier le calcul d'autres limites.

Limites remarquables pour $r > 0$ et $a \in \mathbb{R} \setminus \{0\}$	
$\lim\limits_{x\to 0} \dfrac{\sin(x)}{x} = 1$	$\lim\limits_{x\to 0} \dfrac{\mathrm{sh}(x)}{x} = 1$
$\lim\limits_{x\to 0} \dfrac{\mathrm{tg}(x)}{x} = 1$	$\lim\limits_{x\to 0} \dfrac{\mathrm{th}(x)}{x} = 1$
$\lim\limits_{x\to 0} \dfrac{\arcsin(x)}{x} = 1$	$\lim\limits_{x\to 0} \dfrac{\mathrm{arcsh}(x)}{x} = 1$
$\lim\limits_{x\to 0} \dfrac{\mathrm{arctg}(x)}{x} = 1$	$\lim\limits_{x\to 0} \dfrac{\mathrm{arcth}(x)}{x} = 1$
$\lim\limits_{x\to 0^+} x^x = 1$	$\lim\limits_{x\to +\infty} x^{1/x} = 1$
$\lim\limits_{x\to 0^+} x^r \ln(x) = 0$	$\lim\limits_{x\to +\infty} \dfrac{\ln(x)}{x^r} = 0$
$\lim\limits_{x\to -\infty} x^r \mathrm{e}^x = 0$	$\lim\limits_{x\to \infty} \dfrac{\mathrm{e}^x}{x^r} = +\infty$
$\lim\limits_{x\to 0} \dfrac{\ln(1+x)}{x} = 1$	$\lim\limits_{x\to +\infty} \left(1+\dfrac{a}{x}\right)^x = \mathrm{e}^a$

Remarques. a) La vérification de ces formules remarquables est directe et par conséquent, laissée aux étudiants.

b) Pour $r > 0$, les formules

$$\lim_{x\to 0^+} x^r \ln(x) = 0,\ \lim_{x\to +\infty} \frac{\ln(x)}{x^r} = 0,\ \lim_{x\to -\infty} |x|^r \mathrm{e}^x = 0,\ \text{et}\ \lim_{x\to +\infty} x^{-r}\mathrm{e}^x = +\infty$$

s'énoncent bien souvent de la manière suivante: "en 0^+ et en $+\infty$, les puissances positives de x dominent $\ln(x)$" et "en $-\infty$ et en $+\infty$, e^x domine les puissances positives de x". Elles ont des conséquences importantes en calcul intégral.

Exemple. *Calculer*

$$\lim_{x\to 0^+} \frac{\mathrm{tg}(x) - x}{x - \sin(x)}.$$

On vérifie aisément les hypothèses a), b) et c) du théorème de l'Hospital dans $]0, \pi/2[$. Cela étant, il vient

$$\lim_{x\to 0^+} \frac{\mathrm{D}f(x)}{\mathrm{D}g(x)} = \lim_{x\to 0^+} \frac{1/\cos^2(x) - 1}{1 - \cos(x)} \underset{(*)}{=} \lim_{x\to 0^+} \frac{1 + \cos(x)}{\cos^2(x)} = 2,$$

et ainsi la limite à calculer est égale à 2.

Remarquons que la division haut et bas par $1 - \cos(x)$ pour obtenir l'égalité $(*)$ a simplifié considérablement les calculs.

Exemple. Parfois il faut recommencer les opérations.
Calculer

$$\lim_{x \to 0^+} \frac{\sin(x) - x}{x^3}.$$

Après avoir vérifié les hypothèses a), b) et c) dans $]0, \pi[$, on obtient

$$\lim_{x \to 0^+} \frac{\mathrm{D}f(x)}{\mathrm{D}g(x)} = \lim_{x \to 0^+} \frac{\cos(x) - 1}{3x^2} \underset{(*)}{=} \lim_{x \to 0^+} \frac{-2\sin^2(x/2)}{12(x/2)^2} = -\frac{1}{6}.$$

En $(*)$ on pouvait aussi vérifier à nouveau les hypothèses a), b) et c) et noter que

$$\lim_{x \to 0^+} \frac{\mathrm{D}(\cos(x) - 1)}{\mathrm{D}(3x^2)} = -\lim_{x \to 0^+} \frac{\sin(x)}{6x} = -\frac{1}{6}.$$

Exemple. Parfois l'utilisation du théorème de l'Hospital ne mène à rien, mais on peut obtenir le résultat en passant à $(1/f)/(1/g)$.
Etant donné a, $b > 0$, calculer

$$\lim_{x \to 0^+} \frac{(\ln(ax))^{-1}}{(\ln(bx))^{-1}}.$$

On vérifie aisément les hypothèses a), b) et c) puis il vient

$$\lim_{x \to 0^+} \frac{\mathrm{D}f(x)}{\mathrm{D}g(x)} = \lim_{x \to 0^+} \frac{(\ln(ax))^{-2}}{(\ln(bx))^{-2}}$$

et on vérifie aisément que des essais ultérieurs ne mènent à rien.

Exemple. *Calculer*

$$\lim_{x \to 0^+} \frac{(1+x)^{1/x} - \mathrm{e}}{x}.$$

Comme on vérifie aisément par le théorème de l'Hospital que

$$\lim_{x \to 0^+} (1+x)^{1/x} = \mathrm{e};$$

nous sommes en présence du cas $0/0$. Après avoir vérifié les hypothèses a), b) et c), on calcule

$$\begin{aligned} \lim_{x \to 0^+} \frac{\mathrm{D}f(x)}{\mathrm{D}g(x)} &= \lim_{x \to 0^+} (1+x)^{1/x} \cdot \left(\frac{1}{x(1+x)} - \frac{\ln(1+x)}{x^2} \right) \\ &= \mathrm{e} \cdot \lim_{x \to 0^+} \frac{x - (1+x)\ln(1+x)}{x^2(1+x)} \end{aligned}$$

par un nouveau recours au théorème de l'Hospital. Après avoir vérifié les conditions a), b) et c), on calcule

$$\lim_{x\to 0^+} \frac{1-1-\ln(1+x)}{2x+3x^2}$$

au moyen d'une troisième application du théorème de l'Hospital. On vérifie les conditions a), b) et c) puis on note que

$$-\lim_{x\to 0^+} \frac{1}{(1+x)(2+6x)} = -\frac{1}{2}.$$

Au total, la limite envisagée existe et vaut $-\mathrm{e}/2$.

Exercice. Vérifier que la fonction

$$f\colon \mathbb{R} \to \mathbb{R} \quad x \mapsto \begin{cases} \mathrm{e}^{-1/x} & \text{si} \quad x > 0, \\ 0 & \text{si} \quad x \leq 0 \end{cases}$$

appartient à $\mathrm{C}_\infty(\mathbb{R})$.

7.2 Recherche des extrema

7.2.1 Extrema d'une fonction réelle

Définitions. Soit f une fonction réelle définie sur une partie A de $\mathbb{R}^n$.

Un point $x_0 \in A$ est un *extremum local de f dans A* et plus précisément un *maximum* (resp. *minimum*) *local de f dans A* s'il existe une boule b de centre x_0 telle que

$$f(x_0) \geq \text{(resp. } \leq \text{)}\ f(x), \quad \forall x \in b \cap A. \qquad (*)$$

Parfois on ne spécifie pas le mot local.

En outre, on dit que cet extremum (maximum ou minimum) est

strict si, dans $(*)$, l'égalité est exclue pour $x \neq x_0$;

global si $(*)$ est vrai pour toute boule b de centre x_0.

Remarquons qu'une fonction réelle définie sur $A \subset \mathbb{R}^n$ peut

a) ne pas avoir d'extremum dans A. C'est le cas pour la fonction $f\colon x \mapsto x$ dans $\mathbb{R}$.

b) avoir un nombre fini d'extrema dans A. C'est le cas pour la fonction $|\cdot|$ dans $\mathbb{R}$, pour laquelle $x_0 = 0$ est un minimum strict absolu.

c) avoir une infinité d'extrema dans A. C'est le cas pour la fonction cos dans $\mathbb{R}$, pour laquelle tous les points $2k\pi$ avec $k \in \mathbb{Z}$ sont des maxima stricts locaux et tous les points $\pi + 2k\pi$ avec $k \in \mathbb{Z}$ des minima stricts locaux.

On ne dispose pas de méthode générale permettant la recherche des extrema d'une fonction réelle sur une partie de $\mathbb{R}^n$. On se place dans le cas des fonctions dérivables sur un ouvert; le résultat suivant est à la base de cette étude.

Théorème 7.2.1.1 *Si f est une fonction réelle et dérivable sur l'ouvert Ω de $\mathbb{R}^n$ et si $x_0 \in \Omega$ est un extremum de f dans Ω, on a*

$$[\mathrm{D}_1 f]_{x_0} = \ldots = [\mathrm{D}_n f]_{x_0} = 0.$$

Preuve. De fait, pour $k \leq n$, il vient

$$[\mathrm{D}_k f]_{x_0} = \lim_{h \to 0,\, h \in \mathbb{R}_0} \frac{1}{h}(f(x_0 + he_k) - f(x_0))$$

où pour h suffisamment petit, le quotient change de signe avec h: la limite doit donc être nulle.∎

Malheureusement ce résultat n'admet pas de réciproque. Ainsi la fonction

$$f \colon \mathbb{R}^n \to \mathbb{R} \quad x \mapsto x_1^3 + \ldots + x_n^3$$

appartient à $\mathrm{C}_\infty(\mathbb{R}^n)$ et l'origine est le seul point de $\mathbb{R}^n$ où toutes ses dérivées d'ordre 1 sont nulles. Cependant 0 n'est bien sûr pas un extremum de f dans $\mathbb{R}^n$ puisque pour $x = he_1$ avec $h \in \mathbb{R}$, $f(he_1) = h^3$ change de signe avec h.

Définition. Si f est une fonction réelle et dérivable sur un ouvert Ω de $\mathbb{R}^n$, un *point stationnaire de f dans* Ω est un point de Ω où toutes les dérivées d'ordre 1 de f s'annulent.

On peut alors énoncer le résultat précédent de la manière équivalente suivante: *les extrema éventuels d'une fonction f réelle et dérivable sur un ouvert Ω de $\mathbb{R}^n$ sont des points stationnaires de f dans Ω.*

Remarque. Insistons sur le fait que cette règle ne permet pas de rechercher les extrema d'une fonction qui sont des points où cette fonction n'est pas dérivable. Voici deux exemples typiques où cette situation a lieu.

a) Considérons la fonction $f \colon [0,1] \to \mathbb{R} \quad x \mapsto x$. Elle est réelle et dérivable sur l'ouvert $]0,1[$ et sa dérivée est égale à 1 en tout point de cet ouvert: f n'admet donc pas d'extremum dans $]0,1[$. Cependant on voit directement que 0 est un minimum strict absolu et 1 est un maximum strict absolu de f dans $[0,1]$.

b) Etant donné $r > 0$, considérons la fonction

$$f \colon A = \{\, x \in \mathbb{R}^n : |x| \leq r \} \to \mathbb{R} \quad x \mapsto |x|\,.$$

Bien sûr, elle est réelle et dérivable sur l'ouvert $\{ x \in \mathbb{R}^n : 0 < |x| < r\}$ et, dans cet ouvert, le système

$$[\mathrm{D}_k f]_x = \frac{x_k}{|x|} = 0, \quad (k \leq n),$$

n'a pas de solution. Cependant il est clair que 0 est un minimum strict absolu et que tout $x \in \mathbb{R}^n$ tel que $|x| = r$ est un maximum absolu de f dans A.

7.2.2 Recherche des extrema d'une fonction

La méthode de recherche des extrema d'une fonction réelle et dérivable sur un ouvert Ω de $\mathbb{R}^n$ consiste alors à déterminer les points stationnaires de cette fonction dans Ω puis à examiner chacun de ces points stationnaires. La première partie consiste en la résolution du système d'équations (généralement non linéaires)

$$\mathrm{D}_k f = 0, \quad (k \leq n),$$

ce qui peut être particulièrement redoutable. La deuxième partie peut être allégée au moyen des résultats qui vont suivre; les cas $n = 1$ et $n > 1$ étant différents, nous allons les traiter séparément.

Recherche des extrema d'une fonction réelle et dérivable sur un ouvert de $\mathbb{R}$

Théorème 7.2.2.1 *Soit f une fonction réelle et dérivable sur l'intervalle $]a, b[$ de $\mathbb{R}$. Un point stationnaire x_0 de f dans $]a, b[$*

a) *est un maximum* (resp. *minimum*) *local de f dans $]a, b[$ s'il existe $\varepsilon > 0$ tel que*

$$x_0 - \varepsilon \leq x < x_0 \quad \Rightarrow \quad [\mathrm{D}f]_x \geq \text{ (resp. } \leq) \ 0$$

et

$$x_0 < x \leq x_0 + \varepsilon \quad \Rightarrow \quad [\mathrm{D}f]_x \leq \text{ (resp. } \geq) \ 0.$$

b) *n'est pas un extremum de f dans $]a, b[$ s'il existe $\varepsilon > 0$ tel que $\mathrm{D}f$ garde un signe constant sur $[x_0 - \varepsilon, x_0 + \varepsilon] \setminus \{x_0\}$.*

Preuve. C'est une conséquence directe des propriétés relatives à la croissance et à la décroissance des fonctions réelles et dérivables sur un intervalle ouvert de $\mathbb{R}$. ∎

On peut compléter ce résultat si f est p fois continûment dérivable avec $p > 1$.

Théorème 7.2.2.2 *Si p est un entier strictement supérieur à 1, si f est une fonction réelle appartenant à $\mathrm{C}_p(]x_0 - \varepsilon, x_0 + \varepsilon[)$ avec $\varepsilon > 0$ et si*

$$[\mathrm{D}f]_{x_0} = \ldots = [\mathrm{D}^{p-1} f]_{x_0} = 0 \quad et \quad [\mathrm{D}^p f]_{x_0} \neq 0,$$

a) x_0 *est un extremum local de* f *dans* $]x_0 - \varepsilon, x_0 + \varepsilon[$ *si* p *est pair. C'est un maximum si* $[\mathrm{D}^p f]_{x_0} < 0$ *et c'est un minimum si* $[\mathrm{D}^p f]_{x_0} > 0$.
b) x_0 *n'est pas un extremum de* f *dans* $]x_0 - \varepsilon, x_0 + \varepsilon[$ *si* p *est impair.*

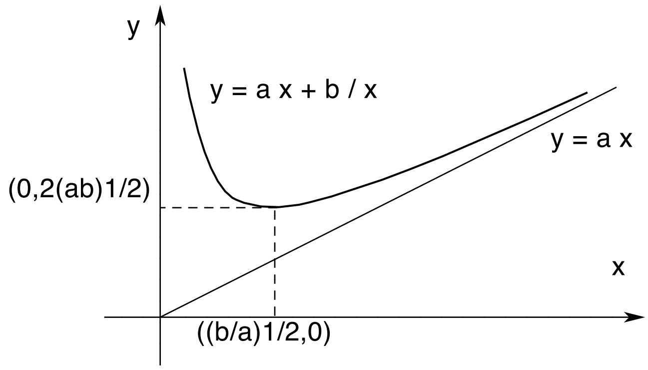
Preuve. Quitte à remplacer ε par $\varepsilon' \in]0, \varepsilon[$, nous pouvons supposer que la fonction continue $\mathrm{D}^p f$ garde un signe constant sur $]x_0 - \varepsilon, x_0 + \varepsilon[$. Vu le théorème de Taylor et l'hypothèse, pour tout $h \in]-\varepsilon, \varepsilon[$, il existe $y \in]x_0 - \varepsilon, x_0 + \varepsilon[$ tel que

$$f(x_0 + h) = f(x_0) + \frac{h^p}{p!}[\mathrm{D}^p f]_y.$$

La conclusion s'ensuit directement.∎

Exercice. *Etant donné* a, $b > 0$, *rechercher les extrema et discuter la croissance de la fonction*

$$f\colon\ ]0, +\infty[\ \to \mathbb{R} \quad x \mapsto ax + b/x.$$



Suggestion. Bien sûr, f appartient à $\mathrm{C}_\infty(]0, +\infty[)$ et on a

$$[\mathrm{D}f]_x = a - bx^{-2}, \quad \forall x \in]0, +\infty[\,.$$

Dès lors, il vient
a) $[\mathrm{D}f]_x > 0$ pour tout $x \in \left](b/a)^{1/2}, +\infty\right[$,
b) $[\mathrm{D}f]_x = 0$ si et seulement si $x = (b/a)^{1/2}$,
c) $[\mathrm{D}f]_x < 0$ pour tout $x \in \left]0, (b/a)^{1/2}\right[$.

On en déduit aussitôt les renseignements suivants:
a) f est strictement décroissant sur $\left]0, (b/a)^{1/2}\right[$ avec $\lim_{x \to 0^+} f(x) = +\infty$,
b) $x_0 = (b/a)^{1/2}$ est un minimum strict absolu,
c) f est strictement croissant sur $\left](b/a)^{1/2}, +\infty\right[$ et $\lim_{x \to +\infty} f(x) = +\infty$.

$* \to$ Ceci implique notamment que $y = ax + b/x$ est un changement de variable régulier d'ordre infini entre $\left]0, (b/a)^{1/2}\right[$ et $\left]2(ab)^{1/2}, +\infty\right[$ d'une part et entre $\left](b/a)^{1/2}, +\infty\right[$ et $\left]2(ab)^{1/2}, +\infty\right[$ d'autre part. $\leftarrow *$∎

Exercice. *Rechercher les extrema et discuter la croissance de la fonction*

$$f\colon\ ]0,+\infty[\ \to \mathbb{R} \quad x \mapsto (1+1/x)^x.$$

Suggestion. Bien sûr, $f(x)$ est défini et strictement positif en tout $x \in\ ]0,+\infty[$, et appartient à $C_\infty(]0,+\infty[)$. De plus, on a

$$[Df]_x = f(x)\cdot\left(\ln(1+\frac{1}{x}) - \frac{1}{1+x}\right), \quad \forall x \in\ ]0,+\infty[\,.$$

Le signe de cette dérivée en x est donc donné par le signe de $\ln(1+1/x)-1/(1+x)$, qui est difficile à trouver directement. On peut cependant procéder comme suit. D'une part,

$$\lim_{x\to+\infty}\left(\ln(1+\frac{1}{x}) - \frac{1}{1+x}\right) = 0$$

et d'autre part,

$$D\left(\ln(1+\frac{1}{x}) - \frac{1}{1+x}\right) = \frac{-1}{x(1+x)^2}, \quad \forall x \in\ ]0,+\infty[\,.$$

Dès lors, $\ln(1+1/x)-1/(1+x)$ est une fonction strictement décroissante sur $]0,+\infty[$ qui tend vers 0 si $x \to +\infty$: elle est donc strictement positive. Il s'ensuit que Df ne s'annule en aucun point de $]0,+\infty[$ et même, plus précisément, que f est une fonction strictement croissante sur $]0,+\infty[$.

$* \to$ Comme on vérifie directement au moyen du théorème de l'Hospital que

$$\lim_{x\to0^+}(1+1/x)^x = 1 \quad \text{et} \quad \lim_{x\to+\infty}(1+1/x)^x = \mathrm{e},$$

on obtient que f est un changement de variable régulier d'ordre infini entre $]0,+\infty[$ et $]1,\mathrm{e}[$. $\leftarrow *$ ∎

Exercice. *Etablir que toute fonction réelle, continue et non monotone sur un intervalle ouvert de $\mathbb{R}$ a un extremum local.*

Suggestion. Soit f une fonction réelle, continue et non monotone sur $]a,b[$.

S'il existe un intervalle compact $[c,d]$ inclus dans l'intervalle ouvert $]a,b[$ tel que $\inf\{\,f(x) : x\in[c,d]\}$ et $\sup\{\,f(x) : x\in[c,d]\}$ ne soient réalisés ni en c ni en d, c'est trivial.

Sinon, pour tout intervalle compact $[c,d]$ inclus dans $]a,b[$, f réalise ses bornes supérieure et inférieure sur $[c,d]$ en c ou en d. On en déduit de suite que f est monotone sur $]a,b[$ car
a) ou f est constant sur $]a,b[$,
b) ou il existe $c,\ d \in\ ]a,b[$ tels que $f(c)\neq f(d)$. Si on a par exemple, $f(c)<f(d)$, on obtient aisément que f est croissant sur $]a,b[$ en recourant au théorème des valeurs intermédiaires. ∎

Recherche des extrema d'une fonction réelle et dérivable sur un ouvert de $\mathbb{R}^n$

Soit x_0 un point stationnaire d'une fonction f réelle et dérivable sur un ouvert de $\mathbb{R}^n$.

Avant d'appliquer les résultats qui vont suivre, il convient de voir si, pour une raison immédiate, x_0 est ou n'est pas un extremum de f. Ainsi la fonction $|\cdot|^2$ appartient à $\mathrm{C}_\infty(\mathbb{R}^n)$, est réelle et admet l'origine pour seul point stationnaire. Or $x_0 = 0$ est évidemment un minimum strict absolu de f sur $\mathbb{R}^n$.

Rappelons que, si f est une fonction réelle appartenant à $\mathrm{C}_2(\Omega)$ où Ω est un ouvert de $\mathbb{R}^n$, la *matrice hessienne associée à f en $x \in \Omega$* est la matrice $H_f(x)$ de dimension $n \times n$, définie sur Ω par

$$[H_f(x)]_{j,k} = [\mathrm{D}_j \mathrm{D}_k f]_x, \quad \forall j, k \leq n.$$

Il s'agit bien sûr d'une matrice réelle et symétrique, donc hermitienne. Cela étant, pour tout point stationnaire x_0 de f dans Ω et tout élément h de $\mathbb{R}^n$ tel que le segment $\{\, x_0 + th : t \in [0,1]\}$ soit inclus dans Ω, le théorème de Taylor affirme qu'il existe $\theta \in]0,1[$ tel que

$$f(x_0 + h) = f(x_0) + \frac{1}{2} \langle H_f(x_0 + \theta h)h, h \rangle .$$

Effectuons également un petit rappel de définitions et propriétés d'algèbre matricielle. Etant donné une matrice hermitienne réelle H, il existe une matrice orthogonale (c'est-à-dire réelle et unitaire) U qui transforme H en $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, c'est-à-dire telle que $UH\tilde{U} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, où $\lambda_1, \ldots, \lambda_n$ sont les différentes valeurs propres de H, comptées avec leur multiplicité. On distingue les cas suivants: H est

a) *défini positif*, en abrégé dp, si on a $\lambda_j > 0$ pour tout $j \leq n$,

b) *défini négatif*, en abrégé dn, si on a $\lambda_j < 0$ pour tout $j \leq n$,

c) *semi-défini positif*, en abrégé sdp, si on a $\lambda_j \geq 0$ pour tout $j \leq n$,

d) *semi-défini négatif*, en abrégé sdn, si on a $\lambda_j \leq 0$ pour tout $j \leq n$,

e) *indéfini* dans les autres cas, c'est-à-dire s'il existe deux valeurs propres λ_j et λ_k de H, de signes différents.

Théorème 7.2.2.3 *Soit $x_0 \in \Omega$ un point stationnaire de la fonction réelle $f \in \mathrm{C}_2(\Omega)$.*

a) *Si la matrice $H_f(x_0)$ est* dp (resp. dn)*, x_0 est un minimum* (resp. *maximum*) *strict local de f dans Ω.*

b) *S'il existe une boule b de centre x_0 incluse dans Ω telle que $H_f(x)$ soit* sdp (resp. sdn) *pour tout $x \in b$, alors x_0 est un minimum* (resp. *maximum*) *local de f dans Ω.*

c) *Si la matrice $H_f(x_0)$ est* sdp (resp. sdn) *et s'il existe une boule b de centre x_0 incluse dans Ω telle que $H_f(x)$ soit* dp (resp. dn) *pour tout $x \in b \setminus \{x_0\}$, alors x_0 est un minimum* (resp. *maximum*) *strict local de f dans Ω.*

d) *Si la matrice $H_f(x_0)$ est indéfinie, alors x_0 n'est pas un extremum local de f dans Ω.*

Preuve. a) Rappelons qu'une matrice hermitienne est
i) dp si et seulement si les déterminants des sous-matrices carrées occupant le coin supérieur gauche sont tous strictement positifs;
ii) dn si et seulement si les déterminants des sous-matrices carrées occupant le coin supérieur gauche sont strictement négatifs si leur dimension est impaire et strictement positifs si leur dimension est paire.

Cela étant, établissons le cas où $H_f(x_0)$ est dp; l'autre se démontre de manière analogue.

Pour tout $j \leq n$, le déterminant de la matrice carrée de dimension j occupant le coin supérieur gauche de $H_f(x_0)$ est donc strictement positif. Par continuité, il existe alors $r_j > 0$ tel que le déterminant de cette matrice soit strictement positif pour tout $x \in \mathbb{R}^n$ tel que $|x - x_0| \leq r_j$. Dès lors, la matrice $H_f(x)$ est dp pour tout élément x de $\mathbb{R}^n$ tel que $|x - x_0| \leq \inf\{r_1, \cdots, r_n\}$. D'où la conclusion par la formule de Taylor car on sait que, si H est une matrice hermitienne définie positive de dimension $n \times n$, on a $\langle Hh, h\rangle > 0$ pour tout $h \in \mathbb{R}^n \setminus \{0\}$.

b) et c) résultent directement de la formule de Taylor et de la valeur de $\langle Hh, h\rangle$ pour tout $h \in \mathbb{R}^n$ si H est une matrice hermitienne dp, dn, sdp ou sdn de dimension $n \times n$.

d) Comme $H_f(x_0)$ est une matrice hermitienne indéfinie, elle admet au moins une valeur propre $\lambda > 0$ et une valeur propre $\lambda' < 0$. Comme toute valeur propre d'une matrice carrée a au moins un vecteur propre, il existe h, $h' \in \mathbb{R}^n \setminus \{0\}$ tels que $H_f(x_0)h = \lambda h$ et $H_f(x_0)h' = \lambda' h'$. On a donc notamment $\langle H_f(x_0)h, h\rangle > 0$ et $\langle H_f(x_0)h', h'\rangle < 0$. Cela étant, vu la formule de Taylor, on obtient aisément

$$\lim_{r \to 0^+} \frac{1}{r^2} \left(f(x_0 + rh) - f(x_0)\right) = \frac{1}{2} \langle H_f(x_0)h, h\rangle > 0$$

et

$$\lim_{r \to 0^+} \frac{1}{r^2} \left(f(x_0 + rh') - f(x_0)\right) = \frac{1}{2} \langle H_f(x_0)h', h'\rangle < 0,$$

ce qui suffit car, pour $r > 0$ suffisamment petit, il vient alors

$$f(x_0 + rh') < f(x_0) < f(x_0 + rh). \blacksquare$$

Remarque. Si x_0 est un point stationnaire de la fonction réelle $f \in \mathrm{C}_2(\Omega)$, si $H_f(x_0)$ est sdp (resp. sdn) et s'il n'existe pas de boule b de centre x_0 incluse dans Ω telle que $H_f(x)$ soit sdp (resp. sdn) pour tout $x \in b$, x_0 peut être ou ne pas être un extremum: on ne peut pas conclure, il faut envisager chaque cas en particulier. Les deux exemples suivants le montrent à suffisance.

Exemple. Considérons la fonction

$$f\colon \mathbb{R}^2 \to \mathbb{R} \quad (x, y) \mapsto x^4 - y^4.$$

Bien sûr, f est réel et appartient à $\mathrm{C}_\infty(\mathbb{R}^2)$. De plus, $(0,0)$ en est le seul stationnaire. Enfin

$$H_f(x, y) = 12 \begin{pmatrix} x^2 & 0 \\ 0 & -y^2 \end{pmatrix}$$

est sd en $(0,0)$ mais il n'existe pas de boule b centrée en $(0,0)$ telle que $H_f(x, y)$ soit sdp (resp. sdn) en tout $(x, y) \in b$. Cependant, il est trivial que $(0,0)$ n'est pas un extremum local de f dans $\mathbb{R}^2$.

Exemple. Considérons la fonction

$$f\colon \mathbb{R}^2 \to \mathbb{R} \quad (x, y) \mapsto (x^2 + y^2 - 1)^2.$$

Bien sûr, f est réel et appartient à $\mathrm{C}_\infty(\mathbb{R}^2)$; de plus, ses points stationnaires sont les solutions du système

$$\begin{cases} 4x(x^2 + y^2 - 1) = 0 \\ 4y(x^2 + y^2 - 1) = 0, \end{cases}$$

à savoir le point $(0,0)$ et chacun des points de la circonférence $\{\, (x, y) : x^2 + y^2 = 1\}$. Enfin la matrice $H_f(x, y)$ s'écrit

$$H_f(x, y) = \begin{pmatrix} 4(x^2 + y^2 - 1) + 8x^2 & 8xy \\ 8xy & 4(x^2 + y^2 - 1) + 8y^2 \end{pmatrix}$$

pour tout $(x, y) \in \mathbb{R}^2$. Il est alors aisé de vérifier que $4(x^2+y^2-1)$ et $12(x^2+y^2)-4$ sont ses valeurs propres. Cela étant, comme $H_f(0,0)$ est dn, il est clair que $(0,0)$ est un maximum local strict de f dans $\mathbb{R}^2$. De plus, en tout point (x, y) de la circonférence $\{\, (x, y) : x^2 + y^2 = 1\}$, $H_f(x, y)$ est sdp mais ne reste sdp sur aucune boule de centre (x, y). Cependant la fonction f ne prend que des valeurs supérieures ou égales à 0 sur $\mathbb{R}^2$ et s'annule en tout point de cette circonférence, donc tout point de cette circonférence est un minimum local global.

Exercice. Rechercher les extrema de la fonction

$$f\colon \Omega =]0, +\infty[^4 \to \mathbb{R} \quad (x, y, z, t) \mapsto xyzt + 1/x + 1/y + 1/z + 1/t.$$

Suggestion. On a vite fait de vérifier que f est réel et appartient à $C_\infty(\Omega)$. Cela étant, ses points stationnaires dans Ω sont les solutions du système

$$\left\{\begin{array}{l} yzt - x^{-2} = 0 \\ xzt - y^{-2} = 0 \\ xyt - z^{-2} = 0 \\ xyz - t^{-2} = 0 \end{array}\right.$$

qui appartiennent à Ω. Comme, dans Ω, x, y, z et t diffèrent de 0, il s'agit des solutions du système

$$xyzt = \frac{1}{x} = \frac{1}{y} = \frac{1}{z} = \frac{1}{t}$$

qui appartiennent à Ω. On en déduit de suite que $(1,1,1,1)$ est le seul point stationnaire de f dans Ω.

Calculons à présent $H_f(1,1,1,1)$: il vient directement

$$H_f(1,1,1,1) = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}.$$

Un calcul aisé prouve alors que 1 est une valeur propre de multiplicité 3 et 5 une valeur propre de multiplicité 1 de cette matrice. Au total, $(1,1,1,1)$ est un minimum strict local (il est même global) de f dans Ω.

Exercice. *Rechercher les extrema de la fonction*

$$f\colon \Omega =]0,\pi[^2 \to \mathbb{R} \quad (x,y) \mapsto \sin(x) + \sin(y) + \sin(x+y).$$

Suggestion. Il est immédiat que f est réel sur Ω et appartient à $C_\infty(\Omega)$. Ses points stationnaires dans Ω sont les solutions du système

$$\left\{\begin{array}{l} \cos(x) + \cos(x+y) = 0 \\ \cos(y) + \cos(x+y) = 0 \end{array}\right.$$

qui appartiennent à Ω. Comme ce système est équivalent à

$$\left\{\begin{array}{l} \cos(x) + \cos(x+y) = 0 \\ \cos(x) = \cos(y) \end{array}\right.$$

et comme la fonction cos est strictement décroissante sur $]0,\pi[$, il s'agit des solutions du système

$$\left\{\begin{array}{l} x = y \\ \cos(x) + \cos(2x) = 0 \end{array}\right.$$

qui appartiennent à Ω. Comme les solutions de $2\cos^2(x)+\cos(x)-1=0$ sont $\cos(x)=-1$ et $\cos(x)=1/2$, on déduit que $(\pi/3,\pi/3)$ est le seul point stationnaire de f dans Ω. On a alors vite fait de vérifier que

$$H_f(\pi/3,\pi/3)=\frac{\sqrt{3}}{2}\begin{pmatrix}-2 & -1\\ -1 & -2\end{pmatrix}$$

et que cette matrice est dn. Dès lors $(\pi/3,\pi/3)$ est un maximum strict local (il est même global) de f dans Ω.

7.3 Opérateurs de dérivation

7.3.1 Généralités

Définitions. Un *opérateur de dérivation* dans un ouvert Ω de $\mathbb{R}^n$ est la donnée d'un nombre fini de dérivées $\mathrm{D}^{\alpha_1}, \ldots, \mathrm{D}^{\alpha_m}$ dans $\mathbb{R}^n$ et d'une fonction F définie sur un ouvert ω de $\Omega\times\mathbb{R}^m$. Etant donné une fonction réelle $f\in \mathrm{C}_p(\Omega)$ avec $p\geq|\alpha_j|$ pour $j=1,\ldots,m$, l'*action* de cet opérateur sur f est la fonction

$$F(x,\mathrm{D}^{\alpha_1}f,\ldots,\mathrm{D}^{\alpha_m}f)$$

sur $\{\, x\in\Omega : (x,\mathrm{D}^{\alpha_1}f(x),\ldots,\mathrm{D}^{\alpha_m}f(x))\in\omega\}$. Elle est aussi notée Ff si aucune confusion n'est possible.

L'*ordre* d'un opérateur de dérivation est le plus petit entier p nécessaire pour pouvoir évaluer son action.

Exemples. Les lois qui, à $f\in\mathrm{C}_1(\Omega)$ (resp. $\mathrm{C}_2(\Omega)$; $\mathrm{C}_3(]a,b[)$) associent

$$(\mathrm{D}_1f)^2+\ldots+(\mathrm{D}_nf)^2$$

$$\left(\text{resp.}\quad \mathrm{D}_1^2f+\ldots+\mathrm{D}_n^2f;\quad \frac{\mathrm{D}^3f}{\mathrm{D}f}+\frac{3x}{2}\left(\frac{\mathrm{D}^2f}{\mathrm{D}f}\right)^2\right)$$

sont des opérateurs de dérivation sur Ω (resp. Ω; $]a,b[$), d'ordre 1 (resp. 2; 3). On les note explicitement

$$\sum_{k=1}^n(\mathrm{D}_k\cdot)^2\quad\left(\text{resp.}\ \sum_{k=1}^n\mathrm{D}_k^2\cdot;\quad\frac{\mathrm{D}^3\cdot}{\mathrm{D}\cdot}+\frac{3x}{2}\left(\frac{\mathrm{D}^2\cdot}{\mathrm{D}\cdot}\right)^2\right),$$

$\cdot$ désignant la place où f doit être écrit. L'emplacement de ce point est alors très important, aussi on préfère bien souvent écrire le résultat de l'action de l'opérateur de dérivation sur une fonction générique. Ainsi les opérateurs $x\mathrm{D}\cdot$ et $\mathrm{D}(x\cdot)$ de dérivation sur $\mathbb{R}$ sont essentiellement differents.

7.3.2 Opérations

Définitions. Soient F, F', $F_1, \ldots, F_J$ des opérateurs de dérivation dans un même ouvert Ω de $\mathbb{R}^n$, avec $J \in \mathbb{N}_0$.

a) Les opérateurs F et F' sont *égaux* s'ils ont le même ordre p et si on a $Ff = F'f$ pour tout $f \in \mathrm{C}_p(\Omega)$.

Ainsi, on voit de suite que les opérateurs de dérivation

$$\mathrm{D}^2\cdot + \frac{n-1}{x}\mathrm{D}\cdot, \quad \frac{1}{x^{n-1}}\mathrm{D}(x^{n-1}\mathrm{D}\cdot) \ \text{ et } \ \frac{1}{x^2}((x\mathrm{D})(x\mathrm{D}) + (n-2)x\mathrm{D})\cdot$$

sont égaux sur $]0, +\infty[$.

b) Le *conjugué* de F est l'opérateur de dérivation $\overline{F}$ dans Ω défini par l'identité $\overline{F}f = \overline{(F\overline{f})}$, c'est-à-dire également $F^- f^- = (Ff)^-$; il a même ordre que F.

