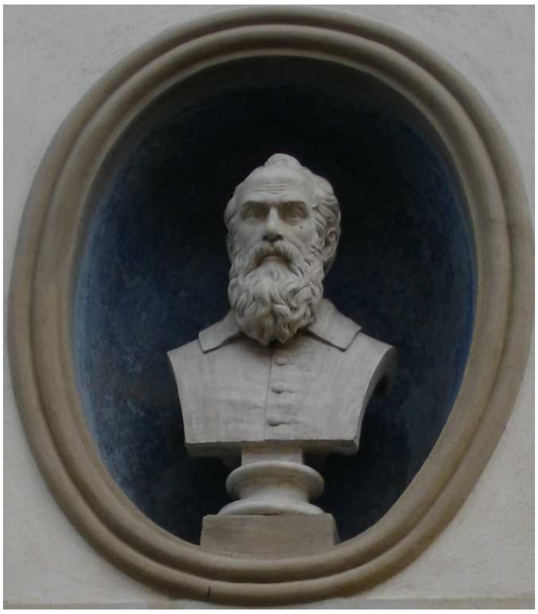
Mathématiques pour physiciens I

LP206, année 2010-2011

Notes de cours

Jesper Lykke Jacobsen *Michel Fioc*

UNIVERSITE PIERRE & MARIE CURIE
LA SCIENCE A PARIS



Nul ne peut lire le grand livre de l'Univers s'il n'en connaît la langue, qui est la langue mathématique. (Galilée) (*1)

1. La citation complète est la suivante : « La philosophie est écrite dans cet immense livre qui se tient toujours ouvert devant nos yeux, je veux dire l'Univers, mais on ne peut le comprendre si on ne s'applique d'abord à en comprendre la langue et à en connaître les caractères avec lesquels il est écrit. Il est écrit dans la langue mathématique et ses caractères sont des triangles, des cercles et autres figures géométriques, sans le moyen desquels il est humainement impossible d'en comprendre un mot. Sans eux, c'est une errance vaine dans un labyrinthe obscur. » (Galilée, *Il Saggiatore*, Rome 1623.)

Préface

Ces notes de cours reprennent largement celles rédigées par nos collègues Claude Aslangul et Jean-Bernard Zuber. Nous les remercions vivement pour avoir mis celles-ci à notre disposition.

Jesper Lykke Jacobsen *Michel Fioc*

Table des matières

Chapitre I

R

Avant de commencer, nous rappelons quelques résultats et notations élementaires qui seront supposés connus par la suite.

I. . Ensembles, logique

$a := b$ ou $a \overset{\text{déf}}{=} b$: la quantité a est *définie* par l'expression b (a est égale par définition à b).

$\forall\, x, \mathcal{P}(x)$: quel que soit x (on dit aussi « pour tout x »), la propriété $\mathcal{P}$ est vérifiée pour x. « $\forall$ » est appelé le quantificateur universel.

$\exists\, x, \mathcal{P}(x)$: il existe au moins un x tel que la propriété $\mathcal{P}$ soit vérifiée pour x. « $\exists$ » est appelé le quantificateur existentiel.

$\exists!\, x, \mathcal{P}(x)$: il existe un *unique* x (« il existe un et un seul x ») vérifiant la propriété $\mathcal{P}$.

$x \in A$: x appartient à l'ensemble A (« x est élément de A »).

$x \notin A$: x n'appartient pas à l'ensemble A.

$\mathcal{P}_1 \Rightarrow \mathcal{P}_2$: la propriété $\mathcal{P}_1$ *implique* la propriété $\mathcal{P}_2$. Une *condition nécessaire* pour que $\mathcal{P}_1$ soit vraie est que $\mathcal{P}_2$ soit vraie.

$\mathcal{P}_1 \Leftarrow \mathcal{P}_2$: une *condition suffisante* pour que $\mathcal{P}_1$ soit vraie est que $\mathcal{P}_2$ soit vraie.

$\mathcal{P}_1 \Leftrightarrow \mathcal{P}_2$: les propriétés $\mathcal{P}_1$ et $\mathcal{P}_2$ sont *équivalentes*, ce qu'on énonce « $\mathcal{P}_1$ si et seulement si ("ssi" en abrégé) $\mathcal{P}_2$ ». Une *condition nécessaire et suffisante* pour que $\mathcal{P}_1$ soit vraie est que $\mathcal{P}_2$ soit vraie.

non $\mathcal{P}$ (aussi noté $\neg\mathcal{P}$) : négation de $\mathcal{P}$. La propriété « non $\mathcal{P}$ » est vraie si $\mathcal{P}$ est fausse et inversement.