L'opérateur F est *réel* si on a $F = \overline{F}$.

c) Etant donné $c_1, \ldots, c_J \in \mathbb{C}$, la *combinaison linéaire* $\sum_{j=1}^J c_j F_j$ est l'opérateur de dérivation dans Ω défini par $(\sum_{j=1}^J c_j F_j)f = \sum_{j=1}^J c_j F_j f$.

On vérifie de suite l'égalité

$$\left(\sum_{j=1}^J c_j F_j\right)^- = \sum_{j=1}^J \overline{c_j}\,\overline{F_j}.$$

d) Le *produit* $F_1 \ldots F_J$ est l'opérateur de dérivation dans Ω défini par

$$(F_1 \ldots F_J)f = F_1(\ldots(F_J(f))).$$

En particulier, si m appartient à $\mathbb{N}_0$, la *m-ème puissance* F^m de F est l'opérateur de dérivation défini par $F^1 = F$ et $F^m = FF^{m-1}$ pour $m > 1$.

On vérifie de suite l'égalité $(F_1 \ldots F_J)^- = (F_1)^- \ldots (F_J)^-$.

En général, on n'a pas $FF' = F'F$. Ainsi $\mathrm{D}(x\mathrm{D})\cdot$ et $x\mathrm{D}(\mathrm{D}\cdot)$ ne sont évidemment pas égaux. Par exception, on dit que F et F' *commutent* si on a $FF' = F'F$.

7.3.3 Linéaires à coefficients constants

Définition. Un opérateur de dérivation F d'ordre p dans l'ouvert Ω de $\mathbb{R}^n$ est *linéaire* s'il vérifie

$$F\left(\sum_{j=1}^J c_j f_j\right) = \sum_{j=1}^J c_j F f_j$$

pour toute combinaison linéaire d'éléments f_j de $\mathrm{C}_p(\Omega)$.

Si l'opérateur F est à la fois linéaire et réel, il vient

$$\Re(Ff) = \frac{1}{2}(Ff + \overline{Ff}) = F\left(\frac{1}{2}(f + \overline{f})\right) = F(\Re f),$$

et

$$\Im(Ff) = \frac{1}{2i}(Ff - \overline{Ff}) = F\left(\frac{1}{2i}(f - \overline{f})\right) = F(\Im f).$$

Introduisons à présent des opérateurs de dérivation très importants.

Définition. Toute combinaison linéaire de dérivées D^α avec $\alpha \in \mathbb{N}^n$ est évidemment un opérateur de dérivation linéaire dans tout ouvert Ω de $\mathbb{R}^n$ et son ordre est égal à la borne supérieure des $|\alpha| = \alpha_1 + \ldots + \alpha_n$ qui y interviennent. Elle est notée soit explicitement:

$$\sum_{(\alpha)} c_\alpha \mathrm{D}^\alpha$$

où $\sum_{(\alpha)}$ indique que la somme est finie, soit encore L. Un tel opérateur est appelé *opérateur de dérivation linéaire à coefficients constants.*

Proposition 7.3.3.1 *On a* $(\sum_{(\alpha)} c_\alpha \mathrm{D}^\alpha)^- = \sum_{(\alpha)} \overline{c_\alpha} \mathrm{D}^\alpha$.

Un tel opérateur $\sum_{(\alpha)} c_\alpha \mathrm{D}^\alpha$ *est réel si et seulement si tous ses coefficients sont réels.*∎

Proposition 7.3.3.2 *Les opérateurs de dérivation linéaires à coefficients constants commutent.*∎

Définition. A l'opérateur de dérivation $\sum_{(\alpha)} c_\alpha \mathrm{D}^\alpha$ linéaire à coefficients constants, on associe son *polynôme caractéristique*, à savoir le polynôme $\sum_{(\alpha)} c_\alpha x^\alpha$ où on a posé $x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ pour tout $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$.

Ainsi $x_1^2 + 2x_1x_2 + x_2^2$ est le polynôme caractéristique de $\mathrm{D}_1^2 + 2\mathrm{D}_1\mathrm{D}_2 + \mathrm{D}_2^2$. Remarquons qu'inversement la donnée du polynôme caractéristique détermine l'opérateur auquel il est associé.

Proposition 7.3.3.3 a) *Le polynôme caractéristique d'une combinaison linéaire d'opérateurs linéaires à coefficients constants est la combinaison linéaire correspondante des polynômes caractéristiques.*

b) Le polynôme caractéristique d'un produit fini d'opérateurs linéaires à coefficients constants est le produit correspondant des polynômes caractéristiques.

Preuve. De fait, les règles qui régissent les combinaisons linéaires ainsi que les produits finis des polynômes et des opérateurs de dérivation linéaires à coefficients constants sont analogues.∎

Corollaire 7.3.3.4 *Toute identité algébrique entre polynômes, qui ne fait intervenir que des combinaisons linéaires et des produits finis, donne lieu à une identité entre les opérateurs de dérivation linéaires à coefficients constants qu'ils caractérisent.*∎

Exemples. Ainsi, pour tout $c \in \mathbb{C}$, $(x-c)(x+c) = x^2 - c^2$ est une identité sur $\mathbb{R}$; par conséquent, on a l'identité

$$(\mathrm{D} - c)(\mathrm{D} + c) = \mathrm{D}^2 - c^2.$$

De même, l'identité $(x_1 - x_2)(x_1 + x_2) = x_1^2 - x_2^2$ entraîne l'identité

$$(\mathrm{D}_1 - \mathrm{D}_2)(\mathrm{D}_1 + \mathrm{D}_2) = \mathrm{D}_1^2 - \mathrm{D}_2^2.$$

7.3.4 Exercices

Exercice. *Etablir que pour tout $m \in \mathbb{Z}$ et tout $k \in \mathbb{N}_0$, on a*

$$\big(x^2\mathrm{D} - (m-1)x\big)^k \cdot = x^{m+k}\mathrm{D}^k(x^{k-m}\cdot)$$

Suggestion. Procédons par récurrence. Comme on a

$$x^{m+1}\mathrm{D}(x^{1-m}\cdot) = (1-m)x \cdot + x^2\mathrm{D}\cdot, \quad \forall m \in \mathbb{Z},$$

c'est vrai pour $k = 1$ et tout $m \in \mathbb{Z}$.Cela étant, supposons l'identité exacte pour tout $k = 1, \ldots, l$ et tout $m \in \mathbb{Z}$ et établissons-la pour $k = l+1$ et tout $m \in \mathbb{Z}$. Il vient successivement

$$\begin{aligned}\big(x^2\mathrm{D} - (m-1)x\big)^{l+1} &= \big(x^2\mathrm{D} - (m-1)x\big)\left(x^{m+l}\mathrm{D}^l(x^{l-m}\cdot)\right) \\ &= (1+l)x^{m+l+1}\mathrm{D}^l(x^{l-m}\cdot) + x^{l+m+2}\mathrm{D}^{l+1}(x^{l-m}\cdot)\end{aligned}$$

et

$$\mathrm{D}^{l+1}\big(x(x^{l-m}\cdot)\big) = x\mathrm{D}^{l+1}(x^{l-m}\cdot) + (l+1)\mathrm{D}^l(x^{l-m}\cdot),$$

ce qui suffit.

Exercice. *Etablir que, pour tout $m \in \mathbb{Z}$ et tout $k \in \mathbb{N}_0$, on a*

$$\big(x\mathrm{D}^2 - (m-1)\mathrm{D}\big)^k \cdot = x^m\mathrm{D}^k(x^{k-m}\mathrm{D}^k\cdot)$$

Suggestion. Procédons par récurrence. Pour $k = 1$ et tout $m \in \mathbb{Z}$, on a évidemment

$$x^m\mathrm{D}(x^{1-m}\mathrm{D}\cdot) = x\mathrm{D}^2\cdot + (1-m)\mathrm{D}\cdot$$

Cela étant, supposons l'identité exacte pour tout $k = 1, \ldots, l-1$ et tout $m \in \mathbb{Z}$ et établissons-la pour $k = l$ et tout $m \in \mathbb{Z}$. Il vient successivement

$$\begin{aligned}&\left(x\mathrm{D}^2 - (m-1)\mathrm{D}\right)^l \cdot \\ &= \left(x\mathrm{D}^2 - (m-1)\mathrm{D}\right)\left(x^m\mathrm{D}^{l-1}(x^{l-1-m}\mathrm{D}^{l-1}\cdot)\right) \\ &= x^m\left(x\mathrm{D}^{l+1}(x^{l+1-m}\mathrm{D}^{l-1}\cdot) + (m+1)\mathrm{D}^l(x^{l-1-m}\mathrm{D}^{l-1}\cdot)\right)\end{aligned}$$

et, par application de la formule $\mathrm{D}^l(x\cdot) = x\mathrm{D}^l\cdot + l\mathrm{D}^{l-1}\cdot$ qui donne lieu à

$$x\mathrm{D}^{l+1}(x^{l+1-m}\mathrm{D}^{l-1}\cdot) = \mathrm{D}^l\big(x\mathrm{D}(x^{l-1-m}\mathrm{D}^{l-1}\cdot)\big) - l\mathrm{D}^l(x^{l-1-m}\mathrm{D}^{l-1}\cdot),$$

il vient

$$x^m\left(\mathrm{D}^l\big(x\mathrm{D}(x^{l-1-m}\mathrm{D}^{l-1}\cdot)\big) + (m+1-l)\mathrm{D}^l(x^{l-1-m}\mathrm{D}^{l-1}\cdot)\right)$$

Il suffit alors d'effectuer $\mathrm{D}(x^{l-1-m}\mathrm{D}^{l-1}\cdot)$ pour conclure.

7.3.5 Changement de variable

Définitions. Soient Ω et Ω' deux ouverts de $\mathbb{R}^n$ et soit p un entier supérieur ou égal à 1, ou $p = \infty$. Un *changement de variable régulier d'ordre p entre Ω et Ω'* est une bijection $\varphi\colon \Omega \to \Omega'$ entre ces deux ouverts telle que

$$[\varphi]_1, \ldots, [\varphi]_n \in \mathrm{C}_p(\Omega) \quad \text{et} \quad [\varphi^{-1}]_1, \ldots, [\varphi^{-1}]_n \in \mathrm{C}_p(\Omega').$$

Très souvent, on pose $\varphi(x) = x'(x)$ et $\varphi^{-1}(x') = x(x')$ et on écrit

$$\begin{aligned} x'(x) \in \mathrm{C}_p(\Omega) &\quad \text{à la place de} \quad [\varphi]_1, \ldots, [\varphi]_n \in \mathrm{C}_p(\Omega), \\ x(x') \in \mathrm{C}_p(\Omega') &\quad \text{à la place de} \quad [\varphi^{-1}]_1, \ldots, [\varphi^{-1}]_n \in \mathrm{C}_p(\Omega'). \end{aligned}$$

On parle alors du changement de variable $x(x')$ entre les ouverts Ω et Ω', $x'(x)$ étant appelé le changement de variable *inverse* entre Ω' et Ω.

Exemples. a) Le premier exemple de changement de variable est particulièrement trivial; il est cependant rendu très important par ses conséquences théoriques.

Le *changement de variable identité* sur un ouvert Ω de $\mathbb{R}^n$ est l'application identité de Ω sur Ω. On vérifie de suite qu'il s'agit bien d'un changement de variable régulier d'ordre ∞ entre Ω et Ω.

b) *Composition de changements de variable.* Si $x(x')$ est un changement de variable régulier d'ordre p entre les ouverts Ω et Ω' de $\mathbb{R}^n$ et si $x'(x'')$ est un changement de variable régulier d'ordre p entre les ouverts Ω' et Ω'' de $\mathbb{R}^n$, alors

$$x_j\left(x'_1(x''), \ldots, x'_n(x'')\right), \quad (j = 1, \ldots, n),$$

est un changement de variable régulier d'ordre p entre les ouverts Ω et Ω''. Cela résulte aussitôt des propriétés de la composition des bijections et du théorème de dérivation multiple des fonctions composées. Ce nouveau changement de variable est noté $x(x'(x''))$.

c) *Changement de variable dans* $\mathbb{R}$. Le théorème de la fonction inverse donne immédiatement lieu à l'interprétation suivante. Si la fonction réelle $x(x')$ appartient à $\mathrm{C}_p(]a', b'[)$ avec $p \geq 1$ et si $\mathrm{D}_{x'}x(x')$ diffère de 0 en tout $x' \in]a', b'[$, alors $x(x')$ est un changement de variable régulier d'ordre p entre

$$\left]\lim_{x' \to a'^+} x(x'), \lim_{x' \to b'^-} x(x')\right[\quad \left(\text{resp.} \quad \left]\lim_{x' \to b'^-} x(x'), \lim_{x' \to a'^+} x(x')\right[\right)$$

et $]a', b'[$ si $x(x')$ est strictement croissant (resp. strictement décroissant). ($* \to$ Au paragraphe suivant, nous allons établir que le jacobien d'un changement de variable ne s'annule pas; il s'ensuit qu'on obtient ainsi tous les changements de variable réguliers d'ordre p entre deux intervalles ouverts de $\mathbb{R}$. $\leftarrow *$)

7.3.6 Matrices jacobiennes et jacobiens

Définitions. Soit $x(x')$ un changement de variable régulier d'ordre p entre les ouverts Ω et Ω' de $\mathbb{R}^n$, de changement de variable inverse noté $x'(x)$.

Les *matrices jacobiennes* qui lui sont associées sont notées

$$\frac{\partial(x)}{\partial(x')} = \frac{\partial(x_1, \ldots, x_n)}{\partial(x'_1, \ldots, x'_n)} \quad \text{et} \quad \frac{\partial(x')}{\partial(x)} = \frac{\partial(x'_1, \ldots, x'_n)}{\partial(x_1, \ldots, x_n)};$$

leurs éléments sont donnés par

$$\left(\frac{\partial(x)}{\partial(x')}\right)_{j,k} = \mathrm{D}_{x'_k} x_j(x') \quad \text{et} \quad \left(\frac{\partial(x')}{\partial(x)}\right)_{j,k} = \mathrm{D}_{x_k} x'_j(x)$$

avec $j,\, k = 1,\, \ldots,\, n$.

Les *jacobiens* qui lui sont associés sont donnés par

$$\mathrm{dtm}\left(\frac{\partial(x)}{\partial(x')}\right) \quad \text{et} \quad \mathrm{dtm}\left(\frac{\partial(x')}{\partial(x)}\right).$$

Il s'agit donc respectivement d'éléments de $\mathrm{C}_{p-1}(\Omega')$ et $\mathrm{C}_{p-1}(\Omega)$.

Ces matrices jacobiennes et jacobiens jouissent des propriétés suivantes.

Proposition 7.3.6.1 *Les matrices jacobiennes du changement de variable identité sur l'ouvert* Ω *de* $\mathbb{R}^n$ *coïncident avec la matrice identité sur* Ω, *à savoir la matrice* $\mathrm{diag}(\chi_\Omega, \ldots, \chi_\Omega)$, *et leurs jacobiens avec la fonction* χ_Ω. ∎

Proposition 7.3.6.2 *Si $x(x')$ et $x'(x'')$ sont des changements de variable réguliers d'ordre $p \geq 1$ entre les ouverts Ω et Ω' d'une part, et Ω' et Ω'' d'autre part, leur composition $x(x'(x'')) = x(x'')$ est un changement de variable régulier d'ordre p entre Ω et Ω'', de matrice jacobienne donnée par*

$$\left(\frac{\partial(x)}{\partial(x'')}\right) = \left(\frac{\partial(x)}{\partial(x')}\right)_{x'(x'')} \left(\frac{\partial(x')}{\partial(x'')}\right).$$

Preuve. Nous savons déjà que $x(x'')$ est un changement de variable régulier d'ordre p entre Ω et Ω''. La description de la matrice jacobienne résulte aussitôt du théorème de dérivation des fonctions composées. De fait, il vient

$$\mathrm{D}_k x_j(x'(x'')) = \sum_{l=1}^{n} [\mathrm{D}_l x_j(x')]_{x'(x'')} \cdot \mathrm{D}_k x'_l(x''). \blacksquare$$

Proposition 7.3.6.3 *Si $x(x')$ est un changement de variable régulier d'ordre $p \geq 1$ entre les ouverts Ω et Ω' de $\mathbb{R}^n$, on a*

$$\frac{\partial(x)}{\partial(x')} = \left(\frac{\partial(x')}{\partial(x)}\right)^{-1}_{x(x')} \quad et \quad \frac{\partial(x')}{\partial(x)} = \left(\frac{\partial(x)}{\partial(x')}\right)^{-1}_{x'(x)}.$$

De plus, les jacobiens sont liés par les relations

$$\mathrm{dtm}\left(\frac{\partial(x)}{\partial(x')}\right) \cdot \mathrm{dtm}\left(\left(\frac{\partial(x')}{\partial(x)}\right)_{x(x')}\right) = \chi_{\Omega'}$$

et

$$\mathrm{dtm}\left(\frac{\partial(x')}{\partial(x)}\right) \cdot \mathrm{dtm}\left(\left(\frac{\partial(x)}{\partial(x')}\right)_{x'(x)}\right) = \chi_{\Omega}.$$

En particulier, les jacobiens ne s'annulent pas.

Preuve. Cela résulte aussitôt des deux propositions précédentes appliquées à la composition des changements de variable $x(x')$ et $x'(x)$ où $x'(x)$ désigne le changement de variable inverse de $x(x')$. En effet, $x(x'(x))$ est alors le changement de variable identité sur Ω et $x'(x(x'))$, le changement de variable identité sur Ω'. $\blacksquare$

7.3.7 Position du problème

Soit $x(x')$ un changement de variable régulier d'ordre $p \geq 1$ entre les ouverts Ω et Ω' de $\mathbb{R}^n$. On sait alors, par le théorème de dérivation multiple des fonctions composées, que la fonction $f(x(x'))$ appartient à $\mathrm{C}_p(\Omega')$ pour tout $f \in \mathrm{C}_p(\Omega)$.

Soit en outre un opérateur de dérivation $F(x, \mathrm{D}_x)$ d'ordre $q \leq p$ dans Ω.

Le problème que nous allons résoudre est le suivant: existe-t-il un opérateur de dérivation $F'(x', \mathrm{D}_{x'})$ d'ordre q dans Ω' tel que

$$F'(x', \mathrm{D}_{x'})f(x(x')) = [F(x, \mathrm{D}_x)f(x)]_{x(x')}, \quad \forall f \in \mathrm{C}_p(\Omega), \quad (*)$$

ou, cela revient au même, tel que

$$[F'(x', \mathrm{D}_{x'})f(x(x'))]_{x'(x)} = F(x, \mathrm{D}_x)f(x), \quad \forall f \in \mathrm{C}_p(\Omega). \quad (**)$$

Nous allons voir que la réponse est affirmative et nous allons même donner un procédé de calcul de $F'(x', \mathrm{D}_{x'})$.

Dans le langage des opérateurs, on écrit

$$F(x, \mathrm{D}_x) = F'(x', \mathrm{D}_{x'}),$$

étant entendu que la signification à donner à cette égalité n'est rien d'autre que $(*)$ ou $(**)$ suivant le cas.

Emettons enfin une remarque: traditionnellement, on exige que, dans la réponse, seules les fonctions $x_j(x')$ avec $j = 1, \ldots, n$, interviennent. Les formules que nous avons établies sur les matrices jacobiennes permettent d'ailleurs de revenir immédiatement aux fonctions $x'_j(x)$ avec $j = 1, \ldots, n$.

7.3.8 Formule fondamentale

Théorème 7.3.8.1 *Si $x(x')$ est un changement de variable régulier d'ordre $p \geq 1$ entre les ouverts Ω et Ω' de $\mathbb{R}^n$, on a*

$$\begin{pmatrix} \mathrm{D}_{x_1} \\ \vdots \\ \mathrm{D}_{x_n} \end{pmatrix} = \left(\frac{\partial(x)}{\partial(x')}\right)^{\sim,-1} \begin{pmatrix} \mathrm{D}_{x'_1} \\ \vdots \\ \mathrm{D}_{x'_n} \end{pmatrix} = \left(\frac{\partial(x')}{\partial(x)}\right)^{\sim}_{x(x')} \begin{pmatrix} \mathrm{D}_{x'_1} \\ \vdots \\ \mathrm{D}_{x'_n} \end{pmatrix}.$$

Preuve. Si f appartient à $\mathrm{C}_1(\Omega)$, définissons la fonction f' sur Ω' par la relation $f'(x') = f(x(x'))$ pour tout $x' \in \Omega'$. Bien sûr, f' appartient à $\mathrm{C}_1(\Omega')$ et on a $f'(x'(x)) = f(x)$ pour tout $x \in \Omega$. Cela étant, il suffit de noter que le théorème de dérivation des fonctions composées donne

$$\mathrm{D}_{x_k} f(x) = \mathrm{D}_{x_k} f'(x'(x)) = \sum_{j=1}^{n} (\mathrm{D}_{x'_j} f') \circ x'(x) \cdot \mathrm{D}_{x_k} x'_j(x)$$

pour tout $k \leq n$, c'est-à-dire

$$\begin{pmatrix} \mathrm{D}_{x_1} \\ \vdots \\ \mathrm{D}_{x_n} \end{pmatrix} f(x) = \left(\frac{\partial(x')}{\partial(x)}\right)^{\sim} \begin{pmatrix} \mathrm{D}_{x'_1} f' \\ \vdots \\ \mathrm{D}_{x'_n} f' \end{pmatrix}_{x'(x)} = \left[\left(\frac{\partial(x)}{\partial(x')}\right)^{\sim,-1} \begin{pmatrix} \mathrm{D}_{x'_1} f' \\ \vdots \\ \mathrm{D}_{x'_n} f' \end{pmatrix}\right]_{x'(x)},$$

sous forme matricielle.

D'où la conclusion.∎

7.3.9 Formule générale

Théorème 7.3.9.1 *Si $x(x')$ est un changement de variable régulier d'ordre $p \geq 1$ entre les ouverts Ω et Ω' de $\mathbb{R}^n$ et si $F(x, \mathrm{D}_x)$ est un opérateur de dérivation d'ordre $q \leq p$ sur Ω, on a*

$$F(x, \begin{pmatrix} \mathrm{D}_{x_1} \\ \vdots \\ \mathrm{D}_{x_n} \end{pmatrix}) = F(x(x'), \left(\frac{\partial(x)}{\partial(x')}\right)^{\sim,-1} \begin{pmatrix} \mathrm{D}_{x'_1} \\ \vdots \\ \mathrm{D}_{x'_n} \end{pmatrix}).$$

Preuve. Cela résulte de suite de la définition même des dérivées multiples d'une fonction.∎

7.3.10 Cas où la matrice jacobienne est à colonnes orthogonales

Soit $x(x')$ est un changement de variable régulier d'ordre $p \geq 1$ entre les ouverts Ω et Ω' de $\mathbb{R}^n$ tel que la matrice jacobienne $\partial(x)/\partial(x')$ soit à colonnes orthogonales.

En désignant par $h_1(x')$, ..., $h_n(x')$ les carrés des modules des colonnes successives de $\partial(x)/\partial(x')$, il vient

$$\left(\frac{\partial(x)}{\partial(x')}\right)^{\sim} \left(\frac{\partial(x)}{\partial(x')}\right) = \mathrm{diag}(h_1(x'), \ldots, h_n(x'))$$

et

$$\left|\mathrm{dtm}\left(\frac{\partial(x)}{\partial(x')}\right)\right| = \sqrt{h_1(x') \ldots h_n(x')}.$$

Cela étant, la formule fondamentale et, par conséquent, la formule générale du changement de variable dans les opérateurs de dérivation admettent une grande simplification.

Théorème 7.3.10.1 *Si $x(x')$ est un changement de variable régulier d'ordre $p \geq 1$ entre les ouverts Ω et Ω' de $\mathbb{R}^n$ et si la matrice $\partial(x)/\partial(x')$ est à colonnes orthogonales, on a*

$$\begin{pmatrix} \mathrm{D}_{x_1} \\ \vdots \\ \mathrm{D}_{x_n} \end{pmatrix} = \frac{\partial(x)}{\partial(x')} \begin{pmatrix} \frac{1}{h_1(x')}\mathrm{D}_{x'_1} \\ \vdots \\ \frac{1}{h_n(x')}\mathrm{D}_{x'_n} \end{pmatrix}.$$

Preuve. Il suffit de noter qu'on a successivement

$$\left(\frac{\partial(x)}{\partial(x')}\right)^{\sim,-1} = \left(\frac{\partial(x)}{\partial(x')}\right)\left(\frac{\partial(x)}{\partial(x')}\right)^{-1}\left(\frac{\partial(x)}{\partial(x')}\right)^{\sim,-1} = \frac{\partial(x)}{\partial(x')}\left(\left(\frac{\partial(x)}{\partial(x')}\right)^{\sim}\frac{\partial(x)}{\partial(x')}\right)^{-1}$$

donc

$$\left(\frac{\partial(x)}{\partial(x')}\right)^{\sim,-1} = \frac{\partial(x)}{\partial(x')}\mathrm{diag}\left(\frac{1}{h_1(x')},\dots,\frac{1}{h_n(x')}\right).\blacksquare$$

7.3.11 Exemples

Changement de variable dans $\mathbb{R}$

Soit $x(x')$ un changement de variable régulier d'ordre $p \geq 1$ entre les ouverts Ω et Ω' de $\mathbb{R}$. Soit en outre $F(x, \mathrm{D})$ un opérateur de dérivation d'ordre $q \leq p$ sur Ω. Il vient alors

$$F(x, \mathrm{D}_x) = F\left(x(x'), \frac{1}{\mathrm{D}_{x'}x(x')}\mathrm{D}_{x'}\right).$$

Par exemple, on vérifie de suite que $x(x') = \mathrm{e}^{x'}$ est un changement de variable régulier d'ordre infini entre $\Omega =]0, +\infty[$ et $\Omega' = \mathbb{R}$, tel que

$$\mathrm{D}_x = \frac{1}{\mathrm{D}_{x'}\mathrm{e}^{x'}}\mathrm{D}_{x'} = \mathrm{e}^{-x'}\mathrm{D}_{x'}$$

donc tel que $x\mathrm{D}_x = \mathrm{D}_{x'}$.

Un changement de variable adéquat peut donc fortement simplifier l'expression d'un opérateur de dérivation; nous allons en voir de nombreux exemples par la suite. ($* \to$ L'exemple que nous venons d'envisager est fort important: il permet de résoudre les équations différentielles d'Euler, cf. § 9.1.7. $\leftarrow *$)

Changement de variable linéaire

Soient A une matrice réelle de dimension $n \times n$ et a un point de $\mathbb{R}^n$. On vérifie de suite que $x(x') = Ax' + a$ est un changement de variable régulier d'ordre infini entre $\mathbb{R}^n$ et $\mathbb{R}^n$ si et seulement si la matrice A est non singulière, auquel cas le changement de variable inverse est donné par $x'(x) = A^{-1}(x - a)$ et les matrices jacobiennes par

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

Dès lors, la formule fondamentale s'écrit

$$\begin{pmatrix} \mathrm{D}_{x_1} \\ \vdots \\ \mathrm{D}_{x_n} \end{pmatrix} = \tilde{A}^{,-1}\begin{pmatrix} \mathrm{D}_{x'_1} \\ \vdots \\ \mathrm{D}_{x'_n} \end{pmatrix}.$$

En particulier, si A s'écrit

$$A = \begin{pmatrix} a_1 & b_1 & \cdots & n_1 \\ a_2 & b_2 & \cdots & n_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_n & b_n & \cdots & n_n \end{pmatrix}$$

avec $\mathrm{dtm}(A) \neq 0$, le changement de variable $x = Ax' + a$ transforme l'opérateur de dérivation

$$\left(\sum_{j=1}^{n} a_j \mathrm{D}_{x_j}\right)\left(\sum_{k=1}^{n} b_k \mathrm{D}_{x_k}\right)\dots\left(\sum_{l=1}^{n} n_l \mathrm{D}_{x_l}\right) \quad \text{en} \quad \mathrm{D}_{x'_1}\mathrm{D}_{x'_2}\dots\mathrm{D}_{x'_n}.$$

De fait, la formule fondamentale s'écrit aussi

$$\widetilde{A}\begin{pmatrix} \mathrm{D}_{x_1} \\ \vdots \\ \mathrm{D}_{x_n} \end{pmatrix} = \begin{pmatrix} \mathrm{D}_{x'_1} \\ \vdots \\ \mathrm{D}_{x'_n} \end{pmatrix}.$$

Ainsi, pour a, $b \in \mathbb{R} \setminus \{0\}$, le changement de variable linéaire

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

transforme l'opérateur de dérivation $a^2\mathrm{D}_x^2 - b^2\mathrm{D}_y^2$ en $\mathrm{D}_{x'}\mathrm{D}_{y'}$. (* Il permet en fait une étude très aisée de l'opérateur des ondes dans $\mathbb{R}$.*)

Opérateur des ondes et transformation de Galilée

Dans $\mathbb{R}^n$, le *laplacien* est l'opérateur de dérivation d'ordre deux Δ défini par

$$\Delta = \sum_{k=1}^{n} \mathrm{D}_k^2 = \mathrm{D}_1^2 + \ldots + \mathrm{D}_n^2.$$

L'*opérateur des ondes* est alors l'opérateur de dérivation

$$\Delta - c^{-2}\mathrm{D}_t^2$$

où c est la vitesse de propagation et où t est une variable indépendante de $x_1, \ldots,$ x_n (le temps). Dans $\mathbb{R}^3$, c'est l'opérateur qui régit la propagation de la lumière dans un milieu homogène.

Afin de simplifier les écritures, nous allons voir comment se transforme cet opérateur de dérivation sur $\mathbb{R}$ au moyen de la *transformation de Galilée,* c'est-à-dire du changement de variable linéaire

$$\begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} 1 & v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix}.$$

Il vient

$$\begin{pmatrix} \mathrm{D}_x \\ \mathrm{D}_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -v & 1 \end{pmatrix} \begin{pmatrix} \mathrm{D}_{x'} \\ \mathrm{D}_{t'} \end{pmatrix}$$

et, après calculs, on trouve

$$\mathrm{D}_x^2 - c^{-2}\mathrm{D}_t^2 = (1 - c^{-2}v^2)\mathrm{D}_{x'}^2 + 2c^{-2}v\mathrm{D}_{x'}\mathrm{D}_{t'} - c^{-2}\mathrm{D}_{t'}^2.$$

On en déduit que l'opérateur des ondes n'est pas invariant pour la transformation de Galilée.

Opérateur des ondes et transformation de Lorentz

Dans $\mathbb{R}$, la *transformation de Lorentz* est le changement de variable linéaire

$$\begin{pmatrix} x \\ t \end{pmatrix} = \frac{1}{\sqrt{1 - v^2/c^2}} \begin{pmatrix} 1 & v \\ v/c^2 & 1 \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix}.$$

Après calculs, on obtient

$$\mathrm{D}_x^2 - c^{-2}\mathrm{D}_t^2 = \mathrm{D}_{x'}^2 - c^{-2}\mathrm{D}_{t'}^2$$

et l'opérateur des ondes est invariant pour la transformation de Lorentz.

Un changement de variable non linéaire

Que devient l'opérateur de dérivation $\sum_{k=1}^n x_k \mathrm{D}_k$ *au moyen du changement de variable* $x_k = x'_1 \dots x'_k, (k = 1, \dots, n)$*?*

Etudions d'abord ce changement de variable. Il s'inverse bien sûr en $x'_1 = x_1$ et $x'_k = x_k/x_{k-1}$ pour $k = 2, \dots, n$. Les fonctions $x_j(x')$ appartiennent visiblement à $\mathrm{C}_\infty(\mathbb{R}^n)$ et les fonctions $x'_j(x)$ à

$$\mathrm{C}_\infty(\mathbb{R}^n \setminus \{\, x : x_1 = 0 \text{ ou } \dots \text{ ou } x_{n-1} = 0\}).$$

On vérifie alors de suite par exemple qu'il s'agit d'un changement de variable régulier d'ordre infini entre les ouverts $\Omega = \{\, x : x_1, \dots, x_n > 0\}$ et $\Omega' = \{\, x' : x'_1, \dots, x'_n > 0\}$ de $\mathbb{R}^n$.

Etudions ensuite ce que devient l'opérateur de dérivation. Afin d'éviter le calcul de l'inverse d'une matrice $n \times n$, notons que

$$\frac{\partial(x')}{\partial(x)} = \begin{pmatrix} 1 & & & & & \\ -x_2/x_1^2 & 1/x_1 & & & & \\ & \cdot & \cdot & & & \\ & & \cdot & \cdot & & \\ & & & \cdot & \cdot & \\ & & & & -x_n/x_{n-1}^2 & 1/x_{n-1} \end{pmatrix}$$

donc

$$\begin{pmatrix} \mathrm{D}_{x_1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \mathrm{D}_{x_n} \end{pmatrix} = \begin{pmatrix} 1 & -x_2/x_1^2 & & & \\ & 1/x_1 & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & \cdot \\ & & & \cdot & -x_n/x_{n-1}^2 \\ & & & & 1/x_{n-1} \end{pmatrix}_{x(x')} \cdot \begin{pmatrix} \mathrm{D}_{x'_1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \mathrm{D}_{x'_n} \end{pmatrix}$$

D'où on tire

$$\begin{aligned} &\sum_{k=1}^{n} x_k \mathrm{D}_{x_k} \\ &= x'_1(\mathrm{D}_{x'_1} - \frac{x'_2}{x'_1}\mathrm{D}_{x'_2}) + x'_1x'_2(\frac{1}{x'_1}\mathrm{D}_{x'_2} - \frac{x'_3}{x'_1x'_2}\mathrm{D}_{x'_3}) + \ldots + \frac{x'_1 \ldots x'_n}{x'_1 \ldots x'_{n-1}}\mathrm{D}_{x'_n} \\ &= x'_1\mathrm{D}_{x'_1}. \end{aligned}$$

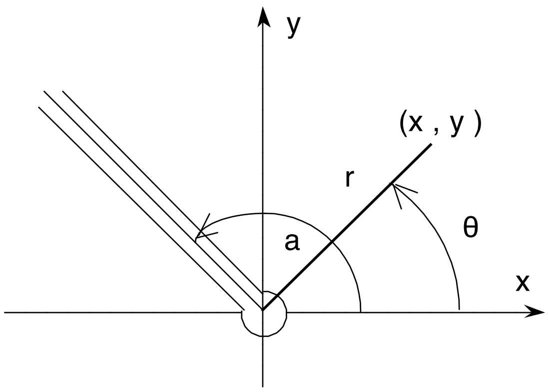
Laplacien et transformation polaire dans $\mathbb{R}^2$

Etudions tout d'abord la *transformation* ou *changement de variable polaire* dans $\mathbb{R}^2$. Il est fourni par les relations

$$\begin{cases} x = r\cos(\theta) \\ y = r\sin(\theta) \end{cases} \quad \text{avec} \quad \begin{cases} r \in\,]0, +\infty[\\ \theta \in\,]a, a+2\pi[\end{cases}$$

avec $a \in \mathbb{R}$ (souvent on choisit $a = 0$ ou $a = \pi$).

Son interprétation géométrique est immédiate.



Il s'inverse en

$$\left\{\begin{array}{l} r = \sqrt{x^2+y^2} \\ \left\{\begin{array}{l} \cos(\theta) = x/r \\ \sin(\theta) = y/r \end{array}\right. \end{array}\right.$$

et nous savons que, pour $r = (x^2+y^2)^{1/2} > 0$, le système $\cos(\theta) = x/r$ et $\sin(\theta) = y/r$ admet une solution unique $\theta(x,y) \in [a, a+2\pi[$ telle que $\theta \in \mathrm{C}_\infty(\Omega_a)$. Il s'agit donc d'un changement de variable régulier d'ordre infini entre

$$\Omega_a = \mathbb{R}^2 \setminus \{\, (r\cos(a), r\sin(a)) : r \geq 0\} \quad \text{et} \quad \Omega'_a =]0, +\infty[\, \times \,]a, a+2\pi[\,.$$

Dans ce cas, la matrice

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}$$

est à colonnes orthogonales et on a $h_1 = 1$ et $h_2 = r^2$. La formule fondamentale s'écrit donc

$$\begin{pmatrix} \mathrm{D}_x \\ \mathrm{D}_y \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} \begin{pmatrix} \mathrm{D}_r \\ r^{-2}\mathrm{D}_\theta \end{pmatrix}.$$

Cela étant, le laplacien devient

$$\begin{aligned} \Delta &= \mathrm{D}_x^2 + \mathrm{D}_y^2 \\ &= \left(\cos(\theta)\mathrm{D}_r - \frac{\sin(\theta)}{r}\mathrm{D}_\theta\right)\left(\cos(\theta)\mathrm{D}_r - \frac{\sin(\theta)}{r}\mathrm{D}_\theta\right) \\ &\quad + \left(\sin(\theta)\mathrm{D}_r + \frac{\cos(\theta)}{r}\mathrm{D}_\theta\right)\left(\sin(\theta)\mathrm{D}_r + \frac{\cos(\theta)}{r}\mathrm{D}_\theta\right) \\ &= \mathrm{D}_r^2 + \frac{1}{r}\mathrm{D}_r + \frac{1}{r^2}\mathrm{D}_\theta^2. \end{aligned}$$

Remarque. Cette dernière formule permet un calcul aisé du laplacien des fonctions radiales appartenant à $C_2(\mathbb{R}^2 \setminus \{0\})$ et montre qu'il s'agit encore d'une fonction radiale. En particulier, on remarquera que la fonction $\ln((x^2+y^2)^{1/2})$ appartient à $C_\infty(\mathbb{R}^2 \setminus \{0\})$ et que son laplacien est la fonction 0 sur $\mathbb{R}^2 \setminus \{0\}$. $* \rightarrow$ On dit que $\ln((x^2+y^2)^{1/2})$ est la solution élémentaire du laplacien sur $\mathbb{R}^2 \setminus \{0\}$. $\leftarrow *$

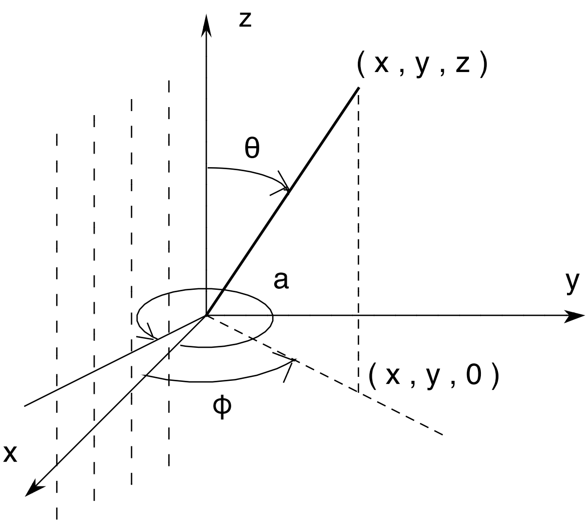
Laplacien et transformation polaire dans $\mathbb{R}^3$

Etudions tout d'abord la *transformation* ou *changement de variable polaire dans* $\mathbb{R}^3$. Il est fourni par les relations

$$\left\{\begin{array}{l} x = r\sin(\theta)\cos(\varphi) \\ y = r\sin(\theta)\sin(\varphi) \\ z = r\cos(\theta) \end{array}\right. \quad \text{avec} \quad \left\{\begin{array}{l} r \in \,]0,+\infty[\\ \varphi \in \,]a, a+2\pi[\\ \theta \in \,]0,\pi[\end{array}\right.$$

avec $a \in \mathbb{R}$ (souvent on choisit $a = 0$).

Son interprétation géométrique est immédiate.



Il s'inverse en

$$\left\{\begin{array}{l} r = \sqrt{x^2+y^2+z^2} \\ \theta = \arccos(z/r) \\ \left\{\begin{array}{l} \cos(\varphi) = \frac{x}{\sqrt{x^2+y^2}} \\ \sin(\varphi) = \frac{y}{\sqrt{x^2+y^2}} \end{array}\right\} \end{array}\right. \quad \begin{array}{l}\text{ce système admet une solution unique} \\ \text{dans } [a, a+2\pi[\text{ telle que } \varphi \in C_\infty(\Omega_a).\end{array}$$

Il s'agit donc d'un changement de variable régulier d'ordre infini entre

$$\Omega_a = \mathbb{R}^3 \setminus \{\, (r\cos(a), r\sin(a), z) : r \in \,]0,+\infty[\,, z \in \mathbb{R}\}$$

et

$$\Omega'_a =]0, +\infty[\times]0, \pi[\times]a, a + 2\pi[\,.$$

Dans ce cas, la matrice

$$\frac{\partial(x,y,z)}{\partial(r,\theta,\varphi)} = \begin{pmatrix} \sin(\theta)\cos(\varphi) & r\cos(\theta)\cos(\varphi) & -r\sin(\theta)\sin(\varphi) \\ \sin(\theta)\sin(\varphi) & r\cos(\theta)\sin(\varphi) & r\sin(\theta)\cos(\varphi) \\ \cos(\theta) & -r\sin(\theta) & 0 \end{pmatrix}$$

est à colonnes orthogonales et on a $h_1 = 1$, $h_2 = r^2$ et $h_3 = r^2 \sin^2(\theta)$. La formule fondamentale s'écrit donc

$$\begin{pmatrix} \mathrm{D}_x \\ \mathrm{D}_y \\ \mathrm{D}_z \end{pmatrix} = \frac{\partial(x,y,z)}{\partial(r,\theta,\varphi)} \begin{pmatrix} \mathrm{D}_r \\ r^{-2}\mathrm{D}_\theta \\ r^{-2}\sin^{-2}(\theta)\mathrm{D}_\varphi \end{pmatrix}.$$

Après de laborieux calculs, on obtient

$$\Delta = \mathrm{D}_x^2 + \mathrm{D}_y^2 + \mathrm{D}_z^2 = \mathrm{D}_r^2 + \frac{2}{r}\mathrm{D}_r + \frac{1}{r^2}\left(\mathrm{D}_\theta^2 + \mathrm{cotg}(\theta)\mathrm{D}_\theta\right) + \frac{1}{r^2\sin^2(\theta)}\mathrm{D}_\varphi^2.$$

Cependant, on peut obtenir cette formule bien plus aisément au moyen des deux changements de variable (polaires dans $\mathbb{R}^2$) successifs suivants:

$$\begin{cases} x = \rho\cos(\varphi) \\ y = \rho\sin(\varphi) \\ z = z \end{cases} \quad \text{et} \quad \begin{cases} z = r\cos(\theta) \\ \rho = r\sin(\theta) \\ \varphi = \varphi \end{cases}$$

Le laplacien devient successivement

$$\Delta = \mathrm{D}_\rho^2 + \frac{1}{\rho}\mathrm{D}_\rho + \frac{1}{\rho^2}\mathrm{D}_\varphi^2 + \mathrm{D}_z^2$$

puis

$$\Delta = \mathrm{D}_r^2 + \frac{1}{r}\mathrm{D}_r + \frac{1}{r^2}\mathrm{D}_\theta^2 + \frac{1}{r\sin(\theta)}\left(\sin(\theta)\mathrm{D}_r + \frac{\cos(\theta)}{r}\mathrm{D}_\theta\right) + \frac{1}{r^2\sin^2(\theta)}\mathrm{D}_\varphi^2.$$

Remarque. Cette dernière formule permet un calcul aisé du laplacien des fonctions radiales appartenant à $\mathrm{C}_2(\mathbb{R}^3 \setminus \{0\})$ et montre qu'il s'agit encore d'une fonction radiale. En particulier, on remarquera que la fonction $(x^2+y^2+z^2)^{-1/2}$ appartient à $\mathrm{C}_\infty(\mathbb{R}^3 \setminus \{0\})$ et que son laplacien est la fonction 0 sur $\mathbb{R}^3 \setminus \{0\}$. ($* \rightarrow$ On dit que $(x^2+y^2+z^2)^{-1/2}$ est la solution élémentaire du laplacien sur $\mathbb{R}^3 \setminus \{0\}$. $\leftarrow *$)

Chapitre 8

Primitivation

8.1 Primitivation dans $\mathbb{R}$

8.1.1 Notation $f \simeq g$

Notation. Soit A une partie de $\mathbb{R}$ et soient f, g deux fonctions définies sur A. On écrit

$$f \simeq_A g \quad \text{ou} \quad f \simeq g \text{ sur } A$$

s'il existe $c \in \mathbb{C}$ tel que $f = g + c$ sur A. Ces notations se lisent *"f est égal à g sur A à une constante additive près"*. On écrit aussi plus simplement $f \simeq g$ si aucune confusion sur A n'est possible.

La relation $\simeq_A$ est bien sûr une relation d'équivalence sur l'ensemble $F(A)$ des fonctions définies sur A car, pour tous f, g, $h \in F(A)$, on a évidemment

$$\begin{gathered} f \simeq f \\ f \simeq g \Rightarrow g \simeq f \\ (f \simeq g, g \simeq h) \Rightarrow f \simeq h. \end{gathered}$$

Voici deux critères qui assurent que $f \simeq g$. Le premier est basé sur la notion de variation d'une fonction.

Définition. Soit f une fonction définie sur $A \subset \mathbb{R}$. Si a et b appartiennent à A, alors la *variation de f entre a et b* est le nombre $f(b) - f(a)$; elle est notée $[f]_a^b$.

Critère 8.1.1.1 *Deux fonctions f, g définies sur $A \subset \mathbb{R}$ sont telles que $f \simeq g$ sur A si et seulement si $[f]_a^b = [g]_a^b$ pour tous a, $b \in A$.*

Preuve. La condition est évidemment nécessaire puisque

$$[f]_a^b = f(b) - f(a) = f(b) + c - f(a) - c = [f + c]_a^b$$

pour tous a, $b \in A$ et tout $c \in \mathbb{C}$.

La condition est suffisante. Etant donné un point x_0 de A, il vient

$$\begin{aligned} f(x) &= [f]_{x_0}^x + f(x_0) = [g]_{x_0}^x + f(x_0) \\ &= g(x) + (f(x_0) - g(x_0)) \end{aligned}$$

pour tout $x \in A$; on a donc $f = g + c$ sur A avec $c = f(x_0) - g(x_0)$.∎

Critère 8.1.1.2 *Si f et g sont des fonctions dérivables sur $]a, b[\subset \mathbb{R}$, alors on a $f \simeq g$ sur $]a, b[$ si et seulement si $\mathrm{D}f$ est égal à $\mathrm{D}g$ sur $]a, b[$.*

Preuve. C'est un cas particulier du théorème de l'ouvert connexe.∎

8.1.2 Primitive d'une fonction

Définition. Soit f une fonction définie sur l'intervalle $]a, b[$ de $\mathbb{R}$. On appelle alors *primitive de f sur $]a, b[$* ou même plus simplement, *primitive de f* si aucune confusion sur $]a, b[$ n'est possible, toute fonction F dérivable sur $]a, b[$ et telle que $\mathrm{D}F = f$ sur $]a, b[$. Si une telle primitive F existe, on dit que f est *primitivable sur* $]a, b[$.

Bien sûr, si F est une primitive sur $]a, b[$ de $f \in \mathrm{C}_p(]a, b[)$, alors on a $F \in \mathrm{C}_{p+1}(]a, b[)$.

Remarque. Il existe des fonctions qui ne sont pas primitivables. Ainsi $\chi_{]0,+\infty[}$ n'est pas primitivable sur $\mathbb{R}$. En effet, une primitive F de $\chi_{]0,+\infty[}$ sur $\mathbb{R}$ serait telle que $\mathrm{D}F(x) = 0$ en tout point x de l'ouvert connexe $]-\infty, 0[$ et F devrait être constant sur $]-\infty, 0[$. De même, on aurait $\mathrm{D}F(x) = 1$ en tout point x de l'ouvert connexe $]0, +\infty[$ et F devrait être égal à $x + c$ sur $]0, +\infty[$ avec $c \in \mathbb{C}$. On vérifie alors de suite qu'une telle fonction ne peut être dérivable en 0.

Vis-à-vis des primitives d'une même fonction, on a les résultats suivants.

Proposition 8.1.2.1 *Si F_1 et F_2 sont des primitives sur $]a, b[\subset \mathbb{R}$ d'une même fonction f, on a $F_1 \simeq F_2$ sur $]a, b[$ et, pour tout $c \in \mathbb{C}$, $F_1 + c$ est une primitive de f sur $]a, b[$.*

Preuve. La première partie est un cas particulier du théorème de l'ouvert connexe, la seconde est triviale.∎

Remarques. a) Il s'ensuit que, si f admet une primitive sur $]a, b[$, il en admet une infinité. Cependant *si f admet une primitive sur $]a, b[$, il existe une primitive et une seule de f sur $]a, b[$ qui, en $x_0 \in]a, b[$, prend une valeur fixée.*

b) Il n'y a pas unicité des primitives d'une fonction f sur $]a, b[$ et la seule notion d'"égalité" entre primitives est l'égalité à une constante additive près.

Notation. Si la fonction f admet une primitive F sur $]a, b[$,

$$\int f(x)\,dx$$

désigne une quelconque primitive de f sur $]a, b[$; on a donc $F(x) \simeq \int f(x)\,dx$. Cela étant, $\int f(x)\,dx$ est essentiellement une fonction dérivable sur $]a, b[$ et de dérivée égale à f.

Proposition 8.1.2.2 *Si f est une fonction réelle et primitivable sur $]a, b[$, il existe une primitive de f qui est une fonction réelle sur $]a, b[$.*

Preuve. De fait, si F est une primitive de f sur $]a, b[$, alors $\Re F$ est une fonction réelle, dérivable et telle que $\mathrm{D}\Re F = f$ sur $]a, b[$.∎

Vis-à-vis de l'existence d'une primitive d'une fonction f, on a le théorème fondamental suivant.

Théorème 8.1.2.3 *Toute fonction $f \in \mathrm{C}_0(]a, b[)$ est primitivable sur $]a, b[$.*∎

Nous allons admettre ce résultat sans démonstration car une preuve ici serait longue et artificielle, alors qu'il est possible d'établir aisément cette propriété en recourant au calcul intégral. Cette position s'admet d'autant mieux que, dans l'exposé oral, nous n'utiliserons pas ce théorème avant le cours de calcul intégral.

Insistons cependant sur l'importance de ce résultat, au moyen des conséquences que voici. La fonction

$$\frac{1}{x} \quad (\text{resp. } \frac{1}{\sqrt{1-x^2}};\ \frac{1}{1+x^2})$$

est réelle et appartient à $\mathrm{C}_0(]0, +\infty[)$ (resp. $\mathrm{C}_0(]-1, 1[)$; $\mathrm{C}_0(\mathbb{R})$) et admet donc une primitive réelle unique sur $]0, +\infty[$ (resp. $]-1, 1[$; $\mathbb{R}$) qui en $x = 1$ (resp. 0; 0), prend la valeur 0. On retrouve ainsi les fonctions

$$\ln \ (\text{resp. } \arcsin;\ \mathrm{arctg})\,.$$

Remarque. $* \rightarrow$ Il existe cependant des fonctions primitivables sur $]a, b[$ qui ne sont pas continues sur $]a, b[$. Il existe en effet des fonctions dérivables et non continûment dérivables sur $]a, b[$. $\leftarrow *$

8.1.3 Théorie du calcul des primitives

Voici quelques résultats relatifs aux primitives; ils constituent la base du calcul effectif des primitives d'une fonction définie sur un intervalle $]a,b[$ de $\mathbb{R}$.

Bien sûr, par définition, on a déjà le résultat suivant: *si f est une fonction primitivable sur $]a,b[$, $\int f(x)dx$ existe et est une fonction dérivable sur $]a,b[$ telle que*

$$\mathrm{D}\int f(x)\,dx = f(x), \quad \forall x \in \,]a,b[\,.$$

Proposition 8.1.3.1 *Si F est une fonction dérivable sur $]a,b[$, $\mathrm{D}F$ est primitivable sur $]a,b[$ et*

$$\int \mathrm{D}F(x)\,dx \simeq F(x).\blacksquare$$

Proposition 8.1.3.2 (combinaison linéaire) *Si $f_1, \ldots, f_J$ sont des fonctions primitivables sur $]a,b[$ en nombre fini, alors toutes leurs combinaisons linéaires sont primitivables sur $]a,b[$ et*

$$\int \sum_{j=1}^{J} c_j f_j(x)\,dx \simeq \sum_{j=1}^{J} c_j \int f_j(x)\,dx.$$

Preuve. C'est une conséquence directe du théorème relatif à la dérivation des combinaisons linéaires.$\blacksquare$

Proposition 8.1.3.3 (parties réelle, imaginaire) a) *La fonction f est primitivable sur $]a,b[$ si et seulement si $\overline{f}$ est primitivable sur $]a,b[$; de plus, on a alors*

$$\int \overline{f}(x)\,dx \simeq \left(\int f(x)\,dx\right)^{-}.$$

b) *Une fonction f est primitivable sur $]a,b[$ si et seulement si $\Re f$ et $\Im f$ sont primitivables sur $]a,b[$; de plus, on a alors*

$$\int f(x)\,dx \simeq \int \Re f(x)\,dx + i\int \Im f(x)\,dx.$$

Preuve. a) résulte aussitôt de la formule $\mathrm{D}\overline{f} = (\mathrm{D}f)^{-}$, valable si f est dérivable.

b) est conséquence directe de a) et de la proposition précédente.$\blacksquare$

Proposition 8.1.3.4 (par parties) *Si f et g sont des fonctions dérivables sur $]a,b[$ et si $f \cdot \mathrm{D}g$ est primitivable sur $]a,b[$, alors $g \cdot \mathrm{D}f$ est primitivable sur $]a,b[$ et*

$$\int g(x) \cdot \mathrm{D}f(x)\,dx \simeq f(x) \cdot g(x) - \int f(x) \cdot \mathrm{D}g(x)\,dx.$$

Preuve. De fait, $f(x) \cdot g(x) - \int f(x) \cdot \mathrm{D}g(x)\,dx$ est une fonction dérivable sur $]a,b[$, dont la dérivée est donnée par $g \cdot \mathrm{D}f + f \cdot \mathrm{D}g - f \cdot \mathrm{D}g = g \cdot \mathrm{D}f$.∎

Proposition 8.1.3.5 (par substitution) *Si f est une fonction primitivable sur $]a,b[$ et si $x(x')$ est une fonction dérivable sur $]a',b'[\subset \mathbb{R}$ dont l'ensemble de variation $\{\, x(x') : x' \in]a',b'[\}$ est inclus dans $]a,b[$, alors la fonction $f(x(x'))\cdot\mathrm{D}x(x')$ est primitivable sur $]a',b'[$ et telle que*

$$\int f(x(x')) \cdot \mathrm{D}x(x')\,dx' \simeq \left[\int f(x)\,dx\right]_{x(x')}.$$

Preuve. C'est une conséquence directe du théorème de dérivation des fonctions de fonction.∎

Proposition 8.1.3.6 (changement de variable) *Soit f une fonction sur l'intervalle $]a,b[$. Si $x(x')$ est un changement de variable régulier d'ordre $p \geq 1$ entre $]a,b[$ et $]a',b'[$ et si $f(x(x')) \cdot \mathrm{D}x(x')$ est primitivable sur $]a',b'[$, alors f est primitivable sur $]a,b[$ et tel que*

$$\int f(x)\,dx = \left[\int f[x(x')] \cdot \mathrm{D}x(x')\,dx'\right]_{x'(x)}.$$

Preuve. C'est une conséquence du théorème de dérivation des fonctions de fonction. De fait,

$$\left[\int f(x(x')) \cdot \mathrm{D}x(x')\,dx'\right]_{x'(x)}$$

est une fonction dérivable sur $]a,b[$ dont la dérivée est donnée par

$$[f(x(x')) \cdot \mathrm{D}x(x')]_{x'(x)} \cdot \mathrm{D}x'(x),$$

c'est-à-dire par $f(x)$, vu la propriété fondamentale des jacobiens relatifs à un changement de variable.∎

8.1.4 Généralités sur le calcul des primitives

Si f est une fonction continue sur $]a, b[\subset \mathbb{R}$, qui s'exprime au moyen d'un nombre fini de fonctions élémentaires par le biais d'opérations algébriques ou de fonctions composées, nous savons que f est primitivable sur $]a, b[$. Le problème qui nous intéresse ici est celui de calculer effectivement une primitive de f au moyen d'un nombre fini de fonctions élémentaires.

En fait, on ne peut résoudre ce problème que dans certains cas particuliers que nous allons étudier. Ainsi on vérifie de suite que

$$\int x\mathrm{e}^{-x}\,dx \simeq -x\mathrm{e}^{-x} - \mathrm{e}^{-x} \quad \text{sur} \quad \mathbb{R},$$

mais par contre la primitive de $\mathrm{e}^x/x \in \mathrm{C}_0(]0, +\infty[)$ ne s'obtient pas au moyen d'un nombre fini de fonctions élémentaires.

En dehors de quelques cas de primitivation systématique que nous allons passer en revue dans les paragraphes suivants, on est amené à faire preuve d'ingéniosité ou à consulter des tables.

Le calcul des primitives est basé sur:

a) les règles de calcul des primitives que nous avons établies au paragraphe précédent;

b) la connaissance à priori de certaines primitives de référence. Nous en avons en fait déjà beaucoup à notre disposition; elles se déduisent directement des propriétés des fonctions élémentaires et de la propriété suivante: si F est une fonction dérivable sur $]a, b[$, alors $\mathrm{D}F$ est primitivable sur $]a, b[$ et on a $\int \mathrm{D}F(x)\,dx \simeq F(x)$. On en déduit aussitôt les formules contenues dans le tableau de la page 265.

Afin d'alléger systématiquement les notations dans la recherche des primitives et sauf mention explicite du contraire, nous allons utiliser les notations suivantes dans les paragraphes 8.1.5 à 8.1.9:

a) a, b, r désignent des nombres réels,

b) α, β, c désignent des nombres complexes,

c) P désigne un polynôme,

d) R désigne une fraction rationnelle.

8.1.5 Primitivation d'un polynôme

La primitive d'un polynôme s'obtient directement au moyen de la formule

$$\int x^m\,dx \simeq \frac{x^{m+1}}{m+1} \quad \text{sur} \quad \mathbb{R}, \quad \forall m \in \mathbb{N},$$

et de la propriété relative à la primitivation d'une combinaison linéaire de fonctions primitivables sur $\mathbb{R}$.

primitive	valable sur
$\int (x-a)^m\,dx \simeq \frac{(x-a)^{m+1}}{m+1}$	$\mathbb{R}$
$\int \frac{dx}{x-a} \simeq \ln(\vert x-a\vert)$	$]-\infty, a[$ ou $]a, +\infty[$
$\int \frac{dx}{(x-a)^p} \simeq \frac{1}{1-p}\frac{1}{(x-a)^{p-1}}$	$]-\infty, a[$ ou $]a, +\infty[$
$\int (x-\alpha)^n\,dx \simeq \frac{(x-\alpha)^{n+1}}{n+1}$	$\mathbb{R}$
$\int x^\beta\,dx \simeq \frac{x^{\beta+1}}{\beta+1}$	$]0, +\infty[$
$\int r^{\gamma x}\,dx \simeq \frac{r^{\gamma x}}{\gamma \ln(r)}$	$\mathbb{R}$
$\int \cos(x)\,dx \simeq \sin(x)$	$\mathbb{R}$
$\int \sin(x)\,dx \simeq -\cos(x)$	$\mathbb{R}$
$\int \frac{dx}{\cos^2(x)} \simeq \mathrm{tg}(x)$	$]-\pi/2, \pi/2[$
$\int \frac{dx}{\sin^2(x)} \simeq -\mathrm{cotg}(x)$	$]0, \pi[$
$\int \mathrm{ch}(x)\,dx \simeq \mathrm{sh}(x)$	$\mathbb{R}$
$\int \mathrm{sh}(x)\,dx \simeq \mathrm{ch}(x)$	$\mathbb{R}$
$\int \frac{dx}{\mathrm{ch}^2(x)} \simeq \mathrm{th}(x)$	$\mathbb{R}$
$\int \frac{dx}{\mathrm{sh}^2(x)} \simeq -\coth(x)$	$]-\infty, 0[$ ou $]0, +\infty[$
$\int \frac{dx}{\sqrt{1-x^2}} \simeq \arcsin(x)$	$]-1, 1[$
$\int \frac{dx}{1+x^2} \simeq \mathrm{arctg}(x)$	$\mathbb{R}$
$\int \frac{dx}{\sqrt{1+x^2}} \simeq \mathrm{arcsh}(x)$	$\mathbb{R}$
$\int \frac{dx}{\sqrt{x^2-1}} \simeq \mathrm{arcch}(x)$	$]1, +\infty[$
$\int \frac{dx}{1-x^2} \simeq \mathrm{arcth}(x)$	$]-1, 1[$
$\int \frac{dx}{1-x^2} \simeq \mathrm{arccoth}(x)$	$]1, +\infty[$

Conditions : $m \in \mathbb{N},\ p \in \mathbb{N}_0 \setminus \{1\},\ n \in \mathbb{Z} \setminus \{-1\},\ a \in \mathbb{R},\ r > 0,$
$\alpha \in \mathbb{C} \setminus \mathbb{R},\ \beta \in \mathbb{C} \setminus \mathbb{Z},\ \gamma \in \mathbb{C} \setminus \{0\}.$

Exemple. *On a bien sûr* $\int (5x^2 - 2x + 1)\, dx \simeq 5x^3/3 - x^2 + x$ *sur* $\mathbb{R}$.

8.1.6 Primitivation d'une exponentielle-polynôme

Définition. Une *fonction exponentielle-polynôme* est une fonction qui s'écrit sous la forme du produit $\mathrm{e}^{\alpha x} P(x)$ d'une exponentielle et d'un polynôme.

Procédé de calcul. Pour calculer une primitive de la fonction exponentielle-polynôme $\mathrm{e}^{\alpha x} P(x)$, on procède comme suit.

a) Si α est égal à 0, on est ramené à calculer la primitive d'un polynôme.

b) Si $P(x)$ est une constante, c'est immédiat car on a $\int \mathrm{e}^{\alpha x}\, dx \simeq \mathrm{e}^{\alpha x}/\alpha$.

c) Dans le cas général, on utilise la formule de primitivation par parties pour abaisser successivement le degré de P jusqu'à 0 et se ramener de la sorte au cas b). Ainsi, pour un polynôme P de degré p, on a successivement

$$\begin{aligned}\int P(x)\mathrm{e}^{\alpha x}\, dx &\simeq \int P(x)\mathrm{D}\frac{\mathrm{e}^{\alpha x}}{\alpha}\, dx \simeq P(x)\frac{\mathrm{e}^{\alpha x}}{\alpha} - \int \frac{\mathrm{D}P(x)}{\alpha}\mathrm{e}^{\alpha x}\, dx \\ &\simeq \quad \ldots \quad \simeq \frac{\mathrm{e}^{\alpha x}}{\alpha}\sum_{k=0}^{p}(-1)^k \frac{\mathrm{D}^k P(x)}{\alpha^k}. \blacksquare\end{aligned}$$

Exemple. *On a*

$$\int x\mathrm{e}^{-x}\, dx \simeq \int x\,(-\mathrm{D}\mathrm{e}^{-x})\, dx \simeq -x\mathrm{e}^{-x} + \int \mathrm{e}^{-x}\, dx \simeq -x\,\mathrm{e}^{-x} - \mathrm{e}^{-x}.$$

Passons à présent aux applications de cette méthode.

1. Calcul de $\int P(x, \cos(ax), \sin(ax), \ldots, \cos(bx), \sin(bx))\, dx$

En utilisant les formules qui expriment une puissance entière de $\cos(rx)$ et de $\sin(rx)$ comme combinaison linéaire des cosinus et des sinus des arcs multiples, puis les formules

$$\begin{aligned}\cos(ax)\cdot\cos(bx) &= \tfrac{1}{2}(\cos((a+b)x) + \cos((a-b)x)), \\ \sin(ax)\cdot\sin(bx) &= \tfrac{1}{2}(\cos((a-b)x) - \cos((a+b)x)), \\ \sin(ax)\cdot\cos(bx) &= \tfrac{1}{2}(\sin((a+b)x) + \sin((a-b)x)),\end{aligned}$$

on voit que $P(x, \cos(ax), \sin(ax), \ldots, \cos(bx), \sin(bx))$ peut s'écrire sous la forme d'une combinaison linéaire de fonctions de la forme $P'(x)\cdot\cos(rx)$ ou $P'(x)\cdot\sin(rx)$, où $P'(x)$ désigne un polynôme et r un nombre réel.

Procédé de calcul. Pour calculer la primitive d'une fonction du type

$$P(x, \cos(ax), \sin(ax), \dots, \cos(bx), \sin(bx)),$$

il suffit de donner un procédé permettant de primitiver une fonction d'un des types $P(x) \cdot \cos(ax)$ ou $P(x) \cdot \sin(ax)$.

a) Si a est égal à 0, on est ramené à calculer la primitive d'un polynôme.

b) Si $P(x)$ est une constante, c'est immédiat car on a

$$\int \cos(ax)\, dx \simeq \frac{\sin(ax)}{a} \quad \text{et} \quad \int \sin(ax)\, dx \simeq -\frac{\cos(ax)}{a}.$$

c) Dans le cas général, on utilise la formule de primitivation par parties pour abaisser successivement le degré de P jusqu'à 0 et ainsi se ramener au cas a).

Exemples. a) *On a*

i) $\displaystyle\int \cos^2(x)\, dx \simeq \int \frac{1+\cos(2x)}{2}\, dx \simeq \frac{x}{2} + \frac{\sin(2x)}{4}$ *sur* $\mathbb{R}$,

ii) $\displaystyle\int \sin^2(x)\, dx \simeq \int \frac{1-\cos(2x)}{2}\, dx \simeq \frac{x}{2} - \frac{\sin(2x)}{4}$ *sur* $\mathbb{R}$.

b) *Etablir une formule de récurrence pour le calcul de*

$$I_m = \int \sin^m(x)\, dx, \quad \forall m \in \mathbb{N}, \ \textit{sur} \ \mathbb{R}.$$

On a bien sûr

$$\int \sin(x)\, dx \simeq -\cos(x) \quad \text{et} \quad \int \sin^2(x)\, dx \simeq \frac{x}{2} - \frac{1}{4}\sin(2x).$$

Pour $m \geq 3$, on a successivement

$$\begin{aligned} \int \sin^m(x)\, dx &\simeq -\int \sin^{m-1}(x) \cdot \mathrm{D}\cos(x)\, dx \\ &\simeq -\cos(x) \cdot \sin^{m-1}(x) + (m-1)\int (1-\sin^2(x)) \sin^{m-2}(x)\, dx \end{aligned}$$

donc

$$mI_m \simeq (m-1)I_{m-2} - \cos(x) \cdot \sin^{m-1}(x),$$

ce qui suffit.

Remarque. Pour le calcul de

$$\int e^{ax}\cos(bx)\,dx \quad \text{ou} \quad \int e^{ax}\sin(bx)\,dx \text{ sur } \mathbb{R} \text{ avec } a,b\in\mathbb{R},$$

on pourrait dès lors être tenté d'utiliser les formules

$$\cos(bx)=\frac{1}{2}(e^{ibx}+e^{-ibx}) \quad \text{et} \quad \sin(bx)=\frac{1}{2i}(e^{ibx}-e^{-ibx}).$$

Il est de loin préférable de procéder comme suit

$$\begin{aligned}\int e^{ax}\left\{\begin{array}{c}\cos(bx)\\ \sin(bx)\end{array}\right\}dx &\simeq \int\left\{\begin{array}{c}\Re\\ \Im\end{array}\right\}e^{(a+ib)x}\,dx \simeq \left\{\begin{array}{c}\Re\\ \Im\end{array}\right\}\int e^{(a+ib)x}\,dx\\ &\simeq \left\{\begin{array}{c}\Re\\ \Im\end{array}\right\}\frac{e^{(a+ib)x}}{a+ib} \simeq \left\{\begin{array}{c}\Re\\ \Im\end{array}\right\}\frac{(a-ib)e^{(a+ib)x}}{a^2+b^2}.\blacksquare\end{aligned}$$

2. Calcul de $\int P(x,\mathrm{ch}(ax),\mathrm{sh}(ax),\ldots,\mathrm{ch}(bx),\mathrm{sh}(bx))\,dx$

Il suffit de procéder de manière analogue.

Exemples. *On a*

a) $\displaystyle\int \mathrm{ch}^2(x)\,dx \simeq \int \frac{1+\mathrm{ch}(2x)}{2}\,dx \simeq \frac{x}{2}+\frac{\mathrm{sh}(2x)}{4}$ *sur* $\mathbb{R}$,

b) $\displaystyle\int \mathrm{sh}^2(x)\,dx \simeq \int \frac{-1+\mathrm{ch}(2x)}{2}\,dx \simeq -\frac{x}{2}+\frac{\mathrm{sh}(2x)}{4}$ *sur* $\mathbb{R}$.

8.1.7 Primitivation d'une fraction rationnelle

Remarques. a) Au moyen du changement de variable $x(t)=1/t$, le calcul d'une primitive d'une fraction rationnelle en x se ramène à celui d'une fraction rationnelle en t. Parfois on assiste à une simplification fort grande dans les calculs.

b) Si la fraction rationnelle $R(x)$ s'écrit aussi $x^{m-1}R'(x^m)$ où $m-1$ appartient à $\mathbb{N}_0$ et où $R'(x)$ est une fraction rationnelle de x, il convient d'effectuer la substitution $x^m=t$ dans la primitive $\int R(x)\,dx$: il vient

$$\left[\frac{1}{m}\int R'(t)\,dt\right]_{t=x^m} \simeq \int x^{m-1}R'(x^m)\,dx \simeq \int R(x)\,dx.$$

C'est toujours le cas pour les fractions rationnelles impaires.

Méthode générale. On sait qu'une fraction rationnelle est égale à une combinaison linéaire de monômes et de fractions du type

$$\frac{c}{(x-\alpha)^m} \quad \text{avec} \quad \alpha,c\in\mathbb{C} \quad \text{et} \quad m\in\mathbb{N}_0. \qquad (*)$$

Pour en calculer une primitive, il suffit de pouvoir primitiver ces fonctions $(*)$. Or nous avons déjà les formules suivantes:

a) $\displaystyle\int \frac{dx}{x-a} \simeq \ln|x-a|$ sur $]-\infty,a[$ et sur $]a,+\infty[$, pour tout $a\in\mathbb{R}$,

b) $\displaystyle\int \frac{dx}{(x-a)^m} \simeq \frac{1}{1-m}\frac{1}{(x-a)^{m-1}}$ sur $]-\infty,a[$ et sur$]a,+\infty[$, pour tout $a\in\mathbb{R}$ et tout entier $m\geq 2$,

c) $\displaystyle\int \frac{dx}{(x-\alpha)^m} \simeq \frac{1}{1-m}\frac{1}{(x-\alpha)^{m-1}}$ sur $\mathbb{R}$ pour tout $\alpha\in\mathbb{C}\setminus\mathbb{R}$ et tout entier $m\geq 2$.