$\mathcal{P}_1$ et $\mathcal{P}_2$ ($\mathcal{P}_1 \wedge \mathcal{P}_2$) : la propriété « $\mathcal{P}_1$ et $\mathcal{P}_2$ » est vraie si et seulement si $\mathcal{P}_1$ et $\mathcal{P}_2$ sont toutes deux vraies.

$\mathcal{P}_1$ ou $\mathcal{P}_2$ ($\mathcal{P}_1 \vee \mathcal{P}_2$) : la propriété « $\mathcal{P}_1$ ou $\mathcal{P}_2$ » est vraie si et seulement si l'une au moins des propriétés $\mathcal{P}_1$ et $\mathcal{P}_2$ est vraie [*1].

Ordre décroissant de priorité des opérations logiques : non, et, ou, $\Rightarrow$, $\Longleftrightarrow$.

$\mathcal{P}_1 \Rightarrow \mathcal{P}_2 \Longleftrightarrow \text{non}\,\mathcal{P}_2 \Rightarrow \text{non}\,\mathcal{P}_1$ (raisonnement par contraposée).

$\text{non}(x \in A) \Longleftrightarrow x \notin A$.

$\text{non}(\forall\, x, \mathcal{P}(x)) \Longleftrightarrow \exists\, x,\ \text{non}\,\mathcal{P}(x)$.

$\text{non}(\forall\, x \in A, \mathcal{P}(x)) \Longleftrightarrow \exists\, x \in A,\ \text{non}\,\mathcal{P}(x)$.

1. Rappelons que, en logique comme en français, le « ou » simple est *inclusif*. Pour indiquer un « ou » *exclusif*, c.-à-d. en excluant le cas où $\mathcal{P}_1$ et $\mathcal{P}_2$ sont toutes deux vraies, on écrira plutôt « $\mathcal{P}_1$ ou bien $\mathcal{P}_2$ » ou « soit $\mathcal{P}_1$, soit $\mathcal{P}_2$ ».

non$(\exists\, x, \mathcal{P}(x)) \Longleftrightarrow \forall\, x,\ \mathrm{non}\,\mathcal{P}(x)$.

non$(\exists\, x \in A, \mathcal{P}(x)) \Longleftrightarrow \forall\, x \in A,\ \mathrm{non}\,\mathcal{P}(x)$.

non(non$\mathcal{P}$) $= \mathcal{P}$.

non$(\mathcal{P}_1$ et $\mathcal{P}_2) \Longleftrightarrow$ non$\mathcal{P}_1$ ou non$\mathcal{P}_2$.

non$(\mathcal{P}_1$ ou $\mathcal{P}_2) \Longleftrightarrow$ non$\mathcal{P}_1$ et non$\mathcal{P}_2$.

$\emptyset$: ensemble vide.

$A \subset B$ [(*2)] : A est un sous-ensemble (on dit aussi « une partie ») de l'ensemble B (« A est inclus dans B »), c.-à-d. que tout élément de A appartient à B :

$$A \subset B \Longleftrightarrow \forall\, x \in A, x \in B.$$

$\{a, b, \ldots\}$: ensemble formé par les éléments a, b, etc.

$\{x \mid \mathcal{P}(x)\}$: ensemble des x satisfaisant la propriété $\mathcal{P}(x)$.

$A \cup B$ (« union de A et B », « A union B ») : $\{x \mid x \in A$ ou $x \in B\}$.

$A \cap B$ (« intersection de A et B », « A inter B ») : $\{x \mid x \in A$ et $x \in B\}$.

$A \setminus B$ (« différence de A et B », « A moins B »), $\complement_A B$ (« complément de B dans A ») : $\{x \mid x \in A$ et $x \notin B\}$.

Ordre décroissant de priorité des symboles relatifs aux ensembles : $\setminus$, $\cap$, $\cup$, $\subset$.

I. . Ensembles de nombres

I. .1. Notations

$\mathbb{N} = \{0, 1, 2, \ldots\}$: ensemble des nombres entiers naturels (c.-à-d. non strictement négatifs).

$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$: ensemble des nombres entiers relatifs (c.-à-d. positifs ou négatifs).

$\mathbb{Q}$: ensemble des nombres rationnels, c.-à-d. des fractions a/b avec $a \in \mathbb{Z}$ et $b \in \mathbb{Z} \setminus \{0\}$.

$\mathbb{R} =]-\infty, \infty[$: ensemble des nombres réels (ou « droite réelle »).

$\mathbb{C}$: ensemble des nombres complexes.