Pour conclure, il suffit de prouver qu'on a aussi

$$\int \frac{dx}{x-\alpha} \simeq \ln|x-\alpha| + i\mathrm{arctg}\left(\frac{x-\Re\alpha}{\Im\alpha}\right)$$

sur $\mathbb{R}$ pour tout $\alpha\in\mathbb{C}\setminus\mathbb{R}$. Pour établir cette formule, on note d'abord que

$$\int \frac{dx}{x-\alpha} \simeq \int \frac{x-\Re\alpha+i\Im\alpha}{(x-\Re\alpha)^2+(\Im\alpha^2)}\,dx,$$

puis on calcule

$$\begin{aligned}\int \frac{x-\Re\alpha}{(x-\Re\alpha)^2+(\Im\alpha^2)}\,dx &\simeq \frac{1}{2}\int \frac{\mathrm{D}((x-\Re\alpha)^2+(\Im\alpha^2))}{(x-\Re\alpha)^2+(\Im\alpha^2)}\,dx\\ &\underset{(*)}{\simeq} \left[\int \frac{dt}{t}\right]_{t=(x-\Re\alpha)^2+(\Im\alpha^2)}\\ &\simeq \frac{1}{2}\ln\left((x-\Re\alpha)^2+(\Im\alpha^2)\right]\end{aligned}$$

en effectuant la substitution $t(x)=(x-\Re\alpha)^2+(\Im\alpha^2)$ en $(*)$, et

$$i\int \frac{\Im\alpha}{(x-\Re\alpha)^2+(\Im\alpha^2)}\,dx \underset{(*)}{\simeq} i\left[\int \frac{dt}{t^2+1}\right]_{t=\frac{x-\Re\alpha}{\Im\alpha}} \simeq i\mathrm{arctg}\left(\frac{x-\Re\alpha}{\Im\alpha}\right)$$

en effectuant le changement de variable linéaire $t=(x-\Re\alpha)/\Im\alpha$ en $(*)$.

Méthode si R est réel. Si R est une fraction rationnelle dont tous les coefficients sont des nombres réels, on peut éviter toute intervention des nombres complexes en procédant de la manière suivante. On sait que, dans ce cas, $R(x)$ est égal à une combinaison linéaire de monômes, de fractions rationnelles de la forme

$$\frac{r}{(x-a)^m}$$

avec a, $r \in \mathbb{R}$ et $m \in \mathbb{N}_0$, et de fractions rationnelles de la forme

$$\frac{mx+n}{(ax^2+bx+c)^p}$$

avec a, b, c, m $n \in \mathbb{R}$ tels que $a \neq 0$, $b^2-4ac<0$ et $p \in \mathbb{N}_0$.

Pour conclure, il suffit donc de donner une méthode de calcul d'une primitive de ces dernières fractions rationnelles. Cela se fait en deux étapes.

La première étape consiste à calculer

$$\int \frac{mx+n}{(ax^2+bx+c)^p}\,dx$$

sur $\mathbb{R}$, en fonction de $\int dt/(t^2+1)^p$. Pour cela, on procède comme suit

$$\begin{aligned}
&\int \frac{mx+n}{(ax^2+bx+c)^p}\,dx \\
&\quad\simeq \frac{m}{2a}\int \frac{2ax+b}{(ax^2+bx+c)^p}\,dx + (n-\frac{bm}{2a})\int \frac{dx}{(ax^2+bx+c)^p} \\
&\quad\simeq \frac{m}{2a}\left[\int \frac{dt}{t^p}\right]_{t=ax^2+bx+c} + (n-\frac{bm}{2a})\int \frac{dx}{\left(a\left((x+\frac{b}{2a})^2+\frac{4ac-b^2}{4a^2}\right)\right)^p} \\
&\quad\simeq \frac{n-\frac{bm}{2a}}{a^p}\sqrt{\frac{4ac-b^2}{4a^2}}\cdot\left(\frac{4a^2}{4ac-b^2}\right)^p\left[\int \frac{dt}{(t^2+1)^p}\right]_{t=(x+\frac{b}{2a})/\sqrt{\frac{4ac-b^2}{4a^2}}} \\
&\qquad + \left\{\begin{array}{ll} \frac{m}{2a}\,\frac{1}{1-p}\,\frac{1}{(ax^2+bx+c)^{p-1}} & \text{si } p\neq 1 \\ \frac{m}{2a}\,\ln\left|ax^2+bx+c\right| & \text{si } p=1 \end{array}\right\}.
\end{aligned}$$

La deuxième étape consiste à calculer une primitive de $(1+t^2)^{-p}$ pour $p \in \mathbb{N}_0$. On procède par récurrence. D'une part, on a

$$\int \frac{dt}{1+t^2} \simeq \operatorname{arctg}(t).$$

D'autre part, pour tout $p>1$, il vient successivement

$$\begin{aligned}
\int \frac{dt}{(t^2+1)^p} &\simeq \int \frac{t^2+1-t^2}{(t^2+1)^p}\,dt \\
&\simeq \int \frac{dt}{(t^2+1)^{p-1}} + \frac{1}{2p-2}\int t\cdot \mathrm{D}\frac{1}{(t^2+1)^{p-1}}\,dt \\
&\simeq \frac{1}{2p-2}\,\frac{t}{(t^2+1)^{p-1}} + \frac{2p-3}{2p-2}\int \frac{dt}{(t^2+1)^{p-1}},
\end{aligned}$$

ce qui suffit.∎

Exemple. *Calculer* $\int dx/(x^2 - a^2)$ *pour tout* $a > 0$. La fonction à primitiver est continue sur $]-\infty, -a[$, sur $]-a, a[$ et sur $]a, +\infty[$. Sur chacun de ces intervalles, nous avons successivement:

$$\int \frac{dx}{x^2 - a^2} \simeq \frac{1}{2a}\left(\int \frac{dx}{x-a} - \int \frac{dx}{x+a}\right) \simeq \frac{1}{2a} \ln \left|\frac{x-a}{x+a}\right|.$$

En particulier, il vient

$$\int \frac{dx}{x^2 - a^2} \simeq -\frac{1}{a}\operatorname{arcth}(x/a) \quad \text{sur} \quad]-a, a[\,,$$

résultat qu'on pouvait obtenir aussitôt au moyen de la substitution $x = at$ dans la recherche d'une primitive de $(t^2 - 1)^{-1}$.

Exemple. *Etant donné* a, b, $c \in \mathbb{R}$ *tels que* $a \neq 0$, *vérifier que*

$$\int \frac{dx}{ax^2 + bx + c}$$

est donné par

a) $\dfrac{1}{\sqrt{b^2 - 4ac}} \ln \left(\dfrac{2ax + b - \sqrt{b^2 - 4ac}}{2ax + b + \sqrt{b^2 - 4ac}}\right)$ sur les intervalles $]-\infty, p[$ et $]q, +\infty[$ si $b^2 - 4ac > 0$ et si on désigne par p et q avec $p < q$ les deux zéros réels du trinôme $ax^2 + bx + c$,

b) $\dfrac{2}{\sqrt{4ac - b^2}} \operatorname{arctg}\left(\dfrac{2ax + b}{\sqrt{4ac - b^2}}\right)$ sur $\mathbb{R}$ si $b^2 - 4ac < 0$,

c) $-\dfrac{2}{2ax + b}$ sur $]-\infty, -b/(2a)[$ et sur $]-b/(2a), +\infty[$ si $b^2 - 4ac = 0$.

Introduisons à présent des exemples de fonctions dont le calcul d'une primitive peut se ramener au calcul d'une primitive d'une fraction rationnelle.

1. Pour tous $m \in \mathbb{N}_0 \setminus \{1\}$ et a, b, c, $d \in \mathbb{R}$ tels que $ad \neq bc$, calcul de

$$\int R\left(x, \sqrt[m]{\frac{ax+b}{cx+d}}\right) dx \ \textit{sur} \ \ I \subset \left\{x : \frac{ax+b}{cx+d} > 0\right\}.$$

Remarque. En effectuant le changement de variable $x = 1/t$, ces expressions se simplifient parfois et peuvent même se calculer par substitution.

Méthode. Comme
$$\mathrm{D}\frac{ax+b}{cx+d} = \frac{ad-bc}{(cx+d)^2}$$
garde un signe constant sur I,
$$t = \sqrt[m]{\frac{ax+b}{cx+d}}$$
est un changement de variable régulier d'ordre ∞ entre I et un intervalle ouvert I'. De plus, on a
$$t^m = \frac{ax+b}{cx+d} \quad \text{et} \quad x = \frac{dt^m - b}{a - ct^m}.$$

2. Etant donné a, b, $c \in \mathbb{R}$ tels que $a \neq 0$, calcul de
$$\int R(x, \sqrt{ax^2+bx+c})\, dx.$$

Comme ax^2+bx+c doit être strictement positif dans l'intervalle où on cherche une primitive et ne peut avoir de zéro double car $(ax^2+bx+c)^{1/2}$ pourrait alors s'écrire $dx+e$ avec d, $e \in \mathbb{R}$, on est amené à n'envisager que les trois cas suivants:

a) $a > 0$ et $b^2 - 4ac > 0$: on travaille sur un intervalle inclus dans $]-\infty, p[$ ou dans $]q, +\infty[$, où p et q sont tels que $p < q$ et désignent les deux zéros réels de ax^2+bx+c,

b) $a > 0$ et $b^2 - 4ac < 0$: on travaille sur un intervalle ouvert de $\mathbb{R}$,

c) $a < 0$ et $b^2 - 4ac > 0$: on travaille sur un intervalle ouvert inclus dans $]p, q[$, où p et q sont tels que $p < q$ et désignent les deux zéros réels de ax^2+bx+c.

Considérons séparément chacun de ces cas.

a) *Cas $a > 0$ et $b^2 - ac > 0$.* La fonction
$$\sqrt{a}\, x + \sqrt{ax^2+bx+c}$$
appartient à $\mathrm{C}_\infty(]-\infty, p[\cup]q, +\infty[)$. Sa dérivée
$$\frac{2\sqrt{a}\sqrt{ax^2+bx+c} + 2ax + b}{2\sqrt{ax^2+bx+c}}$$
a le signe de son numérateur. Or on a
$$2ax + b = 2a\left(x - \frac{p+q}{2}\right).$$
Dès lors, sa dérivée est évidemment strictement positive sur $]q, +\infty[$. De plus, sa dérivée est strictement négative sur $]-\infty, p[$ car on a $2ax + b < 0$ et
$$4a(ax^2+bx+c) < 4a^2x^2 + 4abx + b^2$$

en tout point de cet intervalle. Au total,

$$\sqrt{a}\, x + \sqrt{ax^2 + bx + c} = t$$

est un changement de variable régulier d'ordre infini entre $]-\infty, p[$ (resp. $]q, +\infty[$) et un certain intervalle ouvert de $\mathbb{R}$.

En élevant au carré la relation $\sqrt{ax^2 + bx + c} = -\sqrt{a}\, x + t$, il vient $bx + c = -2\sqrt{a}\, xt + t^2$, d'où on tire $x = (t^2 - c)/(2\sqrt{a}\, t + b)$. Dès lors,

$$\int R(x, \sqrt{ax^2 + bx + c})\, dx$$

s'obtient, par ce changement de variable, au moyen du calcul d'une primitive d'une fraction rationnelle en t, à savoir explicitement

$$\int R\left(\frac{t^2 - c}{2\sqrt{a}\, t + b}, -\sqrt{a}\frac{t^2 - c}{2\sqrt{a}\, t + b} + t\right) \cdot \frac{2\sqrt{a}\, t^2 + 2bt + 2\sqrt{a}\, c}{(2\sqrt{a}\, t + b)^2}\, dt.$$

b) *Cas* $a > 0$ *et* $b^2 - 4ac < 0$. La fonction

$$\sqrt{a}\, x + \sqrt{ax^2 + bx + c}$$

appartient à $\mathrm{C}_\infty(\mathbb{R})$. De plus, en tout $x \in \mathbb{R}$, sa dérivée a le signe de

$$\sqrt{ax^2 + bx + c} + \sqrt{a}\, x + \frac{b}{2\sqrt{a}}$$

donc est strictement positive car on a

$$\sqrt{ax^2 + bx + c} > \left|\sqrt{a}\, x + \frac{b}{2\sqrt{a}}\right| \quad \Leftrightarrow \quad c > \frac{b^2}{4a}.$$

Dès lors,

$$\sqrt{a}\, x + \sqrt{ax^2 + bx + c} = t$$

est un changement de variable régulier d'ordre infini entre $\mathbb{R}$ et un certain intervalle ouvert de $\mathbb{R}$ qui ramène le calcul de la primitive cherchée à celui d'une fraction rationnelle en t.

c) *Cas* $a < 0$ *et* $b^2 - 4ac > 0$. La fonction

$$\sqrt{a\frac{x - q}{x - p}}$$

appartient à $\mathrm{C}_\infty(]p, q[)$. De plus, sa dérivée

$$\frac{1}{2}\sqrt{a\frac{x - p}{x - q}} \cdot \frac{p - q}{(x - p)^2}$$

est strictement négative en tout $x \in \,]p, q[$. De là,

$$\sqrt{a\frac{x-q}{x-p}} = t$$

est un changement de variable régulier d'ordre infini entre $]p, q[$ et $]0, +\infty[$. En élevant cette relation au carré, on obtient

$$x = \frac{aq - pt^2}{a - t^2};$$

de plus, on a bien sûr

$$\sqrt{ax^2 + bx + c} = \sqrt{a(x-p)(x-q)} = (x-p)\sqrt{a\frac{x-q}{x-p}}.$$

Dès lors, le calcul de la primitive se ramène, par ce changement de variable, à celui d'une primitive d'une fraction rationnelle de t, à savoir explicitement

$$\int R\left(\frac{aq - pt^2}{a - t^2}, (\frac{aq - pt^2}{a - t^2} - p)t\right) \cdot \frac{2a(q-p)t}{(a-t^2)^2}\, dt.$$

Exemple. *Calculer* $\int (a^2 + x^2)^{-1/2}\, dx$ *pour tout* $a > 0$.

La fonction $(a^2+x^2)^{-1/2}$ est continue, donc primitivable sur $\mathbb{R}$. Vu ce qui précède, le changement de variable $x + (a^2 + x^2)^{1/2} = t$ est régulier d'ordre ∞ entre $\mathbb{R}$ et $]0, +\infty[$. En élevant au carré la relation $(a^2 + x^2)^{1/2} = t - x$, il vient $a^2 = -2tx + t^2$. Dès lors, on obtient

$$\int \frac{dx}{\sqrt{a^2 + x^2}} \simeq \left[\int \frac{1}{t - x(t)}\, \frac{t - x(t)}{t}\, dt\right]_{t = x + \sqrt{a^2 + x^2}} \simeq \ln(x + \sqrt{a^2 + x^2}),$$

sur $\mathbb{R}$, c'est-à-dire

$$\int \frac{dx}{\sqrt{a^2 + x^2}} \simeq \operatorname{arcsh}(x/a) \quad \text{sur} \quad \mathbb{R},$$

ce qu'on aurait obtenu bien plus directement au moyen du changement de variable $x = at$.

Exemple. *Calculer* $\int (a^2 + x^2)^{1/2}\, dx$ *pour tout* $a > 0$.

La fonction $(a^2 + x^2)^{1/2}$ est continue, donc primitivable, sur $\mathbb{R}$.

On peut éviter la méthode donnée en procédant comme suit. De

$$\begin{aligned} \int \sqrt{a^2 + x^2}\, dx &\simeq \int \frac{a^2 + x^2}{\sqrt{a^2 + x^2}}\, dx \\ &\simeq a^2 \ln(x + \sqrt{a^2 + x^2}) + \int x \mathrm{D}\sqrt{a^2 + x^2}\, dx \\ &\simeq a^2 \ln(x + \sqrt{a^2 + x^2}) + x\sqrt{a^2 + x^2} - \int \sqrt{a^2 + x^2}\, dx, \end{aligned}$$

on tire de suite

$$\int \sqrt{a^2+x^2}\,dx \simeq \frac{a^2}{2}\ln\left(x+\sqrt{a^2+x^2}\right)+\frac{x}{2}\sqrt{a^2+x^2} \quad \text{sur} \quad \mathbb{R}.$$

Exemple. *Calculer* $\int (x^2-a^2)^{-1/2}\,dx$ *pour tout* $a>0$.

La fonction $(x^2-a^2)^{-1/2}$ est continue, donc primitivable, sur $]a,+\infty[$. Ici $x+(x^2-a^2)^{1/2}=t$ est un changement de variable régulier d'ordre infini entre $]a,+\infty[$ et $]a,+\infty[$. Comme en élevant $(x^2-a^2)^{1/2}=-x+t$ au carré, on obtient $-a^2=-2xt+t^2$, il vient

$$\int \frac{dx}{\sqrt{x^2-a^2}} \simeq \left[\int \frac{1}{t-x(t)}\,\frac{t-x(t)}{t}\,dt\right]_{t=x+\sqrt{x^2-a^2}} \simeq \ln\left|x+\sqrt{x^2-a^2}\right|$$

sur $]a,+\infty[$, c'est-à-dire

$$\int \frac{dx}{\sqrt{x^2-a^2}} \simeq \operatorname{arcch}(x/a) \quad \text{sur} \quad]a,+\infty[\,,$$

ce qu'on aurait pu obtenir plus facilement au moyen du changement de variable $x=at$.

Exemple. *Calculer* $\int (x^2-a^2)^{1/2}\,dx$ *pour tout* $a>0$.

La fonction $(x^2-a^2)^{1/2}$ est continue, donc primitivable, sur $]a,+\infty[$.

On peut éviter la méthode donnée en procédant comme suit. De

$$\begin{aligned}
\int \sqrt{x^2-a^2}\,dx &\simeq \int \frac{x^2-a^2}{\sqrt{x^2-a^2}}\,dx \\
&\simeq -a^2\ \ln\left|x+\sqrt{x^2-a^2}\right| + \int x\mathrm{D}\sqrt{x^2-a^2}\,dx \\
&\simeq -a^2\ \ln\left|x+\sqrt{x^2-a^2}\right| + x\,\sqrt{x^2-a^2} - \int \sqrt{x^2-a^2}\,dx,
\end{aligned}$$

on tire de suite

$$\int \sqrt{x^2-a^2}\,dx \simeq -\frac{a^2}{2}\ln\left|x+\sqrt{a^2+x^2}\right|+\frac{x}{2}\sqrt{x^2-a^2} \quad \text{sur} \quad]a,+\infty[\,.$$

Exemple. *Calculer* $\int (a^2-x^2)^{-1/2}\,dx$ *pour tout* $a>0$.

La fonction $(a^2-x^2)^{-1/2}$ est continue, donc primitivable, sur $]-a,a[$. Ici $((-x+a)/(x+a))^{1/2}=t$ est un changement de variable régulier d'ordre infini entre $]-a,a[$ et

$]0, +\infty[$. En élevant $((-x+a)/(x+a))^{1/2} = t$ au carré, on obtient $-x+a = t^2(x+a)$, d'où on tire

$$\begin{aligned}\int \frac{dx}{\sqrt{a^2-x^2}} &\simeq \left[\int \frac{1}{(x(t)+a)t}\,\frac{-2(x(t)+a)t}{1+t^2}\,dt\right]_{t=\sqrt{(-x+a)/(x+a)}}\\ &\simeq -2\,\mathrm{arctg}\left(\sqrt{\frac{-x+a}{x+a}}\right) \quad \text{sur} \quad]-a,a[\,.\end{aligned}$$

Remarquons que le changement de variable $x = at$ aurait procuré bien plus facilement le résultat

$$\int \frac{dx}{\sqrt{a^2-x^2}} \simeq \arcsin(x/a) \quad \text{sur} \quad]-a,a[\,.$$

Exemple. *Calculer* $\int (a^2-x^2)^{1/2}\,dx$ *pour tout* $a > 0$.
La fonction $(a^2-x^2)^{1/2}$ est continue, donc primitivable, sur $]-a,a[$.
On peut éviter la méthode donnée en procédant comme suit. De

$$\begin{aligned}\int \sqrt{a^2-x^2}\,dx &\simeq \int \frac{a^2-x^2}{\sqrt{a^2-x^2}}\,dx \simeq a^2\ \arcsin(x/a) + \int x\mathrm{D}\sqrt{a^2-x^2}\,dx\\ &\simeq a^2\ \arcsin(x/a) + x\,\sqrt{a^2-x^2} - \int \sqrt{a^2-x^2}\,dx,\end{aligned}$$

on tire de suite

$$\int \sqrt{a^2-x^2}\,dx \simeq \frac{a^2}{2}\arcsin(x/a) + \frac{x}{2}\sqrt{a^2-x^2} \quad \text{sur} \quad]-a,a[\,.$$

3. Etant donné a, b, c, $d \in \mathbb{R}$ tels que $ac \neq 0$, calcul de

$$\int R(x, \sqrt{ax+b}, \sqrt{cx+d})\,dx.$$

Remarques. a) La condition $ac \neq 0$ assure la non-trivialité de $(ax+b)^{1/2}$ et de $(cx+d)^{1/2}$.
b) Nous devons bien évidemment nous placer sur un intervalle ouvert I inclus dans $\{\,x : ax+b > 0\} \cap \{\,x : cx+d > 0\}$ où la fonction est continue donc primitivable.

Méthode. La fonction $(ax+b)^{1/2}$ appartient à $\mathrm{C}_\infty(I)$ et sa dérivée garde le signe de a sur I. Dès lors, $t = (ax+b)^{1/2}$ est un changement de variable régulier d'ordre infini entre I et un certain intervalle ouvert I' de $\mathbb{R}$. Comme nous avons alors $t^2 = ax + b$, la primitive à calculer s'écrit aussi

$$\left[\frac{2}{a}\int t\; R\left(\frac{t^2-b}{a}, t, \sqrt{\frac{c}{a}t^2 + \frac{ad-bc}{a}}\right)\, dt\right]_{t=\sqrt{ax+b}}$$

et on est ramené au calcul d'une primitive du type

$$\int R'(t, \sqrt{a't^2 + b't + c'})\, dt.$$

4. Calcul de $\int R(\cos(x), \sin(x), \cos(2x), \sin(2x), \ldots)\, dx$ sur $I \subset \,]-\pi, \pi[$

Méthode générale. En recourant aux formules qui expriment $\cos(mx)$ et $\sin(mx)$ pour $m \in \mathbb{N}_0$ comme combinaisons linéaires des puissances entières des fonctions $\cos(x)$ et $\sin(x)$, la primitive à calculer se met sous la forme suivante $\int R'(\cos(x), \sin(x))\, dx$. On calcule alors la primitive de la fonction rationnelle en t

$$\frac{2}{1+t^2}\; R'\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right).$$

Vu que la fonction $t(x) = \mathrm{tg}(x/2)$ appartient à $\mathrm{C}_\infty(]-\pi, \pi[)$ et donne lieu aux relations

$$\cos(x) = \frac{1-t^2}{1+t^2}, \quad \sin(x) = \frac{2t}{1+t^2}, \quad \mathrm{D}_t x(t) = \frac{2}{1+t^2},$$

on obtient, par changement de variable,

$$\int R'(\cos(x), \sin(x))\, dx \simeq 2\left[\int R'\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)\frac{dt}{1+t^2}\right]_{t=\mathrm{tg}(x/2)}.$$

Exemple. *Calculer $\int dx/\sin(x)$.*

La fonction $1/\sin(x)$ est continue donc primitivable sur l'intervalle $]0, \pi[$. On a donc successivement

$$\int \frac{dx}{\sin(x)} \simeq \left[\int \frac{1+t^2}{2t}\,\frac{2}{1+t^2}\, dt\right]_{t=\mathrm{tg}(x/2)} \simeq \ln\left(\mathrm{tg}(x/2)\right) \;\text{ sur }\;]0, \pi[\,.$$

Cas particuliers. Il y a trois cas particuliers où généralement la méthode que nous venons d'indiquer conduit à des calculs nettement plus longs que les méthodes que nous allons développer maintenant.

4.1. Calcul de $\int R(\cos(x)) \cdot \sin(x)\, dx$

Proposition 8.1.7.1 *Une fraction rationnelle* $R'(\cos(x), \sin(x))$ *s'écrit sous la forme* $R(\cos(x)) \cdot \sin(x)$ *si et seulement si on a*

$$R'(\cos(x), \sin(x)) = -R'(\cos(x), -\sin(x)).$$

Preuve. Bien sûr, $R'(\cos(x), \sin(x))$ peut toujours s'écrire

$$\frac{P_1(\cos(x)) + P_2(\cos(x)) \cdot \sin(x)}{P_3(\cos(x)) + P_4(\cos(x)) \cdot \sin(x)}$$

où P_1, P_2, P_3 et P_4 sont des polynômes: il suffit pour cela de recourir à la formule $\sin^{2m}(x) = (1 - \cos^2(x))^m$ valable pour tout $m \in \mathbb{N}_0$. En multipliant alors haut et bas par $P_3(\cos(x)) - P_4(\cos(x)) \cdot \sin(x)$, le dénominateur devient un polynôme en $\cos(x)$ et le numérateur se présente sous la forme $P'(\cos(x)) + P''(\cos(x)) \cdot \sin(x)$ où $P'(\cos(x))$ est identiquement nul si et seulement si la condition de l'énoncé est réalisée.∎

Méthode. On recommande plutôt d'effectuer la substitution $t = \cos(x)$ dans $\int R(t)\, dt$ car on a

$$\int R(\cos(x)) \cdot \sin(x)\, dx \simeq \left[- \int R(t)\, dt \right]_{t=\cos(x)};$$

on est ainsi ramené au calcul de la primitive d'une fraction rationnelle dont l'expression est généralement plus simple que celle fournie par la méthode indiquée au début.

Exemple. *Calculer* $\int dx/\sin(x)$.
La fonction

$$\frac{1}{\sin(x)} = \frac{\sin(x)}{1 - \cos^2(x)}$$

est continue sur $]0, \pi[$. Le recours à la substitution $t = \cos(x)$ donne aussitôt

$$\int \frac{dx}{\sin(x)} \simeq \left(- \int \frac{dt}{1 - t^2} \right)_{t=\cos(x)} \simeq \frac{1}{2} \ln \left| \frac{\cos(x) - 1}{\cos(x) + 1} \right| \simeq \ln(\operatorname{tg}(x/2))$$

sur $]0, \pi[$.

4.2. Calcul de $\int R(\sin(x)) \cdot \cos(x)\, dx$

Proposition 8.1.7.2 *Une fraction rationnelle $R'(\cos(x), \sin(x))$ s'écrit sous la forme $R(\sin(x)) \cdot \cos(x)$ si et seulement si on a*

$$R'(\cos(x), \sin(x)) = -R'(-\cos(x), \sin(x)).$$

Preuve. La démonstration est analogue à celle relative à la fonction $R(\cos(x)) \cdot \sin(x)$: on recourt à la formule $\cos^{2m}(x) = (1 - \sin^2(x))^m$ valable pour tout $m \in \mathbb{N}_0$.∎

Méthode. On recommande plutôt d'effectuer la substitution $t = \sin(x)$ dans $\int R(t)\, dt$ car on a

$$\int R(\sin(x)) \cdot \cos(x)\, dx \simeq \left[\int R(t)\, dt \right]_{t=\sin(x)}.$$

Exemple. *Calculer $\int dx/\cos(x)$.* La fonction

$$\frac{1}{\cos(x)} = \frac{\cos(x)}{1 - \sin^2(x)}$$

est continue sur $]-\pi/2, \pi/2[$. Le recours à la substitution $t = \sin(x)$ donne aussitôt

$$\begin{aligned} \int \frac{dx}{\cos(x)} &\simeq \left[\int \frac{dt}{1-t^2} \right]_{t=\sin(x)} \simeq \frac{1}{2} \ln \left| \frac{1+\sin(x)}{1-\sin(x)} \right| \\ &\simeq \frac{1}{2} \ln \left| \frac{\cos^2(\pi/4 - x/2)}{\sin^2(\pi/4 - x/2)} \right| \simeq -\ln\left(\operatorname{tg}(\pi/4 - x/2)\right) \end{aligned}$$

sur $]-\pi/2, \pi/2[$.∎

4.3. Calcul de $\int R(\operatorname{tg}(x))\, dx$ sur $I \subset\,]-\pi/2, \pi/2[$

Proposition 8.1.7.3 *Une fraction rationnelle $R'(\cos(x), \sin(x))$ s'écrit sous la forme $R(\operatorname{tg}(x))$ si et seulement si on a*

$$R'(\cos(x), \sin(x)) = R'(-\cos(x), -\sin(x)).$$

Preuve. En recourant systématiquement aux formules

$$\sin(x) = \operatorname{tg}(x) \cdot \cos(x) \text{ et } \cos^2(x) = \frac{1}{1 + \operatorname{tg}^2(x)}$$

$R'(\cos(x), \sin(x))$ peut toujours s'écrire

$$\frac{P_1(\mathrm{tg}(x)) + P_2(\mathrm{tg}(x)) \cdot \cos(x)}{P_3(\mathrm{tg}(x)) + P_4(\mathrm{tg}(x)) \cdot \cos(x)}$$

où P_1, P_2, P_3 et P_4 sont des polynômes. En multipliant alors haut et bas par la fonction $P_3(\mathrm{tg}(x)) - P_4(\mathrm{tg}(x)) \cdot \cos(x)$, le dénominateur devient un polynôme en $\mathrm{tg}(x)$ et le numérateur se présente sous la forme $P'(\mathrm{tg}(x)) + P''(\mathrm{tg}(x)) \cdot \cos(x)$ où $P''(\mathrm{tg}(x)) \cdot \cos(x)$ est identiquement nul vu l'hypothèse, ce qui suffit.∎

Méthode. On recommande plutôt d'effectuer le changement de variable $t = \mathrm{tg}(x)$ car on a

$$\int R(\mathrm{tg}(x))\, dx \simeq \left[\int \frac{R(t)}{1+t^2}\, dt\right]_{t=\mathrm{tg}(x)}.$$

Exemple. *Calculer* $\int \mathrm{tg}(x)\, dx$.

a) Méthode générale. Par le changement de variable $\mathrm{tg}(x/2) = t$ sur $]-\pi/2, \pi/2[$, il vient

$$\begin{aligned}\int \frac{\sin(x)}{\cos(x)}\, dx &\simeq 2\left[\int \frac{2t}{1+t^2}\,\frac{1+t^2}{1-t^2}\,\frac{dt}{1+t^2}\right]_{t=\mathrm{tg}(x/2)} \\ &\simeq 4\left[\int \frac{t\, dt}{(1-t^2)(1+t^2)}\right]_{t=\mathrm{tg}(x/2)} \simeq \ldots\end{aligned}$$

b) Méthode conseillée. Par le changement de variable $\mathrm{tg}(x) = t$, il vient

$$\begin{aligned}\int \mathrm{tg}(x)\, dx &\simeq \left[\int \frac{t}{1+t^2}\, dt\right]_{t=\mathrm{tg}(x)} \simeq \frac{1}{2}\left[\int \frac{\mathrm{D}(1+t^2)}{1+t^2}\, dt\right]_{t=\mathrm{tg}(x)} \\ &\simeq \frac{1}{2}\ln\left(1+\mathrm{tg}^2(x)\right) \simeq -\ln\cos(x) \quad \text{sur} \quad]-\pi/2, \pi/2[\,.\end{aligned}$$

c) On peut aussi procéder de la manière suivante:

$$\int \frac{\sin(x)}{\cos(x)}\, dx \simeq -\int \frac{\mathrm{D}\cos(x)}{\cos(x)}\, dx \simeq -\ln|\cos(x)| \quad \text{sur} \quad]-\pi/2, \pi/2[\,.$$

Exemple. *Calculer* $\int \mathrm{tg}^{-2}(x)\, dx$ *sur* $]0, \pi/2[$.

a) Méthode conseillée. Par le changement de variable $\mathrm{tg}(x) = t$, il vient

$$\begin{aligned}\int \frac{dx}{\mathrm{tg}^2(x)} &\simeq \left[\int \frac{dt}{t^2(1+t^2)}\right]_{t=\mathrm{tg}(x)} \simeq \left[\int \frac{1+t^2-t^2}{t^2(1+t^2)}\, dt\right]_{t=\mathrm{tg}(x)} \\ &\simeq -\mathrm{cotg}(x) - x\end{aligned}$$

et—heureusement—on n'a pas dû procéder à la décomposition de la fraction rationnelle en fractions rationnelles simples.

b) On peut aussi procéder de la manière suivante, qui s'avère beaucoup plus directe

$$\int \frac{dx}{\operatorname{tg}^2(x)} \simeq \int \frac{1-\sin^2(x)}{\sin^2(x)}\, dx \simeq -\operatorname{cotg}(x) - x.$$

5. Calcul de $\int R(\operatorname{ch}(x), \operatorname{sh}(x), \operatorname{ch}(2x), \operatorname{sh}(2x), \ldots)\, dx$

La recherche d'une primitive d'une telle fonction est entièrement parallèle à celle effectuée pour les fonctions rationnelles en les fonctions trigonométriques. Aussi nous allons nous limiter à donner la méthode à suivre.

Méthode générale. La substitution $t = \operatorname{th}(x/2)$ conduit aux formules

$$\operatorname{ch}(x) = \frac{1+t^2}{1-t^2}, \quad \operatorname{sh}(x) = \frac{2t}{1-t^2}, \quad \mathrm{D}_t x(t) = \frac{2}{1-t^2}$$

et

$$\int R'(\operatorname{ch}(x), \operatorname{sh}(x))\, dx \simeq 2\left[\int R'\left(\frac{1+t^2}{1-t^2}, \frac{2t}{1-t^2}\right) \frac{dt}{1-t^2}\right]_{t=\operatorname{th}(x/2)}.$$

Les trois cas particuliers se maintiennent à des modifications évidentes près.

Proposition 8.1.7.4 *Une fraction rationnelle $R'(\operatorname{ch}(x), \operatorname{sh}(x))$ s'écrit également*

$$R(\operatorname{ch}(x)) \cdot \operatorname{sh}(x) \quad (\text{resp. } R(\operatorname{sh}(x)) \cdot \operatorname{ch}(x); R(\operatorname{th}(x)))$$

si et seulement si

$$\begin{array}{rcl} R'(\operatorname{ch}(x), \operatorname{sh}(x)) & = & -R'(\operatorname{ch}(x), -\operatorname{sh}(x)) \\ (\text{resp. } R'(\operatorname{ch}(x), \operatorname{sh}(x)) & = & -R'(-\operatorname{ch}(x), \operatorname{sh}(x)); \\ R'(\operatorname{ch}(x), \operatorname{sh}(x)) & = & R'(-\operatorname{ch}(x), -\operatorname{sh}(x))). \end{array}$$

Preuve. La démonstration est analogue à celle relative à $R'(\cos(x), \sin(x))$. ∎

5.1. Calcul de $\int R(\operatorname{ch}(x)) \cdot \operatorname{sh}(x)\, dx$.

Méthode. On effectue la substitution $t = \operatorname{ch}(x)$; il vient

$$\int R(\operatorname{ch}(x)) \cdot \operatorname{sh}(x)\, dx \simeq \left[\int R(t)\, dt\right]_{t=\operatorname{ch}(x)}.$$

5.2. Calcul de $\int R(\mathrm{sh}(x)) \cdot \mathrm{ch}(x)\, dx$

Méthode. On effectue la substitution $t = \mathrm{sh}(x)$; il vient

$$\int R(\mathrm{sh}(x)) \cdot \mathrm{ch}(x)\, dx \simeq \left[\int R(t)\, dt\right]_{t=\mathrm{sh}(x)}.$$

5.3. Calcul de $\int R(\mathrm{th}(x))\, dx$

Méthode. On effectue la substitution $t = \mathrm{th}(x)$; il vient

$$\int R(\mathrm{th}(x))\, dx \simeq \left[\int \frac{R(t)}{1-t^2}\, dt\right]_{t=\mathrm{th}(x)}.$$

8.1.8 Pour a, b, $c \in \mathbb{R}$ tels que $a \neq 0$, passage de

$$\int R(x, \sqrt{ax^2+bx+c})\, dx \qquad \textbf{à} \qquad \int R'(\cos(\varphi), \sin(\varphi))\, d\varphi$$
$$\textbf{ou à} \qquad \int R'(\mathrm{ch}(\varphi), \mathrm{sh}(\varphi))\, d\varphi$$

Tout est basé sur la formule bien connue

$$ax^2+bx+c = a\left((x+\frac{b}{2a})^2 + \frac{4ac-b^2}{4a^2}\right) \qquad \text{si} \quad a \neq 0.$$

Envisageons séparément les trois cas où on considère $(ax^2+bx+c)^{1/2}$.

a) *Cas $a > 0$ et $b^2 - 4ac < 0$*

La fonction $(ax^2+bx+c)^{1/2}$ appartient à $\mathrm{C}_\infty(\mathbb{R})$ et le changement de variable linéaire

$$x + \frac{b}{2a} = \frac{\sqrt{4ac-b^2}}{2a}\, t$$

donne

$$\begin{aligned}&\int R(x, \sqrt{ax^2+bx+c})\, dx \\ &\simeq \left[\int R\left(-\frac{b}{2a} + \frac{\sqrt{4ac-b^2}}{2a}\, t, \sqrt{\frac{4ac-b^2}{4a}}\ \sqrt{t^2+1}\right) \frac{\sqrt{4ac-b^2}}{2a}\, dt\right]_{t=\ldots}\end{aligned}$$

Cela étant, le changement de variable régulier d'ordre infini $t = \mathrm{sh}(\varphi)$ entre $\mathbb{R}$ et $\mathbb{R}$ donne le résultat.

b) *Cas $a > 0$ et $b^2 - 4ac > 0$*

La fonction $(ax^2 + bx + c)^{1/2}$ appartient à $C_\infty(]-\infty, p[\cup]q, +\infty[)$, où p et q sont les zéros réels de $ax^2 + bx + c$, avec $p < q$. De plus,

$$x + \frac{b}{2a} = \pm \frac{\sqrt{b^2 - 4ac}}{2a}\, t$$

est un changement de variable linéaire entre $\left\{ \begin{array}{c}]q, +\infty[\\ ]-\infty, p[\end{array} \right\}$ et $]1, +\infty[$, qui donne lieu à

$$\begin{aligned} &\textstyle\int R(x, \sqrt{ax^2 + bx + c})\, dx \\ &\simeq \left[\pm \int R\left(-\frac{b}{2a} \pm \frac{\sqrt{b^2 - 4ac}}{2a}\, t, \sqrt{\frac{b^2 - 4ac}{4a}}\, \sqrt{t^2 - 1} \right) \frac{\sqrt{b^2 - 4ac}}{2a}\, dt \right]_{t=\ldots} \end{aligned}$$

Cela étant, le changement de variable régulier d'ordre infini $t = \operatorname{ch}(\varphi)$ entre $]1, +\infty[$ et $]0, +\infty[$ donne le résultat.

c) *Cas $a < 0$ et $b^2 - 4ac > 0$*

La fonction $(ax^2 + bx + c)^{1/2}$ appartient à $C_\infty(]p, q[)$ où p et q sont les zéros réels de $ax^2 + bx + c$, avec $p < q$. Le changement de variable linéaire

$$x + \frac{b}{2a} = \frac{\sqrt{b^2 - 4ac}}{2a}\, t$$

est régulier d'ordre infini entre $]p, q[$ et $]-1, 1[$, et donne

$$\begin{aligned} &\textstyle\int R(x, \sqrt{ax^2 + bx + c})\, dx \\ &\simeq \left[\int R\left(-\frac{b}{2a} + \frac{\sqrt{b^2 - 4ac}}{2a}\, t, \sqrt{\frac{b^2 - 4ac}{4\,|a|}}\, \sqrt{1 - t^2} \right) \frac{\sqrt{b^2 - 4ac}}{2a}\, dt \right]_{t=\ldots} \end{aligned}$$

Cela étant, le changement de variable régulier d'ordre infini $t = \sin(\varphi)$ entre $]-1, 1[$ et $]-\pi/2, \pi/2[$ donne le résultat.

Exemple. *Calculer $\int (x^2 \pm a^2)^{-1/2}\, dx$ pour tout $a > 0$.*

En effectuant le changement de variable régulier d'ordre infini $x = a \operatorname{sh}(\varphi)$ entre $\mathbb{R}$ et $\mathbb{R}$, il vient directement

$$\int \frac{dx}{\sqrt{x^2 + a^2}} \simeq \left(\int \frac{a \operatorname{ch}(\varphi)}{a \operatorname{ch}(\varphi)}\, d\varphi \right)_{\varphi = \operatorname{arcsh}(x/a)} \simeq \operatorname{arcsh}(x/a) \quad \text{sur} \quad \mathbb{R}.$$

De même, le changement de variable régulier d'ordre infini $x = a \operatorname{ch}(\varphi)$ entre $]a, +\infty[$ et $]0, +\infty[$ donne

$$\int \frac{dx}{\sqrt{x^2 - a^2}} \simeq \left[\int \frac{a \operatorname{sh}(\varphi)}{a \operatorname{sh}(\varphi)}\, d\varphi \right]_{\varphi = \operatorname{arcch}(x/a)} \simeq \operatorname{arcch}(x/a)$$

sur $]a, +\infty[$. Enfin le changement de variable régulier d'ordre infini $x = -a\,\mathrm{ch}(\varphi)$ entre $]-\infty, -a[$ et $]0, +\infty[$ donne

$$\int \frac{dx}{\sqrt{x^2 - a^2}} \simeq -\left[\int \frac{a\,\mathrm{sh}(\varphi)}{a\,\mathrm{sh}(\varphi)}\,d\varphi\right]_{\varphi=-\mathrm{arcch}(x/a)} \simeq -\mathrm{arcch}(-x/a)$$

sur $]-\infty, -a[$.

Exemple. *Calculer* $\int (a^2 - x^2)^{-1/2}\,dx$ *pour tout* $a > 0$.

Le changement de variable régulier d'ordre infini $x = a\,\sin(\varphi)$ entre $]-a, a[$ et $]-\pi/2, \pi/2[$ donne

$$\int \frac{dx}{\sqrt{a^2 - x^2}} \simeq \left[\int \frac{a\,\cos(\varphi)}{a\,\cos(\varphi)}\,d\varphi\right]_{\varphi=\arcsin(x/a)} \simeq \arcsin(x/a) \quad \text{sur} \quad]-a, a[\,.\blacksquare$$

8.1.9 Primitivation des fonctions inverses

Méthode. Pour calculer $\int f(x)\cdot g(x)\,dx$ où g est une fonction inverse et où f est une fonction continue sur l'intervalle ouvert I, on commence par calculer une primitive F de f sur I puis on primitive par parties: on obtient

$$\int f(x)\cdot g(x)\,dx \simeq F(x)\cdot g(x) - \int F(x)\cdot \mathrm{D}g(x)\,dx.$$

Exemples. On a successivement:

$$\begin{array}{lll}
\text{a) } \int \ln(x)\,dx & \simeq & \int \ln(x)\cdot \mathrm{D}x\,dx \simeq x\,\ln(x) - x \quad \text{sur} \quad]0, +\infty[\,, \\
\text{b) } \int \arcsin(x)\,dx & \simeq & x\,\arcsin(x) - \int \dfrac{x}{\sqrt{1-x^2}}\,dx \\
 & \simeq & x\,\arcsin(x) + \sqrt{1-x^2} \quad \text{sur} \quad]-1, 1[\,, \\
\text{c) } \int \mathrm{arctg}(x)\,dx & \simeq & x\,\mathrm{arctg}(x) - \int \dfrac{x}{1+x^2}\,dx \\
 & \simeq & x\,\mathrm{arctg}(x) - \frac{1}{2}\ln(1+x^2) \quad \text{sur} \quad \mathbb{R}, \\
\text{d) } \int \mathrm{arcsh}(x)\,dx & \simeq & x\,\mathrm{arcsh}(x) - \int \dfrac{x}{\sqrt{1+x^2}}\,dx \\
 & \simeq & x\,\mathrm{arcsh}(x) - \sqrt{1+x^2} \quad \text{sur} \quad \mathbb{R}, \\
\text{e) } \int \mathrm{arcch}(x)\,dx & \simeq & x\,\mathrm{arcch}(x) - \int \dfrac{x}{\sqrt{x^2-1}}\,dx \\
 & \simeq & x\,\mathrm{arcch}(x) - \sqrt{x^2-1} \quad \text{sur} \quad]1, +\infty[\,, \\
\text{f) } \int \mathrm{arcth}(x)\,dx & \simeq & x\,\mathrm{arcth}(x) - \int \dfrac{x}{1-x^2}\,dx \\
 & \simeq & x\,\mathrm{arcth}(x) + \frac{1}{2}\ln(1-x^2) \qquad \text{sur} \quad]-1, 1[\,, \\
\text{g) } \int \mathrm{arccoth}(x)\,dx & \simeq & x\,\mathrm{arccoth}(x) - \int \dfrac{x}{1-x^2}\,dx \\
 & \simeq & x\,\mathrm{arccoth}(x) + \frac{1}{2}\ln(x^2-1) \quad \text{sur} \quad]1, +\infty[\,.
\end{array}$$

8.2 Primitivation dans $\mathbb{R}^{n>1}$

Note. Ce paragraphe recourt au calcul intégral et notamment au théorème de dérivation sous le signe des intégrales paramétriques ainsi que de la notion d'intégrale curviligne. Son étude ne peut donc être abordée qu'une fois cette théorie vue.

8.2.1 Position du problème

Comme toute fonction dérivable f sur un ouvert de $\mathbb{R}^n$ admet n dérivées premières, à savoir $\mathrm{D}_1 f, \ldots, \mathrm{D}_n f$, le problème de la primitivation sur un ouvert de $\mathbb{R}^n$ se pose de la manière suivante.

Position du problème. Soit Ω un ouvert de $\mathbb{R}^n$ et soient $f_1, \ldots, f_n$ des fonctions définies sur Ω. Ces fonctions $f_1, \ldots, f_n$ *admettent une primitive sur* Ω— on dit aussi qu'elles sont *primitivables sur* Ω—s'il existe une fonction F dérivable sur Ω telle que $\mathrm{D}_1 F = f_1, \ldots, \mathrm{D}_n F = f_n$; F est alors appelé une *primitive de* $f_1, \ldots, f_n$ *sur* Ω.

Formulons quelques remarques, fort importantes pour la suite.

Théorème 8.2.1.1 (extension) *Si les fonctions $f_1, \ldots, f_n$ admettent une primitive F_1 sur Ω_1 et une primitive F_2 sur Ω_2 et si on a $F_1 = F_2$ sur $\Omega_1 \cap \Omega_2$, alors la fonction F définie sur $\Omega_1 \cup \Omega_2$ par*

$$F(x) = \begin{cases} F_1(x) & si \quad x \in \Omega_1 \\ F_2(x) & si \quad x \in \Omega_2 \end{cases}$$

est une primitive de $f_1, \ldots, f_n$ sur $\Omega_1 \cup \Omega_2$.∎

Proposition 8.2.1.2 *Si F est une primitive de $f_1, \ldots, f_n \in \mathrm{C}_p(\Omega)$ sur Ω, alors F appartient à $\mathrm{C}_{p+1}\Omega$.*∎

Voici à présent une condition nécessaire d'existence d'une primitive sur un ouvert de $\mathbb{R}^n$.

Théorème 8.2.1.3 *Si les fonctions $f_1, \ldots, f_n \in \mathrm{C}_1(\Omega)$ admettent une primitive sur l'ouvert Ω de $\mathbb{R}^n$, alors elles vérifient les égalités croisées $\mathrm{D}_j f_k = \mathrm{D}_k f_j$ pour tous $j \neq k$ tels que $1 \leq j, k \leq n$.*

Preuve. De fait, si F est une primitive de $f_1, \ldots, f_n$ sur Ω, F appartient à $\mathrm{C}_2(\Omega)$ et le théorème d'interversion des dérivées donne $\mathrm{D}_j f_k = \mathrm{D}_j \mathrm{D}_k F = \mathrm{D}_k \mathrm{D}_j F = \mathrm{D}_k f_j$.∎

Théorème 8.2.1.4 (unicité) a) *Si F est une primitive des fonctions $f_1, \ldots,$ f_n sur l'ouvert Ω de $\mathbb{R}^n$, alors, pour tout $c \in \mathbb{C}$, $F+c$ est aussi une primitive de $f_1,$ $\ldots, f_n$ sur Ω.*

b) *Si F et F' sont des primitives des fonctions $f_1, \ldots, f_n$ sur l'ouvert connexe Ω de $\mathbb{R}^n$, alors $F - F'$ est constant sur Ω.*

Preuve. a) est trivial.

b) résulte aussitôt du théorème de l'ouvert connexe.∎

Corollaire 8.2.1.5 *Si les fonctions $f_1, \ldots, f_n$ admettent une primitive sur l'ouvert connexe Ω de $\mathbb{R}^n$, alors, pour tout $x_0 \in \Omega$ et tout $c \in \mathbb{C}$, il existe une primitive de $f_1, \ldots, f_n$ sur Ω et une seule qui prend la valeur c en x_0.*∎

8.2.2 Calcul de la valeur d'une primitive

Etablissons un résultat établissant que l'hypothèse "$(\gamma, [a, b])$ est un chemin C_1 tel que $\gamma([a, b]) \subset \Omega$" qui intervient dans l'énoncé du théorème suivant est naturelle.

Proposition 8.2.2.1 *Si Ω est un ouvert connexe de $\mathbb{R}^n$, alors, pour tous x, $y \in \Omega$, il existe un chemin C_1 par morceaux $(\gamma, [a, b])$ d'origine x et d'extrémité y tel que $\gamma([a, b]) \subset \Omega$.*

Preuve. Il suffit d'établir que, pour tout $x \in \Omega$, l'ensemble A_x des points y de Ω pour lesquels il existe un chemin C_1 par morceaux $(\gamma, [a, b])$ d'origine x et d'extrémité y tel que $\gamma([a, b]) \subset \Omega$ est égal à Ω.

Remarquons tout d'abord que, si la boule $b = \{\, z : |y - z| < r\}$ est incluse dans Ω et contient un point z_0 de A_x, alors b est inclus dans A_x. De fait, il existe alors un chemin C_1 par morceaux $(\gamma, [a, b])$ d'origine x et d'extrémité z_0 tel que $\gamma([a, b]) \subset \Omega$ et on vérifie de suite que, pour tout $y_0 \in b$,

$$\gamma_0 : [a, b+1] \to \mathbb{R}^n \quad t \mapsto \begin{cases} \gamma(t) & \text{si} \quad t \in [a, b] \\ z_0 + (t - b)(y_0 - z_0) & \text{si} \quad t \in [b, b+1] \end{cases}$$

est un chemin C_1 par morceaux d'origine x et d'extrémité y_0 pour lequel on a l'inclusion $\gamma_0([a, b+1]) \subset \Omega$.

Cela étant, comme A_x contient au moins le point x, on obtient de suite que A_x est un ouvert non vide inclus dans Ω. En fait, A_x est égal à Ω sinon $\omega = \cup_{y \in \Omega \setminus A_x} A_y$ est aussi un ouvert non vide et $\{A_x, \omega\}$ serait une partition de Ω en deux ouverts non vides.∎

Cela étant, le résultat suivant permet de calculer la valeur en tout point de Ω d'une primitive de $f_1, \ldots, f_n$ sur un ouvert connexe Ω de $\mathbb{R}^n$ si on sait que $f_1, \ldots, f_n$ admettent une primitive sur Ω et si on fixe la valeur de la primitive cherchée en un point de Ω.

Théorème 8.2.2.2 *Si les fonctions $f_1, \ldots, f_n \in \mathrm{C}_0(\Omega)$ admettent une primitive F sur l'ouvert Ω de $\mathbb{R}^n$ et si $(\gamma, [a, b])$ est un chemin C_1 tel que $\gamma([a, b]) \subset \Omega$, alors on a*

$$F(\gamma(b)) = F(\gamma(a)) + \int_a^b \sum_{k=1}^n f_k(\gamma(t)) \cdot \mathrm{D}\gamma_k(t)\, dt.$$

Preuve. Comme la fonction F appartient à $\mathrm{C}_1(\Omega)$, le théorème de dérivation des fonctions composées affirme que $F(\gamma(t))$ appartient aux espaces $\mathrm{C}_0([a, b])$ et $\mathrm{C}_1(]a, b[)$, et donne

$$\mathrm{D}_t F(\gamma(t)) = \sum_{k=1}^n [\mathrm{D}_k F]_{\gamma(t)} \cdot \mathrm{D}\gamma_k(t) = \sum_{k=1}^n f_k(\gamma(t)) \cdot \mathrm{D}\gamma_k(t).$$

Comme la fonction

$$\sum_{k=1}^n f_k(\gamma(t)) \cdot \mathrm{D}\gamma_k(t)$$

est continue sur le compact $[a, b]$, elle est intégrable sur $[a, b]$ et on a

$$\int_a^b \sum_{k=1}^n f_k(\gamma(t)) \cdot \mathrm{D}\gamma_k(t)\, dt = F(\gamma(b)) - F(\gamma(a)).$$

D'où la conclusion.∎

Voici un cas particulier fort utile.

Soit Ω un ouvert de $\mathbb{R}^n$ et soient a et x deux points distincts de Ω tels que le segment d'extrémités a et x soit inclus dans Ω. Cela étant, le segment joignant a à x, à savoir $(a + t(x - a), [0, 1])$, est un chemin C_1. Dès lors, si $f_1, \ldots, f_n \in \mathrm{C}_0(\Omega)$ admettent F comme primitive sur Ω, il vient

$$F(x) = F(a) + \sum_{k=1}^n (x_k - a_k) \cdot \int_0^1 f_k(a + t(x - a))\, dt.$$

8.2.3 Existence d'une primitive

Lemme 8.2.3.1 *Pour tout élément a d'un ouvert Ω de $\mathbb{R}^n$, l'ensemble Ω_a des points x tels que le segment d'extrémités a et x soit inclus dans Ω est un ouvert de $\mathbb{R}^n$.*

Preuve. Procédons par l'absurde.

Supposons qu'il existe un élément x de Ω_a pour lequel il n'existe pas de boule centrée en x incluse dans Ω_a. Il existe alors une suite x_m telle que

$$x_m \in \{\, y : |x - y| \leq 1/m\} \setminus \Omega_a, \quad \forall m \in \mathbb{N}_0.$$

Pour tout $m \in \mathbb{N}_0$, x_m n'appartient pas à Ω_a: il existe donc $r_m \in [0, 1]$ tel que le point $a + r_m(x_m - a)$ n'appartienne pas à Ω. Comme $[0, 1]$ est un compact, on peut extraire de la suite r_m une sous-suite convergente, soit $r_{k(m)} \to r_0 \in [0, 1]$. De là, il vient

$$a + r_{k(m)}(x_{k(m)} - a) \to a + r_0(x - a)$$

car la suite x_m, donc sa sous-suite $x_{k(m)}$, converge évidemment vers x. D'où une contradiction car, par hypothèse, $a + r_0(x - a)$ doit appartenir à Ω et nous venons de démontrer que ce point est limite d'une suite du fermé $\mathbb{R}^n \setminus \Omega$.∎

Lemme 8.2.3.2 *Pour tout compact K de $\mathbb{R}^n$ et tout point a de $\mathbb{R}^n$, l'ensemble*

$$\{\, a + r(x - a) : r \in [0, 1]\,, x \in K\}$$

est compact.

Preuve. D'une part, cet ensemble est borné car on a

$$|a + r(x - a)| \leq 2\,|a| + \sup_{y \in K} |y|\,, \quad \forall x \in K, \quad \forall r \in [0, 1]\,.$$

D'autre part, il est fermé. De fait, soient x_m une suite de K et r_m une suite de $[0, 1]$ et supposons que la suite $a + r_m(x_m - a)$ converge vers y. De la suite x_m, on peut extraire une sous-suite convergente; soit $x_{k(m)} \to x_0 \in K$. De la suite $r_{k(m)}$, on peut extraire une sous-suite convergente; soit $r_{l(k(m))} \to r_0$. D'où la conclusion car on déduit aussitôt que

$$y = \lim \big(a + r_{l(k(m))}(x_{l(k(m))} - a)\big) = a + r_0(x_0 - a).∎$$

Théorème 8.2.3.3 (existence) *Si Ω est un ouvert de $\mathbb{R}^n$, si a appartient à Ω et si $f_1, \ldots, f_n \in \mathrm{C}_1(\Omega)$ vérifient les égalités croisées sur Ω, alors*

$$F(x) = F(a) + \sum_{k=1}^{n} (x_k - a_k) \int_0^1 f_k(a + t(x - a))\, dt$$

est une primitive de $f_1, \ldots, f_n$ sur Ω_a.

Preuve. Pour tout $x \in \Omega_a$ et tout $k \leq n$, $f_k(a + t(x - a))$ est une fonction continue donc intégrable sur $[0, 1]$. La fonction F de l'énoncé est par conséquent définie sur Ω_a.

Etablissons à présent que les hypothèses du théorème de dérivation des intégrales paramétriques sont vérifiées pour chacune des intégrales qui définissent F: pour tout $k \leq n$.

a) vu le théorème de dérivation des fonctions composées, pour tout $t \in [0, 1]$, la fonction $f_k(a + t(x - a))$ appartient à $\mathrm{C}_1(\Omega_a)$ et on a

$$\mathrm{D}_j f_k(a + t(x - a)) = t\ [\mathrm{D}_j f_k]_{a+t(x-a)}.$$

b) pour tout $x \in \Omega_a$ et tout $j \leq n$, les fonctions $f_k(a + t(x - a))$ et $t\ [\mathrm{D}_j f_k]_{a+t(x-a)}$ sont continues et par conséquent, intégrables sur $[0, 1]$.

c) pour tout compact $K \subset \Omega_a$,

$$K' = \{\, a + t(x - a) : x \in K, t \in [0, 1]\}$$

est un compact inclus dans Ω_a et dès lors, pour tout $j \leq n$, il vient

$$\sup_{x \in K} \left| t\ [\mathrm{D}_j f_k]_{a+t(x-a)} \right| \leq \sup_{y \in K'} \left| [\mathrm{D}_j f_k]_y \right| \in \mathrm{L}^1([0, 1]).$$

On obtient donc que F est dérivable sur Ω_a et que

$$\begin{aligned} \mathrm{D}_j F(x) &= \int_0^1 \left(f_j(a + t(x - a)) + t \sum_{k=1}^{n} (x_k - a_k) \cdot [\mathrm{D}_j f_k]_{a+t(x-a)} \right) dt \\ &= \int_0^1 \left(f_j(a + t(x - a)) + t \sum_{k=1}^{n} (x_k - a_k) \cdot [\mathrm{D}_k f_j]_{a+t(x-a)} \right) dt \\ &= \int_0^1 \mathrm{D}_t(t \cdot f_j(a + t(x - a)))\, dt \quad = \quad f_j(x). \end{aligned}$$

D'où la conclusion.∎

Remarque. Il convient d'insister sur le fait que l'existence de la primitive a été établie sur Ω_a et non sur Ω. L'exemple qui suit est éloquent à ce sujet.

Exemple. *Etablir que les fonctions*

$$f_x(x, y) = \frac{-\,y}{x^2 + y^2} \quad \textit{et} \quad f_y(x, y) = \frac{x}{x^2 + y^2}$$

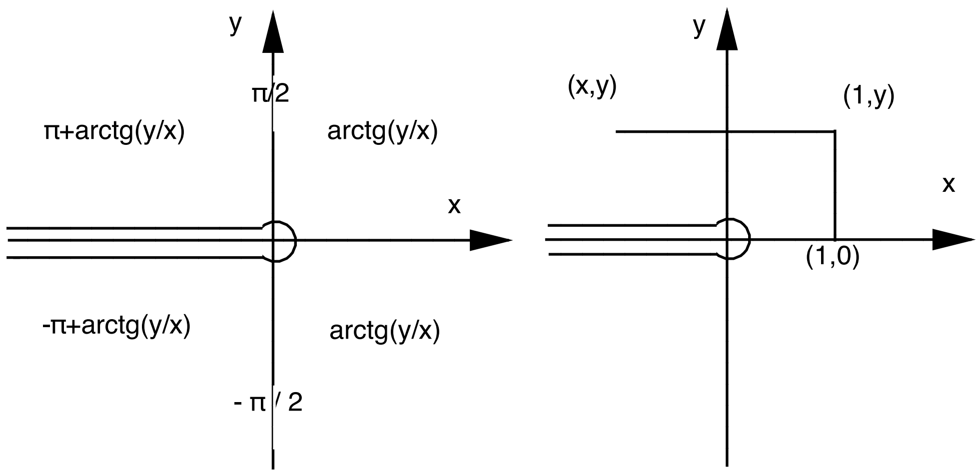
admettent pour primitive sur l'ouvert $\Omega = \mathbb{R}^2 \setminus \{\, (x, 0) : x \leq 0\}$ *la fonction* $\arg(x, y)$ *mais n'admettent pas de primitive sur* $\mathbb{R}^2 \setminus \{(0, 0)\}$.

Suggestion. Les fonctions f_x et f_y appartiennent à $\mathrm{C}_\infty(\mathbb{R}^2 \setminus \{(0,0)\}$ et vérifient les égalités croisées sur $\Omega' = \mathbb{R}^2 \setminus \{(0,0)\}$ vu que

$$\mathrm{D}_y f_x(x,y) = -\frac{x^2-y^2}{(x^2+y^2)^2} = \mathrm{D}_x f_y(x,y).$$

Le point $a = (1,0)$ appartient à Ω' et on vérifie de suite que Ω'_a est égal à l'ouvert Ω de l'énoncé.

Cela étant, on sait que f_x et f_y admettent une primitive sur Ω. Pour calculer sa valeur en un point $(x,y) \in \Omega$, nous pouvons suivre le contour suivant: parcourir le segment d'extrémités $(1,0)$ et $(x,0)$ si y est égal à 0 ou parcourir le segment d'extrémités $(1,0)$ et $(1,y)$ puis le segment d'extrémités $(1,y)$ et (x,y).



On obtient

$$F(x,y) = \begin{cases} F(1,0) + \displaystyle\int_0^y \frac{dt}{1+t^2} - \int_1^x \frac{y}{t^2+y^2}\,dt & \text{si} \quad y \neq 0 \\ F(1,0) - \displaystyle\int_1^x \frac{0}{t^2}\,dt & \text{si} \quad y = 0, \end{cases}$$

c'est-à-dire

$$F(x,y) = \begin{cases} F(1,0) + \operatorname{arctg}(y) + \operatorname{arctg}(1/y) - \operatorname{arctg}(x/y) & \text{si} \quad y \neq 0 \\ F(1,0) & \text{si} \quad y = 0 \end{cases}$$

et en recourant à l'exercice du paragraphe 6.7.9, il vient finalement

$$F(x,y) = \begin{cases} F(1,0) - \operatorname{arctg}(x/y) + \dfrac{\pi}{2}\operatorname{sign}(y) & \text{si} \quad y \neq 0 \\ F(1,0) & \text{si} \quad y = 0, \end{cases}$$

c'est-à-dire la fonction $\arg(x, y)$.

On constate ensuite que, pour $x < 0$, il vient

$$\lim_{y \to 0^+} F(x, y) = \pi \quad \text{et} \quad \lim_{y \to 0^-} F(x, y) = -\pi.$$

Il s'ensuit que f_x et f_y n'admettent pas de primitive sur $\mathbb{R}^2 \setminus \{(0,0)\}$ car une telle primitive devrait être continue sur cet ouvert et égale à une constante additive près à F.∎

Remarque. L'exemple précédent permet de construire la fonction arg et de prouver qu'elle appartient à $C_\infty(\Omega)$.

Définition. Un ouvert connexe Ω de $\mathbb{R}^n$ est *critique* s'il existe des fonctions $f_1, \ldots, f_n \in C_1(\Omega)$ qui vérifient les égalités croisées et qui n'admettent pas de primitive sur Ω.

On pourrait alors énoncer le résultat suivant: *si l'ouvert connexe Ω de $\mathbb{R}^n$ n'est pas critique et si les fonctions $f_1, \ldots, f_n \in C_1(\Omega)$ vérifient les égalités croisées, alors $f_1, \ldots, f_n$ admettent une primitive sur Ω.* Il faudrait ensuite faire l'étude des ouverts connexes non critiques, ce qui sort du cadre de ce cours.

Contentons-nous de signaler que

a) nous venons d'établir que dans $\mathbb{R}^2$, le complémentaire d'un point est critique. L'exemple précédent établit aussi que, dans $\mathbb{R}^2$, le complémentaire d'une boule fermée est critique.

b) dans $\mathbb{R}^3$, on peut établir que le complémentaire d'une droite ou d'une circonférence est critique. Mais en recourant au théorème d'extension, on prouve aisément que le complémentaire d'un point dans $\mathbb{R}^3$ n'est pas critique.

c) s'il existe $a \in \Omega$ tel que $\Omega_a = \Omega$, Ω n'est pas critique.

Méthode. Soient $f_1, \ldots, f_n$ des fonctions définies sur une partie A de $\mathbb{R}^n$.

Pour en rechercher une primitive éventuelle, on procède comme suit:

a) déterminer un plus grand ouvert $\Omega \subset A$ tel que $f_1, \ldots, f_n$ appartiennent à $C_1(\Omega)$,

b) vérifier les égalités croisées,

c) déterminer un ouvert $\Omega' \subset \Omega$ connexe et non critique, le plus grand possible (pour tout $a \in \Omega$, on est sûr que Ω_a convient),

d) calculer une primitive de $f_1, \ldots, f_n$ sur Ω': dans le choix du contour, c'est la nature des fonctions qui sert de guide (cf. les exercices),

e) vérifier si cette primitive ne s'étend pas à Ω.

Ce point e) provient de ce qu'il existe des fonctions $f_1, \ldots, f_n \in C_1(\Omega)$ avec Ω ouvert critique qui admettent cependant une primitive sur Ω. Ainsi la fonction

$\ln\left((x^2+y^2)^{1/2}\right)$ appartient visiblement à $\mathrm{C}_\infty(\Omega)$ où Ω est l'ouvert critique $\mathbb{R}^2 \setminus \{(0,0)\}$ et a pour dérivées

$$f_x(x,y) = \frac{x}{x^2+y^2} \quad \text{et} \quad f_y(x,y) = \frac{y}{x^2+y^2}.$$

Voici enfin une condition nécessaire et suffisante d'existence d'une primitive de $f_1, \ldots, f_n$ sur un ouvert connexe Ω de $\mathbb{R}^n$.

Théorème 8.2.3.4 *Des fonctions $f_1, \ldots, f_n$ continues sur un ouvert connexe Ω de $\mathbb{R}^n$ sont primitivables sur Ω si et seulement si $\int_\gamma \langle f, dx \rangle = 0$ pour tout chemin C_1 par morceaux $(\gamma, [a,b])$ tel que $\gamma([a,b]) \subset \Omega$ et $\gamma(a) = \gamma(b)$.* (On peut même exiger que ces chemins soient remplacés par des contours polygonaux parallèles aux axes).

Une primitive F de $f_1, \ldots, f_n$ sur Ω est alors donnée par la construction suivante:
a) on choisit un point $x_0 \in \Omega$,
b) pour tout point $x \in \Omega$ et tout chemin C_1 par morceaux $(\gamma, [a,b])$, d'origine x_0 et d'extrémité x tel que $\gamma([a,b]) \subset \Omega$, on pose

$$F(x) = F(x_0) + \int_\gamma \langle f, dx \rangle . \qquad (*)$$

Preuve. La condition est nécessaire. Cela résulte aussitôt de la remarque suivante: si $(\gamma, [a,b])$ est un chemin C_1 tel que $\gamma([a,b]) \subset \Omega$, et si F est une primitive de $f_1, \ldots, f_n$ sur Ω, $F(\gamma(t))$ appartient à $\mathrm{C}_1[a,b]$ et est tel que

$$\mathrm{D}_t F(\gamma(t)) = \sum_{k=1}^n f_k(\gamma(t)) \cdot \mathrm{D}\gamma(t)$$

donc tel que

$$\int_\gamma \langle f, dx \rangle = F(\gamma(b)) - F(\gamma(a)).$$

La condition est suffisante. Vu l'hypothèse, nous pouvons utiliser la formule (*) pour définir une fonction F sur Ω. Pour conclure, il suffit alors de vérifier que, pour tout $k \in \{1, \ldots, n\}$, tout $x \in \Omega$ et tout $h \in \mathbb{R}^n$ tel que $0 < |h| < d(x, \mathbb{R}^n \setminus \Omega)$, on a

$$\frac{1}{h}\left(F(x+he_k) - F(x)\right) = \frac{1}{h}\int_0^h f_k(x+te_k)\,dt$$

et que cette dernière intégrale converge vers $f_k(x)$ si $h \to 0$.∎

8.2.4 Application aux champs vectoriels dans $\mathbb{R}^3$

Définition. Rappelons que le champ vectoriel $f = (f_1, f_2, f_3)$ sur l'ouvert Ω de $\mathbb{R}^3$ est l'application qui, à tout point (x, y, z) de Ω associe le point

$$(f_1(x, y, z), f_2(x, y, z), f_3(x, y, z))$$

de $\mathbb{R}^3$. Il est continu, dérivable, p-fois continûment dérivable, différentiable, ... sur Ω si et seulement si les fonctions réelles f_1, f_2 et f_3 le sont.

Cela étant, il *dérive d'un potentiel*

a) *scalaire* s'il existe un fonction F, réelle et différentiable sur Ω, telle que $f = \operatorname{grad}(F)$.

b) *vectoriel* s'il existe un champ vectoriel différentiable g sur Ω tel que $f = \operatorname{rot}(g)$.

Remarques. a) Il est clair que, si le champ vectoriel f sur l'ouvert Ω de $\mathbb{R}^3$ dérive du potentiel scalaire $F \in \mathrm{C}_2(\Omega)$, alors on a $\operatorname{rot}(f) = \operatorname{rot}(\operatorname{grad}(F)) = 0$ sur Ω, vu le théorème d'interversion des dérivées.

b) De même, si le champ vectoriel f sur l'ouvert Ω de $\mathbb{R}^3$ dérive du potentiel vectoriel deux fois continûment dérivable g, alors on a $\operatorname{div}(f) = \operatorname{div}(\operatorname{rot}(g)) = 0$ sur Ω.

Définitions. Le champ vectoriel f sur l'ouvert Ω de $\mathbb{R}^3$ est

a) *irrotationnel* s'il est différentiable sur Ω et tel que $\operatorname{rot}(f) = 0$ sur Ω,

b) *indivergentiel* s'il est différentiable sur Ω et tel que $\operatorname{div}(f) = 0$ sur Ω.

Position du problème. Etant donné un champ vectoriel irrotationnel (resp. indivergentiel) sur un ouvert Ω de $\mathbb{R}^3$, quand peut-on dire qu'il dépend d'un potentiel scalaire (resp. vectoriel)?

En fait, la première question peut aussi s'exprimer sous la forme suivante: étant donné des fonctions f_1, f_2, f_3 réelles, différentiables et telles que $\operatorname{rot}(f_1, f_2, f_3) = 0$ sur Ω (c'est-à-dire vérifiant les égalités croisées), existe-t-il une fonction F réelle et différentiable sur Ω telle que $\mathrm{D}_j F = f_j$ sur Ω pour tout $j \leq n$.

Nous pouvons donc y apporter la réponse suivante.

Théorème 8.2.4.1 (champ irrotationnel) *Si $f = (f_1, f_2, f_3)$ est un champ vectoriel irrotationnel de type C_1 sur l'ouvert connexe et non critique Ω de $\mathbb{R}^3$, alors f dérive d'un potentiel scalaire $F \in \mathrm{C}_2(\Omega)$.*

De plus, la fonction réelle et différentiable G sur Ω est telle que $f = \operatorname{grad}(G)$ sur Ω si et seulement s'il existe $C \in \mathbb{R}$ tel que $G = F + C$ sur Ω. ∎

Dans le cas où l'ouvert connexe Ω est critique, le théorème du champ irrotationnel s'applique notamment à l'ouvert Ω_a quel que soit $a \in \Omega$; on vérifiera alors si le potentiel scalaire peut s'étendre à Ω.

Remarque. Pour tout $f \in \mathrm{C}_0(]0,+\infty[)$ réel, on vérifie de suite que le champ vectoriel $(f(r)x, f(r)y, f(r)z)$ sur $\mathbb{R}^3 \setminus \{0\}$ (où on a posé $r = |(x,y,z)|$) dérive du potentiel scalaire $F(r)$ où F est une primitive réelle de la fonction $tf(t)$ sur $]0,+\infty[$.

En particulier, le champ vectoriel
i) $(x,y,z)/r$ dérive du potentiel scalaire r sur $\mathbb{R}^3 \setminus \{0\}$,
ii) $(x,y,z)/r^2$ dérive du potentiel scalaire $\ln(r)$ sur $\mathbb{R}^3 \setminus \{0\}$,
iii) $-Gm\,(x,y,z)/r^3$ dérive du potentiel scalaire Gm/r sur $\mathbb{R}^3 \setminus \{0\}$.

Considérons à présent le deuxième problème.

Théorème 8.2.4.2 (champ indivergentiel) *Si Ω est un ouvert de $\mathbb{R}^3$ étoilé en $a = (a_1, a_2, a_3)$* (c'est-à-dire que $a \in \Omega$ et $\Omega = \Omega_a$) *et si le champ vectoriel $f = (f_1, f_2, f_3)$ est C_1 et indivergentiel sur Ω, alors f dérive d'un potentiel vectoriel $g \in \mathrm{C}_1(\Omega)$.*

De plus, le champ vectoriel $h \in \mathrm{C}_1(\Omega)$ est tel que $\mathrm{rot}(h) = f$ sur Ω si et seulement s'il existe $H \in \mathrm{C}_2(\Omega)$ tel que $h = g + \mathrm{grad}(H)$ sur Ω.

Preuve. Remarquons d'abord que s'il existe un tel champ vectoriel g, l'énoncé relatif à h résulte aussitôt du théorème précédent appliqué à $g - h$.

Cela étant, nous allons prouver que g défini par

$$g(x) = h(x) \wedge (x - a), \quad \forall x \in \Omega,$$

où

$$h(x) = \int_0^1 t\, f(a + t(x-a))\, dt, \quad \forall x \in \Omega,$$

convient.

On vérifie de suite que les conditions d'application du théorème de dérivation des intégrales paramétriques sont vérifiées. On a donc d'une part $h \in \mathrm{C}_1(\Omega)$ et d'autre part, par exemple, la première composante de $\mathrm{rot}(g)$ est successivement égale à

$$
\begin{aligned}
&\mathrm{D}_2(h \wedge (x-a))_3 - \mathrm{D}_3(h \wedge (x-a))_2 \\
&\quad = \mathrm{D}_2((x_2-a_2)h_1 - (x_1-a_1)h_2) - \mathrm{D}_3[(x_1-a_1)h_3 - (x_3-a_3)h_1] \\
&\quad = 2h_1 + (x_2-a_2)\mathrm{D}_2h_1 - (x_1-a_1)\mathrm{D}_2h_2 \\
&\qquad - (x_1-a_1)\mathrm{D}_3h_3 + (x_3-a_3)\mathrm{D}_3h_1 \\
&\quad = 2\int_0^1 t\, f_1(a+t(x-a))\, dt \\
&\qquad + (x_2-a_2)\int_0^1 t^2\, [(\mathrm{D}_2 f_1]_{a+t(x-a)}\, dt \\
&\qquad - (x_1-a_1)\int_0^1 t^2 \left([\mathrm{D}_2 f_2]_{a+t(x-a)} + [\mathrm{D}_3 f_3]_{a+t(x-a)}\right) dt \\
&\qquad + (x_3-a_3)\int_0^1 t^2\, [\mathrm{D}_3 f_1]_{a+t(x-a)}\, dt \\
&\quad = \int_0^1 \mathrm{D}_t\big(t^2 f_1(a+t(x-a))\big)\, dt \quad = \quad f_1(x)
\end{aligned}
$$

(pour obtenir l'avant dernière égalité, on utilise $\mathrm{div}(f) = 0$).∎

Chapitre 9

Equations différentielles

9.1 Linéaires à coefficients constants

9.1.1 Généralités

Soit $]a, b[$ un intervalle ouvert (borné ou non) de $\mathbb{R}$.

Rappelons qu'un opérateur de dérivation linéaire à coefficients constants d'ordre $p \in \mathbb{N}_0$ sur $]a, b[$ s'écrit

$$L(\mathrm{D}) = c_0 + c_1\mathrm{D} + \ldots + c_p\mathrm{D}^p \quad \text{avec} \quad c_0, c_1, \ldots, c_p \in \mathbb{C} \quad \text{et} \quad c_p \neq 0.$$

On lui associe le polynôme caractéristique

$$L(z) = c_0 + c_1 z + \ldots + c_p z^p.$$

Si $L(\mathrm{D})$ est un opérateur de dérivation linéaire à coefficients constants d'ordre $p \in \mathbb{N}_0$ sur $]a, b[$, rappelons que

a) pour toute combinaison linéaire $\sum_{j=1}^{J} d_j u_j$ d'éléments de $\mathrm{C}_p(]a, b[)$, on a

$$L(\mathrm{D}) \left(\sum_{j=1}^{J} d_j u_j \right) = \sum_{j=1}^{J} d_j L(\mathrm{D}) u_j$$

b) pour tout $u \in \mathrm{C}_p(]a, b[)$, on a $\overline{L}(\mathrm{D})u = \overline{L(\mathrm{D})\overline{u}}$.

Si, de plus, $L(\mathrm{D})$ est réel, ce qui a lieu si et seulement si tous les coefficients c_0, c_1, ..., c_p sont réels, il vient

$$\Re L(\mathrm{D})u = L(\mathrm{D})\Re u \quad \text{et} \quad \Im L(\mathrm{D})u = L(\mathrm{D})\Im u.$$

Enfin toute identité algébrique entre polynômes caractéristiques d'opérateurs de dérivation linéaires à coefficients constants sur $]a, b[$, qui ne comporte que des combinaisons linéaires et des produits, donne lieu à une identité entre les opérateurs de dérivation correspondants. En particulier, tout produit fini d'opérateurs de dérivation linéaires à coefficients constants sur $]a,b[$ est commutatif.

Voici un résultat qui va jouer un rôle essentiel dans la suite; il est complété au paragraphe 9.1.6.

Lemme 9.1.1.1 *Pour tout opérateur de dérivation $L(\mathrm{D})$ linéaire à coefficients constants d'ordre $p \in \mathbb{N}_0$ sur $]a,b[$, tout $c \in \mathbb{C}$ et tout $u \in \mathrm{C}_p(]a,b[)$, on a*

$$L(\mathrm{D})(\mathrm{e}^{cx}u(x)) = \mathrm{e}^{cx}\, L(\mathrm{D}+c)u(x).$$

Preuve. Etablissons tout d'abord par récurrence que, pour tout $k \leq p$, on a

$$\mathrm{D}^k(\mathrm{e}^{cx}u(x)) = \mathrm{e}^{cx}\,(\mathrm{D}+c)^k u(x).$$

Pour $k=0$ et $k=1$, c'est trivial. Cela étant, si c'est vrai pour $k = 0, \ldots, l$ avec $l < p$, établissons que c'est encore vrai pour $k = l+1$. C'est immédiat car on a successivement

$$\mathrm{D}^{l+1}(\mathrm{e}^{cx}u(x)) = \mathrm{D}(\mathrm{e}^{cx}\,(\mathrm{D}+c)^l u(x)) = \mathrm{e}^{cx}\,(\mathrm{D}+c)(\mathrm{D}+c)^l u(x).$$

Cela étant, il suffit de noter que, si $L(\mathrm{D}) = c_0 + c_1\mathrm{D} + \ldots + c_p\mathrm{D}^p$, il vient

$$\begin{aligned} L(\mathrm{D})(\mathrm{e}^{cx}u(x)) &= \sum_{j=0}^{p} c_j \mathrm{D}^j(\mathrm{e}^{cx}u(x)) \\ &= \sum_{j=0}^{p} c_j \mathrm{e}^{cx}\,(\mathrm{D}+c)^j u(x) \;=\; \mathrm{e}^{cx}\, L(\mathrm{D}+c)u(x). \blacksquare \end{aligned}$$

9.1.2 Position du problème

Définitions. Etant donné

a) $]a,b[$ un intervalle ouvert (borné ou non) de $\mathbb{R}$,

b) $L(\mathrm{D}) = \sum_{j=0}^{p} c_j\mathrm{D}^j$ un opérateur de dérivation linéaire à coefficients constants d'ordre p sur $]a,b[$,

c) f un élément de $\mathrm{C}_0(]a,b[)$,

une *solution* de l'équation $L(\mathrm{D})u = f$ sur $]a,b[$ est un élément u de $\mathrm{C}_p(]a,b[)$ tel que $L(\mathrm{D})u = f$.

En particulier, si f est la fonction 0, on dit que le problème est *homogène.*

Position du problème. Nous nous proposons de caractériser l'ensemble des solutions de l'équation $L(\mathrm{D})u = f$ sur $]a,b[$, c'est-à-dire de déterminer

$$\{\, u \in \mathrm{C}_p(]a,b[) : L(\mathrm{D})u = f \quad \text{sur} \quad]a,b[\,\}\,.$$

De la linéarité de l'opérateur de dérivation $L(\mathrm{D})$, on tire de suite que toute combinaison linéaire de solutions de l'équation homogène $L(\mathrm{D})u = 0$ est encore une solution de l'équation homogène.

Remarquons aussi que, si $u \in \mathrm{C}_\infty(]a,b[)$ est une solution de l'équation homogène $L(\mathrm{D})u = 0$, alors, pour tout opérateur de dérivation linéaire à coefficients constants $L'(\mathrm{D})$ sur $]a,b[$, $L'(\mathrm{D})u$ est aussi une solution de l'équation homogène. Cela résulte directement du fait que $L(\mathrm{D})$ et $L'(\mathrm{D})$ commutent: on a $L(\mathrm{D})(L'(\mathrm{D})u) = L'(\mathrm{D})(L(\mathrm{D})u) = L'(\mathrm{D})0 = 0$.

9.1.3 Résolution des équations homogènes

Dans ce paragraphe,

a) $]a,b[$ désigne un intervalle ouvert (borné ou non) de $\mathbb{R}$,

b) $L(\mathrm{D})$ désigne un opérateur de dérivation linéaire à coefficients constants d'ordre $p \in \mathbb{N}_0$ sur $]a,b[$ qui s'écrit $L(\mathrm{D}) = c_0 + c_1\mathrm{D} + \ldots + c_p\mathrm{D}^p$.

Le théorème fondamental suivant règle complètement la résolution des équations homogènes.

Théorème 9.1.3.1 (cas homogène général) *La solution la plus générale de l'équation homogène $L(\mathrm{D})u = 0$ sur $]a,b[$ s'écrit*

$$u(x) = P_{\alpha_1-1}(x)\,\mathrm{e}^{a_1x} + \ldots + P_{\alpha_m-1}(x)\,\mathrm{e}^{a_mx} \quad \textit{sur} \quad]a,b[\,,$$

où

1) *$a_1, \ldots, a_m$ sont les zéros distincts du polynôme caractéristique $L(z)$;*

2) *pour tout $j \leq m$, α_j est la multiplicité de a_j comme zéro de $L(z)$;*

3) *pour tout $\alpha \in \mathbb{N}$, $P_\alpha(x)$ est un polynôme de degré inférieur ou égal à α.*

Preuve. A. Procédons d'abord à l'analyse des solutions.

a) *Equation* $\mathrm{D}^pu = 0$ *sur* $]a,b[$.

Si u est une solution de l'équation homogène $\mathrm{D}^pu = 0$ sur $]a,b[$, $\mathrm{D}^{p-1}u$ appartient à $\mathrm{C}_1(]a,b[)$ et sa dérivée d'ordre 1 est identiquement nulle sur $]a,b[$. Le théorème de l'ouvert connexe affirme alors qu'il existe $C_1 \in \mathbb{C}$ tel que $\mathrm{D}^{p-1}u(x) = C_1$ pour tout $x \in]a,b[$. En répétant p fois cet argument, on obtient que u est égal à un polynôme de degré $p-1$ au plus sur $]a,b[$.

b) *Equation* $(\mathrm{D}-c)^pu = 0$ *sur* $]a,b[$ *avec* $c \in \mathbb{C}$.

Si $u \in \mathrm{C}_p(]a,b[)$ vérifie $(\mathrm{D}-c)^pu = 0$ sur $]a,b[$, posons

$$u^*(x) = \mathrm{e}^{-cx}u(x), \quad \forall x \in]a,b[\,.$$

Bien sûr, u^* appartient à $\mathrm{C}_p(]a,b[)$ et, vu le lemme du paragraphe 9.1.1.1, nous obtenons

$$0 = (\mathrm{D} - c)^p u(x) = \mathrm{e}^{cx} \cdot \mathrm{D}^p u^*(x), \quad \forall x \in \,]a,b[\,.$$

Vu ce qui précède, il existe un polynôme P_{p-1} de degré $p-1$ au plus tel que

$$u(x) = \mathrm{e}^{cx} u^*(x) = \mathrm{e}^{cx} P_{p-1}(x), \quad \forall x \in \,]a,b[\,.$$

c) *Equation* $L(\mathrm{D})u = 0$ *sur* $]a,b[$.

Avec les notations de l'énoncé, nous avons

$$L(z) = c_p \prod_{j=1}^{m} (z - a_j)^{\alpha_j} \;\text{ donc }\; L(\mathrm{D}) = c_p \prod_{j=1}^{m} (\mathrm{D} - a_j)^{\alpha_j}.$$

Par application du théorème de décomposition des fractions rationnelles en fractions rationnelles simples, il existe pour tout $j = 1, \ldots, m$ un polynôme $Q_{\alpha_j - 1}(z)$ de degré $\alpha_j - 1$ au plus tel que

$$\frac{1}{L(z)} = \sum_{j=1}^{m} \frac{Q_{\alpha_j-1}(z)}{(z-a_j)^{\alpha_j}}, \quad \forall z \in \mathbb{C} \setminus \{a_1, \ldots, a_m\}.$$

Par multiplication par $L(z)$ dans les deux membres, il vient

$$1 = c_p \sum_{j=1}^{m} Q_{\alpha_j-1}(z) \cdot (z-a_1)^{\alpha_1} \ldots [(z-a_j)^{\alpha_j}] \ldots (z-a_m)^{\alpha_m}$$

—où le facteur placé entre crochet est omis—pour tout $z \in \mathbb{C} \setminus \{a_1, \ldots, a_m\}$ et même pour tout $z \in \mathbb{C}$ si on remarque que les deux membres de cette égalité sont des polynômes. Dès lors, nous avons l'identité

$$1 = c_p \sum_{j=1}^{m} Q_{\alpha_j-1}(\mathrm{D}) \cdot (\mathrm{D}-a_1)^{\alpha_1} \ldots [(\mathrm{D}-a_j)^{\alpha_j}] \ldots (\mathrm{D}-a_m)^{\alpha_m}.$$

Ces préliminaires étant acquis, soit $u \in \mathrm{C}_p(]a,b[)$ une solution de l'équation $L(\mathrm{D})u = 0$ sur $]a,b[$. En recourant à l'identité obtenue, il vient

$$u(x) = c_p \sum_{j=1}^{m} Q_{\alpha_j-1}(\mathrm{D})\, v_j(x), \quad \forall x \in \,]a,b[\,,$$

si on pose

$$v_j = (\mathrm{D}-a_1)^{\alpha_1} \ldots [(\mathrm{D}-a_j)^{\alpha_j}] \ldots (\mathrm{D}-a_m)^{\alpha_m}\, u, \quad \forall j \le m.$$

Pour tout entier $j \leq m$, la fonction v_j appartient bien sûr à $\mathrm{C}_{\alpha_j}(]a,b[)$ et vérifie l'équation homogène
$$(\mathrm{D} - a_j)^{\alpha_j}\, v_j = 0 \quad \text{sur} \quad]a,b[\,.$$
Vu ce qui précède, il existe donc un polynôme P'_{α_j-1} de degré $\alpha_j - 1$ au plus tel que
$$v_j(x) = \mathrm{e}^{a_j x} P'_{\alpha_j-1}(x) \quad \text{sur} \quad]a,b[\,.$$

Au total, il vient
$$u(x) = c_p \sum_{j=1}^{m} Q_{\alpha_j-1}(\mathrm{D}) \left(\mathrm{e}^{a_j x} P'_{\alpha_j-1}(x) \right) = c_p \sum_{j=1}^{m} \mathrm{e}^{a_j x} Q_{\alpha_j-1}(\mathrm{D} + a_j) P'_{\alpha_j-1}(x)$$
pour tout $x \in \,]a,b[$, ce qui permet évidemment d'affirmer que u s'écrit
$$u(x) = \sum_{j=1}^{m} \mathrm{e}^{a_j x} P_{\alpha_j-1}(x), \quad \forall x \in \,]a,b[\,,$$
où, pour tout $j \leq m$, $P_{\alpha_j-1}(x)$ est un polynôme de degré $\alpha_j - 1$ au plus.

B. Pour conclure, il suffit alors de vérifier que toute fonction u s'écrivant
$$u(x) = \sum_{j=1}^{m} \mathrm{e}^{a_j x} P_{\alpha_j-1}(x) \quad \text{sur} \quad]a,b[\,,$$
est solution de l'équation homogène $L(\mathrm{D})u = 0$. Or nous avons successivement:
$$L(\mathrm{D}) \left(\sum_{j=1}^{m} \mathrm{e}^{a_j x} P_{\alpha_j-1}(x) \right) =$$
$$c_p \sum_{j=1}^{m} \mathrm{e}^{a_j x} (\mathrm{D} - a_1 + a_j)^{\alpha_1} \ldots [\mathrm{D}^{\alpha_j}] \ldots (\mathrm{D} - a_m + a_j)^{\alpha_m} \mathrm{D}^{\alpha_j} P_{\alpha_j-1}(x) = 0.$$

D'où la conclusion.∎

Définition. Chacune des fonctions exponentielle-polynôme
$$x^l\, \mathrm{e}^{a_j x} \quad \text{pour} \quad l = 0, 1, \ldots, \alpha_j - 1 \quad \text{et} \quad j = 1, \ldots, m \qquad (*)$$
est solution de l'équation homogène $L(\mathrm{D})u = 0$ sur $]a,b[$ et toute solution de cette équation homogène est combinaison linéaire de ces fonctions $(*)$, appelées *solutions fondamentales* de l'opérateur $L(\mathrm{D})$. Cela étant, comme on a $\alpha_1 + \ldots + \alpha_m = p$, $L(\mathrm{D})$ a p solutions fondamentales et la solution la plus générale de $L(\mathrm{D})u = 0$ sur $]a,b[$ dépend de p constantes.

Théorème 9.1.3.2 (cas réel homogène) *Si $L(\mathrm{D})$ est un opérateur de dérivation linéaire à coefficients constants et réels et si a_j est un zéro α_j-uple de $L(z)$ tel que $\Im a_j \neq 0$, alors $\overline{a_j}$ est aussi un zéro α_j-uple de $L(z)$ et on peut remplacer*

$$P_{\alpha_j-1}(x)\,\mathrm{e}^{a_j x} + P'_{\alpha_j-1}(x)\,\mathrm{e}^{\overline{a_j} x}$$

par

$$\mathrm{e}^{\Re a_j\, x}\,\Big(P''_{\alpha_j-1}(x)\,\cos(\Im a_j\, x) + P'''_{\alpha_j-1}(x)\,\sin(\Im a_j\, x)\Big)$$

dans l'expression de la solution la plus générale de l'équation homogène.

Preuve. Il suffit de noter que, pour tous a, c, $d \in \mathbb{C}$ et tout $l \in \mathbb{N}$, on a

$$c\,x^l \mathrm{e}^{ax} + d\,x^l \mathrm{e}^{\bar{a}x} = \mathrm{e}^{\Re a\,x}\,\big((c+d)x^l\cos(\Im a\,x) + i\,(c-d)x^l\sin(\Im a\,x)\big)\,,$$

et

$$\mathrm{e}^{\Re a\,x}\,\big(c\,x^l\cos(\Im a\,x) + d\,x^l\sin(\Im a\,x)\big) = \frac{c-id}{2}x^l\mathrm{e}^{ax} + \frac{c+id}{2}x^l\mathrm{e}^{\bar{a}x}.\blacksquare$$

Corollaire 9.1.3.3 *Toute solution de l'équation homogène $L(\mathrm{D})u = 0$ sur $]a,b[$ appartient à $\mathrm{C}_\infty(]a,b[)$.*

Preuve. Cela résulte aissitôt de la forme de la solution la plus générale de cette équation.∎

Théorème 9.1.3.4 *Les solutions fondamentales de $L(\mathrm{D})$ sont linéairement indépendantes.*

Plus précisément, *si on note $u_1, \ldots, u_p$ les solutions fondamentales de $L(\mathrm{D})$, alors, pour tout $x_0 \in]a,b[$ et tous $z_0, \ldots, z_{p-1} \in \mathbb{C}$, il existe des nombres complexes uniques $d_1, \ldots, d_p$ tels que la solution $u = \sum_{k=1}^p d_k u_k$ de l'équation homogène $L(\mathrm{D})u = 0$ sur $]a,b[$ vérifie $[\mathrm{D}^k u]_{x_0} = z_k$ pour tout $k = 0, \ldots, p-1$.*

Preuve. Comme

$$\begin{cases} \displaystyle\sum_{k=1}^p d_k\; u_k(x_0) & = z_0 \qquad (= u(x_0)) \\ \displaystyle\sum_{k=1}^p d_k\; [\mathrm{D}u_k]_{x_0} & = z_1 \qquad (= [\mathrm{D}u]_{x_0}) \\ \vdots & \\ \displaystyle\sum_{k=1}^p d_k\; [\mathrm{D}^{p-1}u_k]_{x_0} & = z_{p-1} \qquad (= [\mathrm{D}^{p-1}u]_{x_0}) \end{cases}$$

est un système de p équations linéaires à p inconnues, il suffit d'établir le cas précisé uniquement lorsque $z_0 = z_1 = \ldots = z_{p-1} = 0$.

Pour conclure il suffit donc d'établir que, si la solution

$$u(x) = \sum_{k=1}^{m} \mathrm{e}^{a_k x} P_{\alpha_k - 1}(x)$$

de l'équation homogène vérifie

$$[\mathrm{D}^k u]_{x_0} = 0 \quad \text{pour tout} \quad k = 0, \ldots, p-1, \qquad (*)$$

alors chacun des polynômes $P_{\alpha_k - 1}$ est identiquement nul.

Fixons un entier $j \leq m$ et considérons

$$\begin{aligned} v_j(x) &= c_p(\mathrm{D} - a_1)^{\alpha_1} \ldots [(\mathrm{D} - a_j)^{\alpha_j}] \ldots (\mathrm{D} - a_m)^{\alpha_m}\ u(x) \\ &= c_p(\mathrm{D} - a_1)^{\alpha_1} \ldots [(\mathrm{D} - a_j)^{\alpha_j}] \ldots (\mathrm{D} - a_m)^{\alpha_m} \left(\mathrm{e}^{a_j x} P_{\alpha_j - 1}(x)\right) \\ &= c_p \mathrm{e}^{a_j x}(\mathrm{D} + a_j - a_1)^{\alpha_1} \ldots [\mathrm{D}^{\alpha_j}] \ldots (\mathrm{D} + a_j - a_m)^{\alpha_m}\ P_{\alpha_j - 1}(x). \end{aligned}$$

La première égalité montre que v_j est une combinaison linéaire de u et de ses dérivées jusqu'à l'ordre $p - \alpha_j$; on en déduit aussitôt les égalités

$$v_j(x_0) = [\mathrm{D} v_j]_{x_0} = \ldots = [\mathrm{D}^{\alpha_j - 1} v_j]_{x_0} = 0.$$

Déduisons-en que le polynôme $P_{\alpha_j - 1}$ a tous ses coefficients nuls, ce qui permet de conclure. Soit $c\,x^l$ le terme de plus haut degré de $P_{\alpha_j - 1}(x)$. On a d'une part $[(\mathrm{D} - a_j)^l\, v_j]_{x_0} = 0$ et d'autre part

$$\begin{aligned} &(\mathrm{D} - a_j)^l\, v_j(x) \\ &\quad = c_p \mathrm{e}^{a_j x}(\mathrm{D} + a_j - a_1)^{\alpha_1} \ldots [\mathrm{D}^{\alpha_j}] \ldots (\mathrm{D} + a_j - a_m)^{\alpha_m}\ \mathrm{D}^l P_{\alpha_j - 1}(x) \\ &\quad = c_p \mathrm{e}^{a_j x}(\mathrm{D} + a_j - a_1)^{\alpha_1} \ldots [\mathrm{D}^{\alpha_j}] \ldots (\mathrm{D} + a_j - a_m)^{\alpha_m}\ c\, l! \end{aligned}$$

c'est-à-dire

$$0 = c_p \mathrm{e}^{a_j x_0}(a_j - a_1)^{\alpha_1} \ldots [(a_j - a_j)^{\alpha_j}] \ldots (a_j - a_m)^{\alpha_m}\ c\, l!,$$

ce qui suffit.∎

Exemple. *Si $a \neq 0$ et b sont des nombres complexes, déterminer la solution la plus générale de l'équation $(a\mathrm{D} + b)\, u = 0$ sur $\mathbb{R}$.*

Bien sûr le polynôme caractéristique $L(z) = az + b$ admet $-b/a$ comme seul zéro. Dès lors, la solution la plus générale de cette équation homogène s'écrit $c\,\mathrm{e}^{-bx/a}$ avec $c \in \mathbb{C}$.

Exemple. *Si $a \neq 0$, b et c sont des nombres complexes, déterminer la solution la plus générale de l'équation $(a\mathrm{D}^2 + b\mathrm{D} + c)\, u = 0$ sur $\mathbb{R}$.*

Si on a $b^2 - 4ac = 0$, le polynôme caractéristique $L(z) = az^2 + bz + c$ de cet opérateur admet $-b/(2a)$ comme zéro double et n'a pas d'autre zéro. La solution la plus générale s'écrit donc

$$(c_1 + c_2 x)\, \mathrm{e}^{-\frac{b}{2a}x} \quad \text{avec} \quad c_1, c_2 \in \mathbb{C}.$$

Si on a $b^2 - 4ac \neq 0$, le polynôme caractéristique admet

$$\frac{-\, b \pm \sqrt{b^2 - 4ac}}{2a}$$

comme zéros simples et la solution la plus générale s'écrit

$$\mathrm{e}^{-\frac{b}{2a}x}(c_1\, \mathrm{e}^{\frac{\sqrt{b^2-4ac}}{2a}x} + c_2\, \mathrm{e}^{-\frac{\sqrt{b^2-4ac}}{2a}x}) \quad \text{avec} \quad c_1, c_2 \in \mathbb{C}.$$

Si a, b, c sont des nombres réels tels que $b^2 - 4ac < 0$, la solution la plus générale peut aussi s'écrire

$$\mathrm{e}^{-\frac{b}{2a}x}\left(c_1\, \cos(\frac{\sqrt{4ac - b^2}}{2a}\, x) + c_2\, \sin(\frac{\sqrt{4ac - b^2}}{2a}\, x)\right) \quad \text{avec} \quad c_1, c_2 \in \mathbb{C}.$$

9.1.4 Généralités sur les équations non homogènes

La structure de la solution la plus générale d'une équation différentielle linéaire à coefficients constants et non homogène est régie par le théorème suivant.

Théorème 9.1.4.1 *La solution la plus générale d'une équation différentielle linéaire à coefficients constants et non homogène $L(\mathrm{D})u = f$ sur $]a, b[$ est la somme d'une solution particulière de $L(\mathrm{D})u = f$ sur $]a, b[$ et de la solution la plus générale de l'équation homogène $L(\mathrm{D})u = 0$ sur $]a, b[$.*

Preuve. D'une part, si u_0 est solution particulière de $L(\mathrm{D})u = f$ sur $]a, b[$ et si u est solution de $L(\mathrm{D})u = 0$ sur $]a, b[$, bien sûr $u_0 + u$ appartient à $\mathrm{C}_p(]a, b[)$ et est tel que $L(\mathrm{D})(u_0 + u) = f$ sur $]a, b[$.

D'autre part, si u_0 et u_1 sont des solutions de $L(\mathrm{D})u = f$ sur $]a, b[$, on a évidemment $u_0 - u_1 \in \mathrm{C}_p(]a, b[)$ et $L(\mathrm{D})(u_0 - u_1) = 0$ sur $]a, b[$. ∎

Les conséquences de ce théorème sont importantes.

Conséquences. a) La solution la plus générale de l'équation non homogène $L(\mathrm{D})(u) = f$ sur $]a, b[$ dépend de p paramètres, tout comme celle de l'équation homogène correspondante.

b) Si l'équation non homogène $L(\mathrm{D})(u) = f$ sur $]a, b[$ admet une solution particulière u_0—nous allons voir au paragraphe suivant que c'est toujours le cas si f appartient à $\mathrm{C}_0(]a, b[)$—, si x_0 appartient à $]a, b[$ et si $z_0, \ldots, z_{p-1}$ sont des nombres complexes, il existe une solution unique u_1 de l'équation $L(\mathrm{D})(u) = f$ sur $]a, b[$ telle que

$$u_1(x_0) = z_0 \text{ et } [\mathrm{D}^k u_1]_{x_0} = z_k, \quad \forall k \leq p - 1. \qquad (*)$$

De fait, d'une part, il existe une solution u_1' de l'équation homogène $L(\mathrm{D})(u) = 0$ sur $]a, b[$ telle que

$$u_1'(x_0) = z_0 - u_0(x_0) \text{ et } [\mathrm{D}^k u_1']_{x_0} = z_k - [\mathrm{D}^k u_0]_{x_0}, \quad \forall k \leq p - 1.$$

On vérifie alors de suite que $u_1 = u_0 + u_1'$ est une solution de l'équation non homogène, qui vérifie les égalités $(*)$.

D'autre part, cette solution est unique car si u_1 et u_2 conviennent, $u_1 - u_2$ est une solution du problème homogène, qui s'annule ainsi que ses $p - 1$ premières dérivées en x_0.

c) Tout revient donc à déterminer une solution particulière de l'équation non homogène (et cela de la manière la plus simple possible).

Remarquons à cet effet que la recherche d'une solution particulière de l'équation $L(\mathrm{D})(u) = \sum_{j=1}^{J} f_j$ sur $]a, b[$ peut s'effectuer en recherchant, pour tout $j \leq J$, une solution particulière de chacune des équations $L(\mathrm{D})(u) = f_j$ sur $]a, b[$ et en en prenant la somme.

Remarquons aussi que, si $L(\mathrm{D})$ est réel et si f est la partie réelle (resp. imaginaire) de $F \in \mathrm{C}_m(]a, b[)$, une solution particulière de $L(\mathrm{D})(u) = f$ sur $]a, b[$ est donnée par la partie réelle (resp. imaginaire) d'une solution particulière de l'équation $L(\mathrm{D})(u) = F$ sur $]a, b[$.

9.1.5 Méthode de la variation des constantes

La méthode d'obtention d'une solution particulière de l'équation non homogène $L(\mathrm{D})u = f$ sur $]a, b[$ que nous allons d'écrire ci-dessous a une double importance:

a) elle établit que toute équation non homogène admet au moins une solution particulière;

b) elle donne une méthode de calcul effectif d'une solution particulière—on n'est donc limité que par la possibilité de mener à bien ces calculs.

Théorème 9.1.5.1 (Variation des constantes) *Soit*

$$L(\mathrm{D}) = c_0 + c_1\mathrm{D} + \ldots + c_p\mathrm{D}^p$$

un opérateur de dérivation linéaire à coefficients constants d'ordre $p \in \mathbb{N}_0$, soit $]a,b[$ un intervalle ouvert de $\mathbb{R}$ et soit f un élément de $\mathrm{C}_l(]a,b[)$.

a) *On met en évidence les p constantes dont dépend la solution la plus générale de l'équation homogène $L(\mathrm{D})u = 0$ sur $]a,b[$ en l'écrivant $u = \sum_{k=1}^{p} C_k\, u_k$ où u_1, ..., u_p sont les solutions fondamentales de $L(\mathrm{D})$.*

b) *On forme le système d'équations*

$$\left\{\begin{array}{lcl} \displaystyle\sum_{k=1}^{p} C_k\, u_k & = & 0 \\ \displaystyle\sum_{k=1}^{p} C_k\, \mathrm{D}u_k & = & 0 \\ \ldots & & \\ \displaystyle\sum_{k=1}^{p} C_k\, \mathrm{D}^{p-2}u_k & = & 0 \\ \displaystyle\sum_{k=1}^{p} C_k\, \mathrm{D}^{p-1}u_k & = & \dfrac{f}{c_p} \end{array}\right.$$

c) *Pour tout $x \in\,]a,b[$, ce système admet une solution unique $C_1(x)$, ..., $C_p(x)$ et $C_1(x)$, ..., $C_p(x)$ sont p fonctions sur $]a,b[$ qui appartiennent en fait à $\mathrm{C}_l(]a,b[)$.*

d) *La fonction*

$$u(x) = \sum_{k=1}^{p} \left(\int C_k(x)\, dx \right) \cdot u_k(x)$$

est une solution particulière de $L(\mathrm{D})u = f$ et appartient à $\mathrm{C}_{l+p}(]a,b[)$.

Preuve. a) et b) ne posent aucun problème. Pour a), il suffit de résoudre $L(\mathrm{D})u = 0$. Pour b), on note que c_p diffère de 0.

c) Pour tout $x \in\,]a,b[$, nous avons vu en effet qu'il existe des nombres uniques $C_1, \ldots, C_p$ tels que la solution $u = C_1\, u_1 + \ldots + C_p\, u_p$ de l'équation homogène vérifie les égalités

$$u(x) = [\mathrm{D}u]_x = \ldots = [\mathrm{D}^{p-2}u]_x = 0 \;\text{ et }\; [\mathrm{D}^{p-1}u]_x = \frac{f(x)}{c_p}.$$

De plus, la formule de Cramer donne de suite l'appartenance de $C_1(x)$, ..., $C_p(x)$ à $\mathrm{C}_l(]a,b[)$ car chacun des $u_1(x)$, ..., $u_p(x)$ appartient à $\mathrm{C}_\infty(]a,b[)$.

d) Posons

$$u(x) = \sum_{k=1}^{p} \left(\int C_k(x)\, dx \right) \cdot u_k(x)$$

On vérifie aisément, en recourant notamment à la définition des fonctions $C_k(x)$, que les égalités suivantes ont lieu

$$\begin{aligned}
\mathrm{D}u(x) &= \sum_{k=1}^{p} C_k(x)\cdot u_k(x) + \sum_{k=1}^{p}\left(\int C_k(x)\,dx\right)\cdot \mathrm{D}u_k(x)\\
&= \sum_{k=1}^{p}\left(\int C_k(x)\,dx\right)\cdot \mathrm{D}u_k(x)\\
&\vdots\\
\mathrm{D}^{p-1}u(x) &= \sum_{k=1}^{p} C_k(x)\cdot \mathrm{D}^{p-2}u_k(x) + \sum_{k=1}^{p}\left(\int C_k(x)\,dx\right)\cdot \mathrm{D}^{p-1}u_k(x)\\
&= \sum_{k=1}^{p}\left(\int C_k(x)\,dx\right)\cdot \mathrm{D}^{p-1}u_k(x)\\
\mathrm{D}^{p}u(x) &= \sum_{k=1}^{p} C_k(x)\cdot \mathrm{D}^{p-1}u_k(x) + \sum_{k=1}^{p}\left(\int C_k(x)\,dx\right)\cdot \mathrm{D}^{p}u_k(x)\\
&= \frac{f}{c_p} + \sum_{k=1}^{p}\left(\int C_k(x)\,dx\right)\cdot \mathrm{D}^{p}u_k(x).
\end{aligned}$$

On en tire de suite que $L(\mathrm{D})u = c_0\,u + c_1\,\mathrm{D}u + \ldots + c_p\,\mathrm{D}^p u$ est égal à

$$f + \sum_{k=1}^{p}\left(\int C_k(x)\,dx\right)\cdot L(\mathrm{D})u_k = f.$$

De plus, u appartient à $\mathrm{C}_{l+p}(]a,b[)$: on a pu en effet dériver p fois la fonction u et cette dérivée $\mathrm{D}^p u$ appartient à $\mathrm{C}_l(]a,b[)$ comme on le vérifie de suite.∎

Remarque. Pour un lecteur déjà au courant de la théorie de l'intégration, précisons que, si x_0 appartient à $]a,b[$,

$$u(x) = \sum_{k=1}^{p}\left(\int_{x_0}^{x} C_k(t)\,dt\right)\cdot u_k(x)$$

est la solution particulière de l'équation $L(\mathrm{D})u = f$ qui s'annule en x_0 ainsi que ses $p-1$ premières dérivées. La vérification est immédiate.

Exemple. *Résoudre l'équation* $\mathrm{D}u + u = \mathrm{e}^{\mathrm{e}^x}$ *sur* $\mathbb{R}$.

a) L'équation homogène $\mathrm{D}u + u = 0$ sur $\mathbb{R}$ admet bien sûr $C\mathrm{e}^{-x}$ avec $C \in \mathbb{C}$ comme solution la plus générale.

b) Appliquons la méthode de la variation des constantes. L'équation $C\mathrm{e}^{-x} = \mathrm{e}^{\mathrm{e}^x}$ admet évidemment $C(x) = \mathrm{e}^x\mathrm{e}^{\mathrm{e}^x}$ comme solution. Dès lors,

$$u(x) = \left(\int \mathrm{e}^x\mathrm{e}^{\mathrm{e}^x}\,dx\right)\mathrm{e}^{-x}$$

est une solution particulière de l'équation non homogène.

c) La solution la plus générale du problème posé s'écrit donc

$$\mathrm{e}^{-x}\mathrm{e}^{\mathrm{e}^x} + C\mathrm{e}^{-x} \quad \text{avec} \quad C \in \mathbb{C}.$$

Exemple. *Etant donné* $\omega \in \left]0, +\infty\right[$, *résoudre l'équation*

$$(\mathrm{D}^2 + \omega^2)u = \frac{1}{\cos(\omega x)} \quad \text{sur} \quad \left]-\pi/(2\omega), \pi/(2\omega)\right[.$$

Bien sûr, le second membre est continu sur l'intervalle considéré.

a) Résolvons l'équation homogène.

L'équation caractéristique s'écrit $z^2 + \omega^2 = 0$ et a donc $z = i\omega$ et $z = -i\omega$ pour seuls zéros, ces zéros étant différents et de multiplicité 1. Dès lors, la solution la plus générale de l'équation homogène s'écrit

$$C_1\mathrm{e}^{i\omega x} + C_2\mathrm{e}^{-i\omega x} \quad \text{avec} \quad C_1, C_2 \in \mathbb{C},$$

ou

$$C_1\cos(\omega x) + C_2\sin(\omega x) \quad \text{avec} \quad C_1, C_2 \in \mathbb{C}.$$

b) Appliquons la méthode de la variation des constantes.

Vu la forme du second membre, nous allons utiliser la deuxième forme de la solution la plus générale de l'équation homogène. Résolvons le système

$$\left\{\begin{array}{rcl} C_1\cos(\omega x) + C_2\sin(\omega x) & = & 0 \\ -\,C_1\omega\sin(\omega x) + C_2\omega\cos(\omega x) & = & \dfrac{1}{\cos(\omega x)}. \end{array}\right.$$

Si on multiplie la première équation par $\omega\sin(\omega x)$ et la seconde par $\cos(\omega x)$ et si on somme les résultats obtenus, il vient $C_2\omega = 1$. Il s'ensuit que la solution de ce système est donnée par

$$C_1(x) = -\frac{1}{\omega}\,\mathrm{tg}(\omega x) \quad \text{et} \quad C_2(x) = \frac{1}{\omega} \quad \text{sur} \quad \left]-\pi/(2\omega), \pi/(2\omega)\right[.$$

Une solution particulière est donc donnée par

$$\begin{aligned} & -\frac{1}{\omega}\cos(\omega x)\int \mathrm{tg}(\omega x)\,dx + \frac{1}{\omega}\sin(\omega x)\int dx \\ \simeq\ & \frac{1}{\omega^2}\ln\left(\cos(\omega x)\right)\cdot\cos(\omega x) + \frac{x}{\omega}\cdot\sin(\omega x). \end{aligned}$$

c) Cela étant,la solution la plus générale du problème posé s'écrit

$$\frac{1}{\omega^2}\ln\left(\cos(\omega x)\right)\cdot\cos(\omega x)+\frac{x}{\omega}\cdot\sin(\omega x)+C_1\cos(\omega x)+C_2\sin(\omega x)$$

avec $C_1, C_2 \in \mathbb{C}$.

9.1.6 Méthode des exponentielles-polynômes

La méthode d'obtention d'une solution particulière de l'équation non homogène $L(\mathrm{D})u = f$ qui va suivre ne s'applique que si f est une fonction exponentielle-polynôme, c'est-à-dire une fonction du type $f(x) = \mathrm{e}^{cx}P(x)$ où c est un nombre complexe et $P(x)$ un polynôme. Elle s'étend bien sûr par linéarité, aux combinaisons linéaires de telles fonctions.

Proposition 9.1.6.1 *Si $L(\mathrm{D})$ est un opérateur de dérivation linéaire à coefficients constants d'ordre $p \in \mathbb{N}_0$ sur $]a, b[$, si c est un nombre complexe et si $P(x)$ est un polynôme de degré q, l'équation*

$$L(\mathrm{D})u(x) = \mathrm{e}^{cx}P(x) \quad \textit{sur} \quad]a, b[$$

admet une solution de la forme $\mathrm{e}^{cx}\mathcal{P}_q(x)x^\gamma$ où $\mathcal{P}_q(x)$ est un polynôme de degré égal à q et où γ est la multiplicité de c comme zéro de $L(z)$.

Preuve. La méthode de la variation des constantes permet d'affirmer qu'il existe au moins une solution particulière, soit u_0. Cela étant de

$$L(\mathrm{D})u_0(x) = \mathrm{e}^{cx}P(x) \quad \text{sur} \quad]a, b[\,,$$

on tire

$$\begin{aligned}(\mathrm{D}-c)^{q+1}L(\mathrm{D})u_0(x) &= (\mathrm{D}-c)^{q+1}\left(\mathrm{e}^{cx}P(x)\right)\\ &= \mathrm{e}^{cx}\mathrm{D}^{q+1}P(x) = 0 \quad \text{sur} \quad]a, b[\,.\end{aligned}$$

Il s'ensuit que u_0 est solution de l'équation homogène

$$(\mathrm{D}-c)^{q+1}L(\mathrm{D})u = 0 \quad \text{sur} \quad]a, b[$$

et s'écrit donc

$$u_0(x) = \mathrm{e}^{a_1x}P_{\alpha_1-1}(x) + \ldots + \mathrm{e}^{a_lx}P_{\alpha_l-1}(x) + \mathrm{e}^{cx}P_{q+\gamma}(x)$$

si $a_1, \ldots, a_l$ sont les zéros de $L(z)$ différents de c et de multiplicité $\alpha_1, \ldots, \alpha_l$ et si c est un zéro de $L(z)$ de multiplicité γ éventuellement nulle.

Cependant, on peut éliminer automatiquement une partie des termes de cette somme. En effet, on peut aussi la mettre sous la forme

$$[\mathrm{e}^{a_1 x} P_{\alpha_1 - 1}(x) + \ldots + \mathrm{e}^{a_l x} P_{\alpha_l - 1}(x) + \mathrm{e}^{cx} P_{\gamma-1}(x)] + \mathrm{e}^{cx} x^{\gamma} \mathcal{P}_q(x)$$

où le terme entre crochets est annulé par l'opérateur $L(\mathrm{D})$. Il s'ensuit que la fonction $\mathrm{e}^{cx} x^{\gamma} \mathcal{P}_q(x)$ est aussi une solution particulière de $L(\mathrm{D})u = f$, ce qui suffit.∎

Cas d'application. La méthode des exponentielles-polynômes s'applique de façon directe dans les cas suivants:
a) si $f(x)$ est un polynôme: cela revient à considérer $c = 0$;
b) si $f(x)$ est une combinaison linéaire d'exponentielles, donc en particulier si $f(x)$ est un polynôme en $\mathrm{ch}(mx)$, $\mathrm{sh}(mx)$, $\cos(mx)$, $\sin(mx)$ avec $m \in \mathbb{N}$.

Ainsi la méthode s'applique notamment si

$$f(x) = \mathrm{e}^{\alpha x} \left\{ \begin{array}{c} \cos(\beta x) \\ \sin(\beta x) \end{array} \right\} \cdot P(x), \quad \text{avec} \quad \alpha, \beta \in \mathbb{C}.$$

Remarque. Si $L(\mathrm{D})$ et $P(x)$ sont réels, notons qu'une solution particulière de l'équation

$$L(\mathrm{D})u(x) = \mathrm{e}^{\alpha x} \cos(\beta x) \cdot P(x) \quad (\text{resp. } \mathrm{e}^{\alpha x} \sin(\beta x) \cdot P(x))$$

sur $\mathbb{R}$ avec $\alpha, \beta \in \mathbb{R}$, est donnée par la partie réelle (resp. imaginaire) d'une solution particulière de l'équation

$$L(\mathrm{D})u(x) = \mathrm{e}^{(\alpha + i\beta)x} P(x).$$

Méthode. Pour résoudre l'équation différentielle $L(\mathrm{D})u(x) = \mathrm{e}^{cx} P(x)$,
a) on détermine la solution la plus générale de l'équation homogène;
b) tout revient alors à déterminer une solution particulière de l'équation non homogène, c'est-à-dire à trouver γ et $\mathcal{P}_q(x)$. Pour γ, c'est immédiat. Pour $\mathcal{P}_q(x)$, on recourt à la méthode d'identification des coefficients. La disposition pratique des calculs qui découle du lemme suivant permet d'alléger au maximum le calcul de $L(\mathrm{D})(\mathrm{e}^{cx} \mathcal{P}_q(x) x^{\gamma})$. Ce lemme précise en fait celui du paragraphe 9.1.1.1.

Lemme 9.1.6.2 *Si l'opérateur de dérivation linéaire à coefficients constants $L(\mathrm{D})$ est d'ordre $p \in \mathbb{N}_0$ et si f appartient à $\mathrm{C}_p(]a, b[)$, on a*

$$L(\mathrm{D})(\mathrm{e}^{cx} f(x)) = \mathrm{e}^{cx} \sum_{j=0}^{p} \frac{1}{j!} \mathrm{D}^j f(x) \cdot [\mathrm{D}_z^j L(z)]_{z=c}$$

pour tout $c \in \mathbb{C}$.

Preuve. De fait, si $L(\mathrm{D})$ s'écrit $\sum_{l=0}^{p} c_l \mathrm{D}^l$, il vient successivement

$$\begin{aligned}
\sum_{l=0}^{p} c_l \mathrm{D}^l(\mathrm{e}^{cx} f(x)) &\underset{(*)}{=} \sum_{l=0}^{p} c_l \sum_{j=0}^{l} \mathrm{C}_l^j \mathrm{D}^{l-j} \mathrm{e}^{cx} \cdot \mathrm{D}^j f(x) \\
&= \mathrm{e}^{cx} \sum_{l=0}^{p} c_l \sum_{j=0}^{l} \frac{l!}{j!(l-j)!} c^{l-j} \cdot \mathrm{D}^j f(x) \\
&\underset{(**)}{=} \mathrm{e}^{cx} \sum_{j=0}^{p} \frac{1}{j!} \left(\sum_{l=j}^{p} c_l \frac{l!}{(l-j)!} c^{l-j} \right) \cdot \mathrm{D}^j f(x)
\end{aligned}$$

en recourant à la formule de Leibniz en $(*)$ et en permutant les deux sommes en $(**)$. D'où la conclusion car on vérifie aisément que, pour $j \leq p$, il vient

$$\mathrm{D}_z^j L(z) = \sum_{l=j}^{p} c_l \frac{l!}{(l-j)!} z^{l-j}. \blacksquare$$

Disposition pratique. Cela étant, pour obtenir $L(\mathrm{D})(\mathrm{e}^{cx} x^{\gamma} \mathcal{P}_q(x))$, il est conseillé de construire le tableau suivant:

j	$\mathrm{D}_z^j L(z)$	$[\mathrm{D}_z^j L(z)]_{z=c}$	$\mathrm{D}^j(\mathcal{P}_q(x) x^{\gamma})$	$1/j!$
0	$L(z)$	$L(c)$	$\cdot$	1
1	$\mathrm{D}_z L(z)$	$\cdot$	$\cdot$	1
2	$\mathrm{D}_z^2 L(z)$	$\cdot$	$\cdot$	1/2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Exemple. *Résoudre l'équation* $\mathrm{D}u + u = \cos(x)$ *sur* $\mathbb{R}$.

Comme on a $\cos \in \mathrm{C}_\infty(\mathbb{R})$, le problème a un sens.

a) Résolvons d'abord l'équation homogène. Bien sûr, -1 est le seul zéro de $L(z)$ et sa multiplicité est égale à 1. La solution la plus générale de l'équation homogène $L(\mathrm{D})u = 0$ s'écrit donc $C\mathrm{e}^{-x}$ avec $C \in \mathbb{C}$.

b) Recherchons une solution particulière de l'équation proposée.

Comme $L(\mathrm{D})$ est réel et comme $\cos(x)$ est la partie réelle de la fonction $\mathrm{e}^{ix} \in \mathrm{C}_\infty(\mathbb{R})$, on peut prendre comme solution particulière u_0, la partie réelle d'une solution particulière de l'équation $\mathrm{D}v + v = \mathrm{e}^{ix}$. Or il existe une telle solution particulière de la forme $\mathrm{e}^{ix} x^{\gamma} \mathcal{P}_m(x)$, c'est-à-dire de la forme $A\mathrm{e}^{ix}$ avec $A \in \mathbb{C}$. Du tableau suivant

j	$\mathrm{D}_z^j L(z)$	$[\mathrm{D}_z^j L(z)]_{z=i}$	$\mathrm{D}^j A$	$1/j!$
0	$z+1$	$1+i$	A	1
1	1	1	0	1

on déduit de suite que A doit vérifier l'égalité $(1+i)A = 1$, c'est-à-dire que v_0 est donné par $v_0(x) = (1-i)\mathrm{e}^{ix}/2$. Par conséquent, il vient

$$u_0(x) = \Re\left(\frac{1-i}{2}\mathrm{e}^{ix}\right) = \frac{\cos(x)+\sin(x)}{2}.$$

c) Au total, la solution la plus générale de l'équation proposée s'écrit

$$u(x) = \frac{\cos(x)+\sin(x)}{2} + C\mathrm{e}^{-x} \quad \text{avec} \quad C \in \mathbb{C}.$$

Exemple. *Résoudre l'équation* $\mathrm{D}^3u - u = x(1+\mathrm{e}^x)$ *sur* $\mathbb{R}$.
Comme on a $x(1+\mathrm{e}^x) \in \mathrm{C}_\infty(\mathbb{R})$, le problème a un sens.
a) Résolvons d'abord l'équation homogène. Bien sûr,

$$1, \frac{-1+i\sqrt{3}}{2} \quad \text{et} \quad \frac{-1-i\sqrt{3}}{2}$$

sont les seuls zéros de $L(z)$ et leur multiplicité est chaque fois égale à 1. Dès lors, la solution la plus générale de l'équation homogène peut s'écrire:

$$C_1\mathrm{e}^x + \mathrm{e}^{-x/2}\left(C_2\cos(\frac{\sqrt{3}}{2}x) + C_3\sin(\frac{\sqrt{3}}{2}x)\right) \quad \text{avec} \quad C_1, C_2, C_3 \in \mathbb{C}.$$

b) On remarque directement que $u(x) = -x$ est une solution particulière de l'équation $\mathrm{D}^3u - u = x$ sur $\mathbb{R}$.

c) Recherchons une solution particulière de l'équation $\mathrm{D}^3u - u = x\mathrm{e}^x$ sur $\mathbb{R}$. Comme le second membre est une exponentielle-polynôme, il en existe une solution particulière de la forme $\mathrm{e}^x x^\gamma \mathcal{P}_m(x)$, c'est-à-dire qu'elle peut s'écrire $\mathrm{e}^x x(Ax+B)$ avec A, $B \in \mathbb{C}$. Du tableau ci-dessous, on déduit de suite que A et B doivent vérifier l'identité $6Ax + 3B + 6A = x$; on a donc $A = 1/6$ et $B = -1/3$.

j	$\mathrm{D}_z^j L(z)$	$[\mathrm{D}_z^j L(z)]_{z=c}$	$\mathrm{D}^j(Ax^2+Bx)$	$1/j!$
0	z^3-1	0	Ax^2+Bx	1
1	$3z^2$	3	$2Ax+B$	1
2	$6z$	6	$2A$	1/2

d) Au total, la solution la plus générale de l'équation à résoudre s'écrit:

$$-x + \left(\frac{x^2}{6} - \frac{x}{3}\right)\mathrm{e}^x + C_1\mathrm{e}^x + \mathrm{e}^{-x/2}\left(C_2\cos(\frac{\sqrt{3}}{2}x) + C_3\sin(\frac{\sqrt{3}}{2}x)\right)$$

avec C_1, C_2, $C_3 \in \mathbb{C}$.

9.1.7 Equations d'Euler

Définition. Une *équation d'Euler* est une équation différentielle du type

$$c_p(x\mathrm{D})^p u + \ldots + c_1(x\mathrm{D})u + c_0 u = f$$

où p appartient à $\mathbb{N}_0$, où c_0, c_1, …, c_p sont des nombres complexes tels que $c_p \neq 0$ et où f appartient à $\mathrm{C}_l(]a, b[)$ avec $l \in \mathbb{N}$ et $]a, b[$ un intervalle ouvert inclus dans $]0, +\infty[$. Ses solutions sont les fonctions $u \in \mathrm{C}_p(]a, b[)$ qui vérifient cette équation.

Remarques. a) Toute équation différentielle du type

$$d_p x^p \mathrm{D}^p u + \ldots + d_1 x\mathrm{D}u + d_0 u = f$$

où p appartient à $\mathbb{N}_0$, où d_0, d_1, …, d_p sont des nombres complexes tels que $d_p \neq 0$ et où f appartient à $\mathrm{C}_l(]a, b[)$ avec $l \in \mathbb{N}$ et $]a, b[$ un intervalle ouvert inclus dans $]0, +\infty[$, est en fait une équation d'Euler et inversement. Pour le vérifier, il suffit de prouver que, pour tout $m \geq 1$, il existe des nombres réels $r_{m,1}$, …, $r_{m,m} = 1$ tels que

$$(x\mathrm{D})^m = \sum_{j=1}^{m} r_{m,j} x^j \mathrm{D}^j.$$

Or, pour $m = 1$, c'est trivial et, si c'est vrai pour $m = 1$, …, k, il vient

$$(x\mathrm{D})^{k+1} = x\mathrm{D}\left(\sum_{j=1}^{k} r_{k,j} x^j \mathrm{D}^j\right) = \sum_{j=1}^{k} \left(j r_{k,j} x^j \mathrm{D}^j + r_{k,j} x^{j+1} \mathrm{D}^{j+1}\right),$$

ce qui suffit.

b) Bien sûr, une équation d'Euler est essentiellement une équation différentielle linéaire à coefficients *non* constants. Cependant le changement de variable d'Euler ramène sa résolution à celle d'une équation différentielle linéaire à coefficients constants.

Proposition 9.1.7.1 *La solution la plus générale de l'équation d'Euler*

$$c_p(x\mathrm{D})^p u + \ldots + c_1(x\mathrm{D})u + c_0 u = f \quad \textit{avec} \quad f \in \mathrm{C}_l(]a, b[)$$

s'écrit $u(x) = v(\ln(x))$ *où* v *est la solution la plus générale de l'équation différentielle linéaire à coefficients constants*

$$c_p \mathrm{D}^p v + \ldots + c_1 \mathrm{D}v + c_0 v = g$$

où $g \in \mathrm{C}_l(]\ln(a), \ln(b)[)$ *est défini par* $g(y) = f(\mathrm{e}^y)$.

Preuve. De fait, $x = \mathrm{e}^y$ est un changement de variable régulier d'ordre infini entre $]a, b[$ et $]\ln(a), \ln(b)[$ et ainsi il existe $u \in \mathrm{C}_p]a, b[$ tel que

$$c_p(x\mathrm{D})^p u + \ldots + c_1(x\mathrm{D})u + c_0 u = f \quad \text{sur} \quad]a, b[$$

si et seulement s'il existe $v \in \mathrm{C}_p(]\ln(a), \ln(b)[)$ tel que

$$c_p\mathrm{D}^p v + \ldots + c_1\mathrm{D}v + c_0 v = g \quad \text{sur} \quad]\ln(a), \ln(b)[\,,$$

à savoir $v(y) = u(\mathrm{e}^y)$ ou $u(x) = v \ln(x)$.∎

Exemple. *Si $a \neq 0$ et b sont des nombres complexes, déterminer la solution la plus générale de l'équation d'Euler homogène*

$$ax\mathrm{D}u + bu = 0 \quad \textit{sur} \quad]0, +\infty[\,.$$

Vu la forme de la solution la plus générale de l'équation différentielle linéaire à coefficients constants correspondante, cette solution s'écrit $cx^{-b/a}$ avec $c \in \mathbb{C}$.

Exemple. *Si $a \neq 0$, b et c sont des nombres complexes, déterminer la solution la plus générale de l'équation d'Euler homogène*

$$a(x\mathrm{D})^2 u + bx\mathrm{D}u + cu = 0 \quad \textit{sur} \quad]0, +\infty[\,.$$

En procédant comme ci-dessus, on obtient
a) $(c_1 \ln(x) + c_2)x^{-b/(2a)}$ avec c_1, $c_2 \in \mathbb{C}$ si $b^2 - 4ac = 0$,
b) $c_1 x^{r_1} + c_2 x^{r_2}$ avec c_1, $c_2 \in \mathbb{C}$ si on a $b^2 - 4ac \neq 0$ et si r_1 et r_2 sont les deux zéros distincts du trinôme du second degré $az^2 + bz + c$.

Exercice. *Résoudre l'équation d'Euler $x\mathrm{D}u + u = \mathrm{e}^x$ sur $]0, +\infty[$.*

Suggestion. a) Résolvons d'abord l'équation différentielle linéaire à coefficients constants $\mathrm{D}v + v = \mathrm{e}^{\mathrm{e}^x}$ sur $\mathbb{R}$.

i) L'équation homogène $\mathrm{D}v + v = 0$ sur $\mathbb{R}$ admet bien sûr $C\mathrm{e}^{-x}$ avec $C \in \mathbb{C}$ comme solution la plus générale.

ii) Appliquons la méthode de la variation des constantes. L'équation $C\mathrm{e}^{-x} = \mathrm{e}^{\mathrm{e}^x}$ admet évidemment $C(x) = \mathrm{e}^x \mathrm{e}^{\mathrm{e}^x}$ comme solution. Dès lors

$$v_0(x) = \mathrm{e}^{-x} \cdot \int \mathrm{e}^x \mathrm{e}^{\mathrm{e}^x}\, dx = \mathrm{e}^{\mathrm{e}^x} \mathrm{e}^{-x}$$

est une solution particulière de l'équation non homogène.

b) Dès lors, la solution la plus générale de l'équation proposée s'écrit sous la forme

$$u(x) = \frac{C}{x} + \frac{\mathrm{e}^x}{x} \quad \text{avec} \quad C \in \mathbb{C}.$$

De la même manière, en recourant au changement de variable $x = -\mathrm{e}^y$, on obtient de suite le résultat suivant.

Proposition 9.1.7.2 *Si on a* $p \in \mathbb{N}_0$; $c_0, c_1, \ldots, c_p \in \mathbb{C}$; $c_p \neq 0$ *et* $f \in \mathrm{C}_l(]a,b[)$ *avec* $l \in \mathbb{N}$ *et* $]a,b[\subset]-\infty, 0[$, *la solution la plus générale de l'équation*

$$c_p(x\mathrm{D})^p u + \ldots + c_1(x\mathrm{D})u + c_0 u = f \quad sur \quad]a,b[$$

s'écrit $u(x) = v(\ln(-x))$ *où* v *est la solution la plus générale de l'équation différentielle linéaire à coefficients constants*

$$c_p\mathrm{D}^p v + \ldots + c_1\mathrm{D}v + c_0 v = g \quad sur \quad]\ln(-b), \ln(-a)[$$

où $g \in \mathrm{C}_l(]\ln(-b), \ln(-a)[)$ *est défini par* $g(y) = f(-\mathrm{e}^y)$. ∎

Remarque. Cela étant
a) l'équation (très simple en apparence) $x\mathrm{D}u + u = 0$ sur $\mathbb{R}$ admet une et une seule solution, à savoir $u = 0$,
b) l'équation (très simple en apparence) $(x\mathrm{D})^2u - 2x\mathrm{D}u + u = 0$ sur $\mathbb{R}$ admet comme solutions, les fonctions $u(x) = cx$ avec $c \in \mathbb{C}$,
c) l'équation (très simple en apparence) $(x\mathrm{D})^2u + 2x\mathrm{D}u + u = 0$ sur $\mathbb{R}$ admet une et une seule solution, à savoir $u = 0$.

9.2 Equations différentielles ordinaires

9.2.1 Généralités

Définitions. Une *équation différentielle ordinaire* d'*ordre* $p \in \mathbb{N}_0$ sur l'intervalle $]a,b[$ de $\mathbb{R}$ (on dit aussi, plus simplement, équation différentielle ordinaire s'il n'y a pas d'ambiguité sur p et $]a,b[$) est une équation de la forme

$$F(x, u, \mathrm{D}u, \ldots, \mathrm{D}^p u) = f$$

où F est un opérateur de dérivation réel d'ordre p sur $]a,b[$ et où f est une fonction réelle appartenant à $\mathrm{C}_l(]a,b[)$ avec $l \in \mathbb{N}$.

Résoudre l'équation différentielle ordinaire $Fu = f$ d'ordre p dans l'intervalle $]a,b[$ de $\mathbb{R}$, c'est trouver toutes les *solutions* de cette équation, c'est-à-dire toutes les fonctions réelles $u \in \mathrm{C}_p(]a,b[)$ vérifiant $Fu = f$ sur un ouvert de $]a,b[$. Il s'agit là d'un point de vue bien souvent trop ambitieux (réalisé par exemple dans le cas des équations d'Euler) et on doit régulièrement se contenter d'*intégrer* l'équation différentielle, c'est-à-dire trouver des équations implicites (faisant intervenir x et u, mais pas les dérivées de u), que doivent vérifier toutes les solutions u.

Remarque. L'intégration des équations différentielles ordinaires est un problème très délicat qui ne peut être abordé en toute généralité en première candidature. Dans cette introduction, nous nous limitons à quelques types classiques, rencontrés couramment dans les applications. Parfois nous devons même nous limiter à donner une méthode sans guère la justifier; dans un tel cas, la vérification de la réponse s'impose.

Définition. $* \rightarrow$ Etant donné un ouvert Ω de $\mathbb{R}^2$ et une fonction $f \in \mathrm{C}_1(\Omega)$ réelle, une *solution* de l'équation différentielle $\mathrm{D}_x u = f(x,u)$ sur l'intervalle ouvert I de $\mathbb{R}$ est une fonction réelle $u \in \mathrm{C}_1(I)$ telle que

$$\begin{cases} (x, u(x)) \in \Omega, & \forall x \in I, \\ [\mathrm{D}u]_x = f(x, u(x)), & \forall x \in I. \end{cases}$$

Dans ce contexte, *le problème de Cauchy* de *condition initiale* $(x_0, y_0) \in \Omega$ consiste en la recherche des solutions u de l'équation $\mathrm{D}_x u = f(x,u)$ telles que $u(x_0) = y_0$. $\leftarrow *$

Dans le cadre de ce cours, contentons-nous de signaler le résultat suivant qui sera développé et généralisé au cours d'analyse non linéaire.

Théorème 9.2.1.1 (unicité) *Le problème de Cauchy* $\mathrm{D}_x u = f(x,u)$ *de condition initiale* $(x_0, y_0) \in \Omega$ *a une et une seule solution dans tout intervalle ouvert* I *de* $\mathbb{R}$ *contenant* x_0.∎

Remarque. En général, on remarque que la solution la plus générale d'une équation différentielle ordinaire d'ordre p dépend de p constantes. On remarque aussi parfois l'existence de solutions particulières.

9.2.2 Equations différentielles du type $\mathrm{D}u = f(x,u)$

1. Equations exactes

Une *équation différentielle exacte* est une équation différentielle ordinaire qui s'écrit

$$\mathrm{D}u = -\frac{\mathrm{D}_1 U(x,u)}{\mathrm{D}_2 U(x,u)} \quad \text{sur} \quad]a,b[\qquad (*)$$

où $U(x,u)$ appartient à $\mathrm{C}_1(]a,b[\times \mathbb{R})$ et est réel.

Proposition 9.2.2.1 *Si* $U \in \mathrm{C}_1(]a,b[\times \mathbb{R})$ *est une fonction réelle, alors une fonction réelle* $u \in \mathrm{C}_1(]a,b[)$ *vérifie l'équation*

$$[\mathrm{D}_1 U]_{(x,u(x))} + [\mathrm{D}_2 U]_{(x,u(x))} \cdot \mathrm{D}u(x) = 0 \quad sur \quad]a,b[$$

si et seulement si la fonction $U(x, u(x))$ *est constante sur* $]a,b[$.

Preuve. C'est immédiat à partir du théorème de dérivation des fonctions composées et du théorème de l'ouvert connexe.∎

Méthode. La résolution de l'équation exacte $(*)$ revient donc à déterminer les constantes $C \in \mathbb{R}$ et les fonctions réelles $u \in \mathrm{C}_1(]a,b[)$ telles que $U(x,u(x)) = C$ pour tout $x \in]a,b[$.

Remarques. a) $* \rightarrow$ Si $f(x,u)$, $g(x,u) \in \mathrm{C}_1(]a,b[)$ sont réels et tels que

$$\mathrm{D}_2 f(x,u) = \mathrm{D}_1 g(x,u), \quad \forall (x,u) \in]a,b[\times \mathbb{R}, \qquad (*)$$

l'équation $\mathrm{D}u = -f(x,u)/g(x,u)$ sur $]a,b[$ est exacte. En effet, pour tout point $(x_0,u_0) \in]a,b[\times \mathbb{R}$, la fonction U définie sur $]a,b[\times \mathbb{R}$ par

$$U(x,u) = \int_{u_0}^{u} g(x_0,t)\,dt + \int_{x_0}^{x} f(t,u)\,dt \quad \text{ou} \quad \int_{x_0}^{x} f(t,u_0)\,dt + \int_{u_0}^{u} g(x,t)\,dt$$

convient (justification par les primitives dans $\mathbb{R}^2$). $\leftarrow *$

b) Si $f(x,u)$, $g(x,u) \in \mathrm{C}_1(]a,b[)$ sont réels et ne vérifient pas la relation (*) de la remarque précédente, il peut se faire qu'il existe $h \in \mathrm{C}_1]a,b[\times \mathbb{R}$ réel tel que

$$\mathrm{D}_2(f(x,u)h(x,u)) = \mathrm{D}_1(g(x,u)h(x,u)), \quad \forall (x,u) \in]a,b[\times \mathbb{R}.$$

A ce moment, on peut appliquer la remarque 1) à l'équation différentielle

$$\mathrm{D}u = -\frac{f(x,u)\cdot h(x,u)}{g(x,u)\cdot h(x,u)} \quad \text{sur} \quad]a,b[\,.$$

Une telle fonction h est appelée *facteur intégrant* —malheureusement, on ne dispose pas d'un critère général donnant l'existence d'un facteur intégrant et encore moins sa valeur.

Exemple. *Trouver, si elle existe, une solution* $u \in \mathrm{C}_1(]-1,1[)$ *réelle de l'équation*

$$\mathrm{D}u = -\frac{2xu+1}{x^2+4u}$$

vérifiant $u(0) = 1$ (resp. $u(0) = 0$).

Comme on a $2xu+1$, $x^2+4u \in \mathrm{C}_\infty(\mathbb{R}^2)$ et

$$\mathrm{D}_2(2xu+1) = 2x = \mathrm{D}_1(x^2+4u),$$

on cherche une fonction $U(x,u) \in \mathrm{C}_1(]-1,1[\times \mathbb{R})$ telle que

$$\mathrm{D}_1 U(x,u) = 2xu+1 \quad \text{et} \quad \mathrm{D}_2 U(x,u) = x^2+4u.$$

On trouve aisément $U(x,u) = x^2u + x + 2u^2$. Il suffit alors de déterminer la ou les solutions réelles $u \in \mathrm{C}_1(]-1,1[)$ du système

$$\left\{\begin{array}{l} u(0) = 1 \quad [\text{resp. } u(0) = 0] \\ 2u^2 + x^2u + x = C \text{ avec } \; C \in \mathbb{R}. \end{array}\right.$$

La première équation impose la valeur $C = 2$ et on trouve

$$u = \frac{-x^2 + \sqrt{x^4 - 8x + 16}}{4} \in \mathrm{C}_1(]-1,1[)$$

comme solution (resp. la première équation impose la valeur $C = 0$; cela étant, comme $x^4 - 8x$ est strictement négatif sur $]0,1[$, il n'existe pas de solution).

Exemple. *Intégrer l'équation différentielle* $\mathrm{D}u = -(2u+1)/x$.

Bien sûr, les fonctions $2u+1$ et x appartiennent à $\mathrm{C}_\infty(\mathbb{R}^2)$ mais on a $\mathrm{D}_u(2x+1) \neq \mathrm{D}_x x$. Cependant on vérifie directement que $h(x,u) = x$ est un facteur intégrant. Cela étant, on trouve de suite $U(x,u) = ux^2 + x^2/2$. L'équation implicite régissant l'équation proposée s'écrit donc $ux^2 + x^2/2 = C$ avec $C \in \mathbb{R}$.

Remarque. a) Si on considère le dernier exemple sur un intervalle $]a,b[$ inclus dans $]-\infty,0[$ ou $]0,+\infty[$, c'est aussi une équation d'Euler.

b) Sur un intervalle $]a,b[$ contenant 0, il existe une et une seule solution, à savoir $-\frac{1}{2}\chi_{]a,b[}$.

2. Equations à second membre séparé

Une *équation différentielle à second membre séparé* est une équation différentielle ordinaire qui s'écrit

$$\mathrm{D}u = f(x)g(u).$$

Théorème 9.2.2.2 *Etant donné des fonctions réelles* $f \in \mathrm{C}_0(]a,b[)$ *et* $g \in \mathrm{C}_0(]c,d[)$*, et des points* $x_0 \in \,]a,b[$ *et* $u_0 \in \,]c,d[$*,*

a) *si on a* $g(u_0) \neq 0$*, il existe* $\alpha, \beta \in \mathbb{R}$ *tels que* $a \leq \alpha < x_0 < \beta \leq b$ *et une fonction réelle* $u \in \mathrm{C}_1(]\alpha,\beta[)$ *unique telle que*

$$\left\{\begin{array}{l} u(]\alpha,\beta[) \subset \,]c,d[\\ u(x_0) = u_0 \\ \mathrm{D}u(x) = f(x)g(u(x)), \quad \forall x \in \,]\alpha,\beta[\,. \end{array}\right.$$

b) *si on a* $g(u_0) = 0$*, la fonction* $u = u_0\chi_{]a,b[}$ *appartient à* $\mathrm{C}_\infty(]a,b[)$*, est réelle et vérifie* $u(]a,b[) \subset \,]c,d[$*,* $u(x_0) = u_0$ *et* $\mathrm{D}u(x) = f(x) \cdot g(u(x))$ *pour tout* $x \in \,]a,b[$*.* (Il s'agit d'une solution particulière de l'équation $\mathrm{D}u = f(x) \cdot g(u)$; on a perdu l'unicité).

Preuve. b) est trivial; établissons a).

Soit F la primitive réelle de f sur $]a,b[$ vérifiant l'égalité $F(x_0)=0$. Comme on a $g(u_0)\neq 0$ avec $g\in \mathrm{C}_0(]c,d[)$, il existe γ, $\delta\in\mathbb{R}$ tels que $c\leq\gamma<u_0<\delta\leq d$ ainsi que $g(t)\neq 0$ pour tout $t\in\,]\gamma,\delta[$. Soit alors G la primitive réelle de $1/g$ sur $]\gamma,\delta[$ telle que $G(u_0)=0$. Comme g garde un signe constant sur $]\gamma,\delta[$, G détermine un changement de variable régulier d'ordre $\neq 1$ entre $]\gamma,\delta[$ et un intervalle ouvert I de $\mathbb{R}$, contenant 0.

Analyse. Si on a de tels α, β et u, il existe un voisinage V de x_0 tel que $u(x)$ appartienne à $]\gamma,\delta[$ pour tout $x\in V$. Dès lors, la fonction de fonction $G(u)$ appartient à $\mathrm{C}_1(V)$ et vérifie

$$[\mathrm{D}G(u)]_x=\frac{\mathrm{D}u(x)}{g(u(x))}=f(x),\quad \forall x\in V.$$

Il s'ensuit qu'on a $G(u)\simeq F$ sur V et même $G(u(x))=F(x)$ pour tout $x\in V$ car on a $G(u(x_0))=G(u_0)=0=F(x_0)$. Ceci assure l'unicité de u.

Synthèse. Il existe α, $\beta\in\mathbb{R}$ tels que $a\leq\alpha<x_0<\beta\leq b$ et $F(]\alpha,\beta[)\subset I$. Cela étant, $u=G^{-1}(F)$ est une fonction réelle appartenant à $\mathrm{C}_1(]\alpha,\beta[)$ et telle que $u(]\alpha,\beta[)\subset\,]\gamma,\delta[$, $u(x_0)=u_0$ et

$$\mathrm{D}u(x)=\frac{1}{[\mathrm{D}G]_{G^{-1}(F(x))}}\cdot f(x)=g(u(x))\cdot f(x),\quad \forall x\in\,]\alpha,\beta[\,,$$

ce qui suffit.∎

Méthode. a) La solution générale de l'équation à second membre séparé $\mathrm{D}u=f(x)g(u)$ s'obtient en résolvant l'équation

$$\int\frac{du}{g(u)}=\int f(x)\,dx+C\quad\text{avec}\quad C\in\mathbb{R}$$

par rapport à u.

b) Si g s'annule en c, $u=c\chi_{]a,b[}$ est une solution généralement particulière.

Exemple. *Trouver, si elles existent, toutes les solutions de l'équation différentielle* $\mathrm{D}u=-x/u$ *qui vérifient* $u(1)=1$.

Ici on a $f(x)=-x$ et $g(u)=1/u$. Les hypothèses sont alors réalisées pour $]a,b[\,=\,]-\infty,\infty[$ et $]c,d[$ égal à $]-\infty,0[$ ou $]0,+\infty[$; comme on veut avoir $u(1)=1$, c'est $]c,d[\,=\,]0,+\infty[$ que nous choisissons. La fonction g ne s'annule en aucun point de $]0,+\infty[$: il n'y a pas de solution particulière. De

$$\int u\,du=-\int x\,dx+C\quad\text{avec}\quad C\in\mathbb{R},$$

on tire l'équation implicite $u^2 = -x^2 + C$ avec $C \in \mathbb{C}$. La condition $u(1) = 1$ impose $C = 2$. Il s'ensuit que, pour tout intervalle ouvert $]a, b[$ contenant 1 et inclus dans $]-\sqrt{2}, \sqrt{2}[$, il existe une et une seule solution, à savoir $u(x) = (2 - x^2)^{1/2}$ et qu'il n'existe pas de solution sur un intervalle contenant 1 et non inclus dans $]-\sqrt{2}, \sqrt{2}[$.

Exemple. *Trouver, si possible, un intervalle ouvert I contenant 0 et une fonction réelle $u \in \mathrm{C}_1(I)$ vérifiant $u(0) = 0$ et l'équation différentielle ordinaire*

$$\mathrm{e}^x \sin(x) - \mathrm{e}^u \cos(u)\mathrm{D}u = 0.$$

Dans ce problème, on a d'une part $f(x) = \mathrm{e}^x \sin(x) \in \mathrm{C}_\infty(\mathbb{R})$ ainsi que $g(u) = 1/(\mathrm{e}^u \cos(u)) \in \mathrm{C}_\infty(]-\pi/2, \pi/2[)$. On en déduit qu'il existe une solution unique déterminée par le système

$$\begin{cases} u(0) = 0 \\ \int \mathrm{e}^u \cos(u)\, du = \int \mathrm{e}^x \sin(x)\, dx + C \quad \text{avec} \quad C \in \mathbb{R}. \end{cases}$$

Tout revient à résoudre le système

$$\begin{cases} u(0) = 0 \\ \dfrac{\sin(u) + \cos(u)}{2}\mathrm{e}^u = \dfrac{\sin(x) - \cos(x)}{2}\mathrm{e}^x + C \quad \text{avec} \quad C \in \mathbb{R}. \end{cases}$$

On doit donc avoir $C = 1$; u est solution de

$$(\sin(u) + \cos(u))\mathrm{e}^u = (\sin(x) - \cos(x))\mathrm{e}^x + 2.$$

La résolution de cette équation est impraticable; il s'agit d'une *solution sous forme implicite.*

Exemple. *Résoudre l'équation différentielle* $\mathrm{D}u = -(2u + 1)/x$. *Existe-t-il une solution u telle que $u(1) = 1$?*

Il s'agit d'une équation à second membre séparé.

Comme on a $-2u - 1 \in \mathrm{C}_\infty(\mathbb{R})$ et $1/x \in \mathrm{C}_\infty(]0, +\infty[)$, la solution générale s'obtient en résolvant

$$\int \frac{du}{2u + 1} = -\int \frac{dx}{x} + C \quad \text{avec} \quad C \in \mathbb{R}$$

par rapport à u. Cela revient à résoudre $\frac{1}{2} \ln |2u + 1| = -\ln |x| + C$ avec $C \in \mathbb{R}$, c'est-à-dire $\ln(|2u + 1|\, x^2) = C$ avec $C \in \mathbb{R}$. Dès lors, pour $2u + 1 > 0$, on obtient $(2u + 1)x^2 = C'$ avec $C' > 0$, et, pour $2u + 1 < 0$, $(2u + 1)x^2 = C'$ avec $C' < 0$; de là, on tire $u(x)$.

Comme $2u + 1 = 0$ a lieu pour $u = -1/2$, $u(x) = -1/2$ est une solution particulière qu'on ne trouve d'ailleurs pour aucune valeur de C'.

Pour avoir $u(1) = 1$, on doit avoir $C' = 3$ et alors $u_0(x) = 3x^{-2}/2 - 1/2$ est la solution sur $]0, +\infty[$ vérifiant $u_0(1) = 1$.

2*. Equation $\mathrm{D}u = f(ax + bu + c)$ **où** a**,** b**,** $c \in \mathbb{R}$**;** $a \neq 0$**,** $b \neq 0$

Méthode. La fonction réelle $u \in \mathrm{C}_1(I)$ vérifie l'équation envisagée $\mathrm{D}u = f(ax + bu + c)$ si et seulement si la fonction $v = ax + bu + c$ est réelle, appartient à $\mathrm{C}_1(I)$ et vérifie l'équation $a + b\mathrm{D}u = \mathrm{D}v$. Il revient donc au même d'intégrer l'équation différentielle à second membre séparé $\mathrm{D}v = bf(v) + a$.

Exemple. *Intégrer l'équation différentielle* $\mathrm{D}u = -\sin^2(x + u)$.

Posons $x + u = v$ et intégrons $\mathrm{D}v = \cos^2(v)$. Il s'agit d'une équation à second membre séparé. Sa solution générale sur $I \subset \mathbb{R}$ s'obtient en résolvant

$$\int \frac{dv}{\cos^2(v)} = \int dx + C \quad \text{avec} \quad C \in \mathbb{R};$$

il vient donc $\mathrm{tg}(v) = x + C$ avec $C \in \mathbb{R}$, c'est-à-dire $x + u = \mathrm{arctg}(x + C) + k\pi$ avec $C \in \mathbb{R}$ et $k \in \mathbb{Z}$.

Comme on a $\cos^2(v) = 0$ pour $v = (2k+1)\pi/2$ avec $k \in \mathbb{Z}$, on a aussi les solutions particulières $u = -x + (2k + 1)\pi/2$ avec $k \in \mathbb{Z}$.

3. Equations à second membre homogène par rapport à x **et** u

Une équation différentielle *à second membre homogène par rapport à* x *et* u est une équation différentielle qui s'écrit $\mathrm{D}u = f(u/x)$.

Méthode. Le cas $f(u/x) = u/x$ est connu: on peut l'envisager correspondant à une équation d'Euler ou à une équation à second membre séparé.

Si on cherche une solution sur un intervalle I inclus dans $\mathbb{R} \setminus \{0\}$ ou au voisinage d'un point $x_0 \neq 0$, on peut introduire la fonction auxiliaire $v(x) = u(x)/x$. On a alors $u(x) = xv(x)$ sur l'ensemble considéré et l'équation devient $\mathrm{D}v = (f(v) - v)/x$, c'est-à-dire une équation différentielle à second membre séparé.

Exemple. *Intégrer l'équation différentielle* $\mathrm{D}u = u^2/x^2$.

Posons $v = u/x$. L'équation devient alors $\mathrm{D}v = (v^2 - v)/x$, c'est-à-dire une équation à second membre séparé, dont la solution générale éventuelle s'obtient en résolvant

$$\int \frac{dv}{v^2 - v} = \ln|x| + C \quad \text{avec} \quad C \in \mathbb{R},$$

c'est-à-dire

$$\ln\left|\frac{v - 1}{v}\right| = \ln|x| + C \quad \text{avec} \quad C \in \mathbb{R}.$$

Cela étant, la solution générale s'obtient en résolvant

$$\left|\frac{u - x}{ux}\right| = C \quad \text{avec} \quad C > 0.$$

De plus, $v^2 = v$ a lieu pour $v = 0$ et $v = 1$ et ainsi $u = 0$ et $u = x$ sont des solutions particulières.

3*. Equation $\mathrm{D}u = u/x + f(x)g(u/x)$

La méthode utilisée pour résoudre les équations à second membre homogène par rapport à x et u s'applique également aux équations de ce type.

En effet, en posant $v = u/x$, il vient $u = vx$ donc $\mathrm{D}u = v + x\mathrm{D}v$ et l'équation se transforme en l'équation à second membre séparé $\mathrm{D}v = g(v) \cdot f(x)/x$.

Exemple. *Intégrer l'équation différentielle* $\mathrm{D}u = u/x + u^2/x^3$.

En posant $v = u/x$, on est amené à résoudre l'équation différentielle $\mathrm{D}v = v^2/x^2$, dont la solution générale éventuelle est obtenue en résolvant l'équation $-1/v = -1/x + C$ avec $C \in \mathbb{R}$. Dès lors, la solution générale éventuelle est donnée par $u = x^2/(1 + Cx)$.

De plus, $v^2 = 0$ a lieu pour $v = 0$ et ainsi $u = 0$ est une solution particulière.

4. Equations à second membre linéaire en u

Une *équation différentielle à second membre linéaire en* u est une équation différentielle ordinaire qui s'écrit $\mathrm{D}u = a(x)u + b(x)$ sur I avec a, $b \in \mathrm{C}_0(I)$ réels, I étant un intervalle ouvert de $\mathbb{R}$.

Bien sûr toute équation différentielle ordinaire qui s'écrit $c(x)\mathrm{D}u = a(x)u + b(x)$ sur I avec a, b, $c \in \mathrm{C}_0(I)$ réels, I étant un intervalle ouvert de $\mathbb{R}$ et $c(x) \neq 0$ pour tout $x \in I$ est une équation à second membre linéaire en u.

Proposition 9.2.2.3 *La solution la plus générale de l'équation à second membre linéaire* $\mathrm{D}u = a(x)u + b(x)$ *sur l'intervalle ouvert* I *de* $\mathbb{R}$ *s'écrit*

$$u(x) = \mathrm{e}^{\int a(x)\,dx} \cdot \left(\int b(x)\mathrm{e}^{-\int a(x)\,dx}\,dx + C \right) \quad \textit{avec} \quad C \in \mathbb{R}.$$

Preuve. **Analyse.** Si v, $w \in \mathrm{C}_1(I)$ sont réels et tels que le produit $u = vw$ soit solution de l'équation proposée, on a

$$v\mathrm{D}w + w\mathrm{D}v = a(x)vw + b(x) \text{ sur } I$$

donc

$$v(\mathrm{D}w - a(x)w) + w\mathrm{D}v = b(x) \text{ sur } I.$$

Synthèse. Cela étant, $\mathrm{D}w = a(x)w$ sur I est une équation à second membre séparé dont

$$w_0(x) = \mathrm{e}^{\int a(x)\,dx} \in \mathrm{C}_1(I)$$

est une solution qui ne s'annule en aucun point de I. Dès lors, $u \in \mathrm{C}_1(I)$ réel est solution de l'équation proposée si et seulement si $v = u/w_0$ est solution de l'équation différentielle $\mathrm{D}v = b/w_0$ sur I, c'est-à-dire si et seulement si u/w_0 est une primitive de b/w_0 sur I.

D'où la conclusion.∎

Exemple. *Résoudre l'équation différentielle* $\mathrm{D}u = -u/x + x^2$.

Sur $]-\infty, 0[$ ou sur $]0, +\infty[$, on a directement

$$\begin{aligned} u(x) &= \mathrm{e}^{-\int \frac{dx}{x}} \left(\int x^2 \mathrm{e}^{\int \frac{dx}{x}} \, dx + C' \right) = \frac{1}{|x|} \left(\int x^2 \, |x| \, dx + C' \right) \\ &= \frac{1}{x} \left(\int x^3 \, dx + C \right) = \frac{x^3}{4} + \frac{C}{x} \quad \text{avec} \quad C \in \mathbb{R}. \end{aligned}$$

4*. Equations de Bernoulli

Une *équation de Bernoulli* est une équation différentielle de la forme

$$\mathrm{D}u = a(x)u + b(x)u^r,$$

où r est un nombre réel différent de 1 et de 0.

Si r est égal à 1, on a en fait une équation à second membre séparé; s'il est égal à 0, c'est une équation à second membre linéaire en u.

Si r diffère de 1 et de 0, en divisant les deux membres par u^r, on obtient

$$\mathrm{D}u^{1-r} = (1-r)a(x)u^{1-r} + (1-r)b(x),$$

qui est en fait une équation à second membre linéaire en u^{1-r}.

De plus, si r est strictement positif, $u = 0$ est une solution généralement singulière.

Exemple. *Résoudre l'équation différentielle* $\mathrm{D}u = -u/x + u^2/x^2$.

Après division des deux membres par u^2 et après avoir posé $v = 1/u$, on obtient l'équation $\mathrm{D}v = v/x - 1/x^2$. Cela étant, il vient successivement sur $]-\infty, 0[$ et sur $]0, +\infty[$

$$\begin{aligned} \frac{1}{u} &= v = \mathrm{e}^{\int \frac{dx}{x}} \cdot \left(- \int \frac{1}{x^2} \mathrm{e}^{-\int \frac{dx}{x}} \, dx + C' \right) \\ &= |x| \cdot \left(- \int \frac{dx}{x^2 \, |x|} + C' \right) = x \left(- \int \frac{dx}{x^3} + C \right) = \frac{1}{2x} + Cx \end{aligned}$$

avec $C \in \mathbb{R}$. De plus, $u = 0$ est une solution singulière.

4**. Equations renversées

Une *équation renversée* est une équation différentielle qui s'écrit

$$\mathrm{D}u = \frac{1}{a(u)x + b(u)} \quad \text{ou} \quad \mathrm{D}u = \frac{1}{a(u)x + b(u)x^r}$$

avec $r \in \mathbb{R}$ différent de 1 et de 0.

La méthode consiste à prendre x comme inconnue et u comme variable: sous des conditions adéquates, on a alors $\mathrm{D}_x u = 1/\mathrm{D}_u x$ et on est ramené à une équation à second membre linéaire en x ou à une équation de Bernoulli.

Exemple. *Résoudre l'équation* $\mathrm{D}u = \cos(u)/(1 - x\sin(u))$.

Nous allons plutôt résoudre l'équation $\mathrm{D}_u x = -x\mathrm{tg}(u) + 1/\cos(u)$. C'est une équation à second membre linéaire en l'inconnue; on a donc

$$\begin{aligned} x &= \mathrm{e}^{-\int \mathrm{tg}(u)\,du} \cdot \left(\int \frac{1}{\cos(u)} \mathrm{e}^{\int \mathrm{tg}(u)\,du}\, du + C' \right) \\ &= |\cos(u)| \cdot \left(\int \frac{du}{\cos(u)\,|\cos(u)|} + C' \right) \\ &= \cos(u) \cdot \left(\int \frac{du}{\cos^2(u)} + C \right) = \sin(u) + C\cos(u) \end{aligned}$$

avec $C \in \mathbb{R}$.

Exemple. *Résoudre l'équation différentielle* $\mathrm{D}u = -1/(ux + 2u^3x^2)$.

Nous allons plutôt résoudre l'équation de Bernoulli, $\mathrm{D}_u x = -ux - 2u^3x^2$ obtenue en inversant les rôles de u et de x. Elle devient $\mathrm{D}_u 1/x = u/x + 2u^3$ après division des deux membres par x^2. Cela étant, la solution est donnée sous forme implicite par l'équation suivante

$$\begin{aligned} \frac{1}{x} &= \mathrm{e}^{\int u\,du} \cdot \left(\int 2u^3 \mathrm{e}^{-\int u\,du}\, du + C \right) = \mathrm{e}^{u^2/2} \cdot \left(\int 2u^3 \mathrm{e}^{-u^2/2}\, du + C \right) \\ &= -2u^2 - 4 + C\mathrm{e}^{u^2/2} \end{aligned}$$

avec $C \in \mathbb{R}$.

9.2.3 Equations linéaires en x, u, à coefficients fonctions de $\mathrm{D}u$

1. Equations du type $x = f(\mathrm{D}u)$ avec $f \in \mathrm{C}_1(\mathbb{R})$ réel

Méthode. On cherche x et u en fonction du paramètre $\lambda = \mathrm{D}u$. On est donc amené à résoudre le système

$$\left\{\begin{array}{lcl} x & = & f(\lambda) \\ \mathrm{D}_\lambda u & = & \lambda\, \mathrm{D}_\lambda f(\lambda). \end{array}\right.$$

La deuxième équation donnant la dérivée de u par rapport à λ comme étant une fonction continue de λ, il vient

$$u(\lambda) = \int \lambda \mathrm{D}_\lambda f(\lambda)\, d\lambda + C.$$

Tout revient alors à éliminer λ dans le système d'équations

$$\left\{\begin{array}{lcl} x & = & f(\lambda) \\ u & = & \int \lambda\, \mathrm{D}_\lambda f(\lambda)\, d\lambda + C \quad \text{avec} \quad C \in \mathbb{R}. \end{array}\right.$$

En guise de “justification”, on note que

$$\mathrm{D}_\lambda u = \mathrm{D}_x u \cdot \mathrm{D}_\lambda x = \lambda\, \mathrm{D}_\lambda f(\lambda).$$

Exemple. *Résoudre l'équation différentielle $x = \mathrm{arctg}(\mathrm{D}u)$.*
Eliminons λ dans le système d'équations

$$\left\{\begin{array}{lcl} x & = & \mathrm{arctg}(\lambda) \\ u & = & \displaystyle\int \frac{\lambda}{1+\lambda^2}\, d\lambda + C \quad \text{avec} \quad C \in \mathbb{R}, \end{array}\right.$$

c'est-à-dire dans

$$\left\{\begin{array}{lcl} \mathrm{tg}(x) & = & \lambda \\ u & = & \frac{1}{2} \ln(1+\lambda^2) + C \quad \text{avec} \quad C \in \mathbb{R}. \end{array}\right.$$

Il vient

$$u = \frac{1}{2} \ln(1 + \mathrm{tg}^2(x)) + C = -\ln|\cos(x)| + C \quad \text{avec} \quad C \in \mathbb{R}.$$

Remarque. On aurait pu aussi ramener cette équation à $\mathrm{D}u = \mathrm{tg}(x)$.

2. Equations du type $u = f(\mathrm{D}u)$, avec $f \in \mathrm{C}_1(\mathbb{R})$ réel

Méthode. On cherche x et u en fonction du paramètre $\lambda = \mathrm{D}u$. On est donc amené à résoudre le système

$$\begin{cases} u &= f(\lambda) \\ \mathrm{D}_\lambda x &= \dfrac{\mathrm{D}_\lambda f(\lambda)}{\lambda}. \end{cases}$$

La deuxième équation donne de suite $x(\lambda) = \int \mathrm{D}_\lambda f(\lambda)/\lambda \, d\lambda$. Tout revient alors à éliminer λ dans le système d'équations

$$\begin{cases} u &= f(\lambda) \\ x &= \displaystyle\int \frac{\mathrm{D}_\lambda f(\lambda)}{\lambda}\, d\lambda + C \quad \text{avec} \quad C \in \mathbb{R}. \end{cases}$$

En guise de "justification", on note que

$$\mathrm{D}_\lambda x = \mathrm{D}_u x \cdot \mathrm{D}_\lambda u = \frac{\mathrm{D}_\lambda u}{\mathrm{D}_x u} = \frac{\mathrm{D}_\lambda f(\lambda)}{\lambda}.$$

Exemple. *Résoudre l'équation différentielle* $u = (\mathrm{D}u)^2 \cdot \ln(\mathrm{D}u)$.
Tout revient à éliminer λ dans le système d'équations

$$\begin{cases} u &= \lambda^2 \ln(\lambda) \\ x &= \displaystyle\int \frac{2\lambda \ln(\lambda) + \lambda}{\lambda}\, d\lambda + C \\ &= \int (2\ln(\lambda) + 1)\, d\lambda + C = 2\lambda \ln(\lambda) - \lambda + C \quad \text{avec} \quad C \in \mathbb{R}. \end{cases}$$

Remarque. On voit bien que cette entreprise d'élimination de λ est généralement difficile et même bien souvent irréalisable sans moyens supplémentaires.

3. Equations de Lagrange et de Clairaut

Il s'agit des équations du type

$$u = a(\mathrm{D}u)x + b(\mathrm{D}u),$$

la fonction a n'étant pas identiquement nulle. Une telle équation est de *Lagrange* si $\mathrm{D}u \neq a(\mathrm{D}u)$; elle est de *Clairaut* si $\mathrm{D}u = a(\mathrm{D}u)$.

Méthode. On cherche x et u en fonction du paramètre $\lambda = \mathrm{D}u$. On est donc amené à résoudre le système

$$\begin{cases} u &= a(\lambda)x + b(\lambda) \\ \lambda &= x\mathrm{D}_\lambda a(\lambda) \cdot \mathrm{D}_x \lambda + a(\lambda) + \mathrm{D}_\lambda b(\lambda) \cdot \mathrm{D}_x \lambda \end{cases}$$

où la deuxième équation est obtenue en dérivant la première par rapport à x.

Cela étant, la méthode diffère selon qu'il s'agit d'une équation de Lagrange ou d'une équation de Clairaut.

a) *Equation de Lagrange.* La deuxième équation s'écrit également

$$\mathrm{D}\lambda = \frac{\lambda - a(\lambda)}{x\mathrm{D}_\lambda a(\lambda) + \mathrm{D}_\lambda b(\lambda)};$$

c'est donc une équation renversée. On résout alors l'équation

$$\mathrm{D}_\lambda x = \frac{x\mathrm{D}_\lambda a(\lambda) + \mathrm{D}_\lambda b(\lambda)}{\lambda - a(\lambda)};$$

elle est à second membre linéaire en x si $b(\lambda)$ n'est pas constant et à second membre séparé si $b(\lambda)$ est constant. En outre, pour tout λ_0 tel que $\lambda_0 = a(\lambda_0)$, $u = \lambda_0 x + b(\lambda_0)$ est une solution, généralement particulière.

b) *Equation de Clairaut.* La deuxième équation s'écrit

$$(x\mathrm{D}_\lambda a(\lambda) + \mathrm{D}_\lambda b(\lambda)) \cdot \mathrm{D}_x\lambda = 0.$$

D'une part, $\mathrm{D}_x\lambda = 0$ admet pour solution générale $\lambda = C$ avec $C \in \mathbb{R}$; il s'ensuit que $u = Cx + b(C)$ est la solution générale de l'équation proposée. D'autre part, on a le système

$$\left\{\begin{array}{lcl} u & = & \lambda x + b(\lambda) \\ x + \mathrm{D}_\lambda b(\lambda) & = & 0 \end{array}\right.$$

qu'il faut résoudre en éliminant λ entre ces deux équations.

Exemple. *Résoudre l'équation différentielle* $u = -x\mathrm{D}u + (\mathrm{D}u)^2$.
C'est une équation de Lagrange. Posons $\mathrm{D}u = \lambda$ et résolvons le système

$$\left\{\begin{array}{lcl} u & = & -\lambda x + \lambda^2 \\ \lambda & = & -\lambda - x\mathrm{D}_x\lambda + 2\lambda\mathrm{D}_x\lambda. \end{array}\right.$$

La deuxième équation s'écrit aussi $\mathrm{D}\lambda = 2\lambda/(2\lambda - x)$. En l'inversant, on est amené à résoudre $\mathrm{D}_\lambda x = -x/(2\lambda) + 1$, équation à second membre linéaire en x; sa solution générale s'écrit

$$\begin{aligned} x &= \mathrm{e}^{-\int \frac{d\lambda}{2\lambda}} \cdot \left(\int \mathrm{e}^{\int \frac{d\lambda}{2\lambda}}\, d\lambda + C\right) \\ &= \frac{1}{\sqrt{|\lambda|}}\left(\int \sqrt{|\lambda|}\, d\lambda + C\right) = \frac{2}{3}\,|\lambda| + \frac{C}{\sqrt{|\lambda|}} \quad \text{avec} \quad C \in \mathbb{R}. \end{aligned}$$

Pour $\lambda > 0$, on a donc $x = 2\lambda/3 + C\lambda^{-1/2}$ et pour $\lambda < 0$, $x = -2\lambda/3 + C(-\lambda)^{-1/2}$. Il reste alors à éliminer λ entre ces relations et l'équation $u = -\lambda x + \lambda^2$.

De plus, on a $a(\lambda_0) = \lambda_0$ si et seulement si $\lambda_0 = 0$ et ainsi $u = 0$ est une solution particulière.

Exemple. *Résoudre l'équation différentielle* $u = x\mathrm{D}u + (\mathrm{D}u)^2$.

C'est une équation de Clairaut. Posons $\mathrm{D}u = \lambda$ et résolvons le système

$$\left\{\begin{array}{lcl} u & = & \lambda x + \lambda^2 \\ \lambda & = & \lambda + x\mathrm{D}_x\lambda + 2\lambda\mathrm{D}_x\lambda. \end{array}\right.$$

La deuxième équation s'écrit aussi $(x + 2\lambda)\mathrm{D}_x\lambda = 0$. Cela étant, la solution générale s'écrit $u = Cx + C^2$ avec $C \in \mathbb{R}$ tandis que $u = -x^2/4$ est une solution particulière.

9.2.4 Equations différentielles d'ordre deux

1. Equations où u manque

Il s'agit donc des équations du type $\mathrm{D}^2u = f(x, \mathrm{D}u)$.

Méthode. On pose $\mathrm{D}u = v$. L'équation devient $\mathrm{D}v = f(x, v)$, c'est-à-dire une équation qui est éventuellement résoluble par une des méthodes du paragraphe 9.2.2. C'est le cas notamment si $f(x, v)$ s'écrit sous l'une des formes suivantes:

$$\begin{array}{lcll} f(x,v) & = & g(x)h(v) & \text{(à second membre séparé),} \\ f(x,v) & = & g(v/x) & \text{(à second membre homogène),} \\ f(x,v) & = & g(x)\cdot v + h(x) & \text{(à second membre linéaire).} \end{array}$$

A ce moment sa solution s'écrit $v = F(x, C)$ où C est un paramètre réel. Il reste alors à résoudre l'équation différentielle triviale $\mathrm{D}u = F(x, C)$, dont la solution nécessitera la présence d'une deuxième constante.

Exemple. *Résoudre l'équation différentielle* $\mathrm{D}^2u = (\mathrm{D}u)^2$.

Si on pose $\mathrm{D}u = v$, on est amené à résoudre l'équation à second membre séparé $\mathrm{D}v = v^2$. Elle admet

$$\int \frac{dv}{v^2} = -\frac{1}{v} = x + C \quad \text{avec} \quad C \in \mathbb{R}$$

pour solution générale et $v = 0$ comme solution singulière.

Cela étant, d'une part,

$$u = -\int \frac{dx}{x + C} + C' = -\ln|x + C| + C' \quad \text{avec} \quad C, C' \in \mathbb{R}$$

est la solution générale et, d'autre part, $u = C$ avec $C \in \mathbb{R}$ est une solution particulière.

Exemple. *Résoudre l'équation différentielle* $\mathrm{D}^2u = (\mathrm{D}u)/x + 1$.

Si on pose $\mathrm{D}u = v$, on est amené à résoudre l'équation à second membre homogène par rapport à v et x suivante $\mathrm{D}v = v/x+1$. En posant $v/x = w$, elle devient $\mathrm{D}w = 1/x$, d'où on tire $w = \ln|x| + C$, donc $v = x\ln|x| + Cx$ avec $C \in \mathbb{R}$.

Il reste alors à résoudre l'équation différentielle $\mathrm{D}u = x \cdot \ln|x| + Cx$ avec $C \in \mathbb{R}$. On obtient de suite

$$u(x) = \int (x\ln(x) + Cx)\,dx + C' = \frac{x^2}{2}\ln(x) - \frac{x^2}{4} + C\frac{x^2}{2} + C'$$

sur $]0, +\infty[$, par exemple.

Exemple. *Résoudre l'équation différentielle* $(1 - x^2)\mathrm{D}^2u = 2 + x\mathrm{D}u$.

En posant $\mathrm{D}u = v$, on est amené à résoudre l'équation différentielle à second membre linéaire en v suivante

$$\mathrm{D}v = \frac{x}{1-x^2}v + \frac{2}{1-x^2}.$$

Sa solution générale s'écrit

$$\begin{aligned} v &= \mathrm{e}^{\int \frac{x}{1-x^2}\,dx} \cdot \left(\int \frac{2}{1-x^2}\mathrm{e}^{-\int \frac{x}{1-x^2}\,dx}\,dx + C \right) \\ &= \frac{1}{\sqrt{|1-x^2|}} \cdot \left(\int \frac{2\sqrt{|1-x^2|}}{1-x^2}\,dx + C \right) \end{aligned}$$

avec $C \in \mathbb{R}$.

Sur $]-1, 1[$, on a donc

$$v = \frac{1}{\sqrt{1-x^2}}\,(2\arcsin(x) + C)$$

et par conséquent,

$$u = \arcsin^2(x) + C\arcsin(x) + C' \text{ avec } \; C, C' \in \mathbb{R}.$$

Sur $]-\infty, -1[$ et sur $]1, +\infty[$, il vient

$$v = \frac{1}{\sqrt{x^2-1}}(\mp 2\operatorname{arcch}(|x|) + C)$$

et par conséquent,

$$u = -\operatorname{arcch}^2(|x|) \pm C\operatorname{arcch}(|x|) + C' \text{ avec } \; C, C' \in \mathbb{R}.$$

2. Equations où x manque

Il s'agit donc des équations du type $\mathrm{D}^2 u = f(u, \mathrm{D}u)$.

Méthode. *Analyse.* Posons $v(x) = \mathrm{D}_x u$. Si $u(x)$ est un changement de variable régulier d'ordre ≥ 2 entre I et J, nous pouvons définir la fonction

$$w \colon J \to \mathbb{R} \quad u \mapsto v(x(u)).$$

On a alors $w \in \mathrm{C}_1(J)$ et $w(u(x)) = v(x)$ donc

$$[\mathrm{D}_u w]_{u(x)} \cdot \mathrm{D}_x u = \mathrm{D}_x v = \mathrm{D}_x^2 u = f(u, \mathrm{D}_x u)$$

c'est-à-dire $\frac{1}{2}\mathrm{D}_u w^2 = w \cdot \mathrm{D}_u w = f(u, w)$ car $[\mathrm{D}_x u]_{x(u)} = v(x(u)) = w(u)$ et toute solution w de cette équation donne lieu à $\mathrm{D}_x u = v(x) = w(u(x))$.

Synthèse. On résout une des deux équations $\mathrm{D}_u w = f(u, w)$ ou $\frac{1}{2}\mathrm{D}_u w^2 = f(u, w)$ puis l'équation $\mathrm{D}_x u = w(u)$.

Exemple. *Résoudre l'équation différentielle* $\mathrm{D}^2 u = -u^{-3}$.

Prenons la forme $\mathrm{D}_u(\mathrm{D}u)^2 = -2u^{-3}$. Il s'agit d'une équation à second membre séparé: il vient donc

$$(\mathrm{D}u)^2 = -2\int u^{-3}\, du + C = u^{-2} + C \quad \text{avec} \quad C \in \mathbb{R}.$$

De là, on tire $\mathrm{D}u = \pm\sqrt{u^{-2} + C}$. Si, par exemple, u est une fonction strictement positive et si on a $C > 0$, on peut encore écrire

$$\mathrm{D}u = \pm\frac{1}{u}\sqrt{1 + Cu^2};$$

c'est une équation à second membre séparé dont la solution générale s'obtient en résolvant

$$x + C' = \pm\int \frac{u}{\sqrt{1 + Cu^2}}\, du = \pm\frac{\sqrt{1 + Cu^2}}{C} \quad \text{avec} \quad C > 0.$$

Remarquons que la position $\mathrm{D}_u(\mathrm{D}u) = -1/(u^3\mathrm{D}u)$ ne pose pas de problème non plus: il s'agit d'une équation à second membre séparé.

Exemple. *Résoudre l'équation différentielle* $\mathrm{D}^2 u = (1 + (\mathrm{D}u)^2)/u$.

Ici la position $\mathrm{D}_u(\mathrm{D}u)^2 = (2 + 2(\mathrm{D}u)^2)/u$ est toute indiquée. En posant $(\mathrm{D}u)^2 = v$, on obtient une équation à second membre linéaire en v (mais aussi séparé) et on a donc

$$\begin{aligned} v &= \mathrm{e}^{\int \frac{2}{u}\, du} \cdot \left(\int \frac{2}{u}\mathrm{e}^{-\int \frac{2}{u}\, du}\, du + C\right) \\ &= u^2 \cdot \left(\int \frac{2}{u^3}\, du + C\right) = -1 + Cu^2 \quad \text{avec} \quad C \in \mathbb{R}. \end{aligned}$$

Il reste alors à résoudre $(\mathrm{D}u)^2 = Cu^2 - 1$ (et on doit donc exiger $C > 0$), c'est-à-dire

$$x + C' = \pm \int \frac{du}{\sqrt{Cu^2 - 1}} = \pm \frac{1}{\sqrt{C}} \mathrm{arcch}(\sqrt{C}\, u) \quad \text{avec} \quad C > 0.$$

On obtient finalement

$$u = \frac{1}{\sqrt{C}} \mathrm{ch}(\sqrt{C}(x + C')) \quad \text{avec} \quad C > 0, C' \in \mathbb{R}.$$

3. Equations homogènes en u, $\mathrm{D}u$ et $\mathrm{D}^2 u$

Méthode. On pose $\mathrm{D}u = \lambda u$. Il vient

$$\mathrm{D}^2 u = \lambda \mathrm{D}u + u \mathrm{D}\lambda = u \cdot (\lambda^2 + \mathrm{D}\lambda)$$

et on peut alors simplifier par une puissance de u, ce qui ramène le problème à une équation différentielle du premier ordre en λ.

Exemple. *Résoudre l'équation différentielle* $u\mathrm{D}^2 u - (\mathrm{D}u)^2 = 6xu^2$.

Posons $\mathrm{D}u = \lambda u$. L'équation devient $u^2(\lambda^2 + \mathrm{D}\lambda) - \lambda^2 u^2 = 6xu^2$. De là, $u = 0$ est une solution singulière et on est amené à résoudre $\mathrm{D}\lambda = 6x$. On obtient $\lambda = 3x^2 + C$ avec $C \in \mathbb{R}$ et par conséquent $\mathrm{D}u = (3x^2 + C)u$ avec $C \in \mathbb{R}$, équation à second membre séparé. On tire $\ln(|u|) = x^3 + Cx + C'$ donc

$$u(x) = C^* \mathrm{e}^{x^3 + Cx} \quad \text{avec} \quad C \in \mathbb{R}, C^* \neq 0.$$

Table des matières