$\mathbb{K}$: $\mathbb{R}$ ou $\mathbb{C}$.

$\mathbb{Q}^+ = \{x \in \mathbb{Q} \mid x \geqslant 0\}$.

$\mathbb{Z}^- = \{x \in \mathbb{Z} \mid x \leqslant 0\}$.

$\mathbb{C}^* = \{x \in \mathbb{C} \mid x \neq 0\}$.

$\mathbb{R}^n = \{x = (x_1, \ldots, x_n)$, où $x_i \in \mathbb{R}\}$.

I. .2. Intervalles

Une partie I de $\mathbb{R}$ est un intervalle si $\forall\, (y, z) \in I^2, \forall\, x \in \mathbb{R}, y \leqslant x \leqslant z \Rightarrow x \in I$.

$[a, b] = \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$ (intervalle fermé).

2. Dans l'usage français, « $A \subset B$ » désigne une inclusion large, c.-à-d. qu'on peut avoir $A = B$; l'inclusion stricte, excluant le cas $A = B$, peut être notée « $A \subsetneq B$ ». En revanche, dans l'usage anglais, « $A \subset B$ » désigne une inclusion stricte ; l'inclusion large est alors notée « $A \subseteq B$ ».

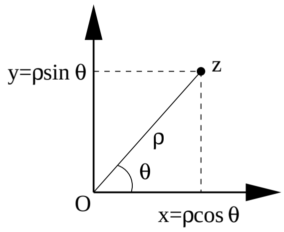


Figure I.1 – Interprétation géométrique d'un nombre complexe.

$]a, b[= \{x \in \mathbb{R} \mid a < x < b\}$ (intervalle ouvert).

$]a, b] = \{x \in \mathbb{R} \mid a < x \leqslant b\}$, $[a, b[= \{x \in \mathbb{R} \mid a \leqslant x < b\}$ (intervalles semi-ouverts).

$[\![a, b]\!]$ (notation non standard) : $\{x \in \mathbb{Z} \mid a \leqslant x \leqslant b$ et $(a, b) \in \mathbb{Z}^2\}$.

Ces notations impliquent que $a \leqslant b$.

I. .3. Nombres complexes

On note $\mathbb{C}$ l'ensemble $\mathbb{R}^2$ muni de l'addition et de la multiplication définies ci-dessous. Soit $z = (x, y)$ un nombre complexe : les réels x et y sont la *partie réelle* et la *partie imaginaire* de z ; on les note « $\operatorname{Re} z$ » et « $\operatorname{Im} z$ ».

Par définition, si $z' = (x', y')$ est un autre nombre complexe,

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

et [*3]

$$z \times z' = (x \times x' - y \times y', x \times y' + x' \times y).$$

Les symboles d'opération définis à gauche du signe « = » agissent sur des nombres complexes ; ceux utilisés à droite sont l'addition et la multiplication usuelles dans $\mathbb{R}$.

L'ensemble $\mathbb{R}$ s'identifie à l'ensemble des nombres complexes de partie imaginaire nulle. En effet, si $y = y' = 0$, on a $z + z' = (x + x', 0)$ et $z \times z' = (x \times x', 0)$. On peut donc sans danger écrire $(x, 0) = x$ et utiliser les mêmes symboles pour les opérations entre complexes que pour celles entre réels.

Avec ces opérations, $\mathbb{C}$ possède comme $\mathbb{R}$ les propriétés suivantes : commutativité, associativité, distributivité de la multiplication par rapport à l'addition, mêmes éléments neutres ($(0, 0) = 0$ pour l'addition et $(1, 0) = 1$ pour la multiplication). En revanche, la relation d'ordre $\leqslant$ n'est pas définie sur $\mathbb{C}$: $z = z'$ et $z \neq z'$ ont un sens, mais pas $z \leqslant z'$ ou $z > z'$.

On note « i » l'élément $(0, 1)$. On a $z = (x, y) = (x, 0) + \mathrm{i} \times (y, 0)$. On peut donc écrire $z = x + \mathrm{i}\, y$. Par ailleurs, $\mathrm{i}^2 = (0, 1) \times (0, 1) = (-1, 0) = -1$. On retrouve bien, en utilisant la règle de distributivité de la multiplication par rapport à l'addition, que

$$z\, z' = (x + \mathrm{i}\, y) \times (x' + \mathrm{i}\, y') = x\, x' + \mathrm{i}\, x\, y' + \mathrm{i}\, y\, x' + \mathrm{i}^2\, y\, y' = (x\, x' - y\, y') + \mathrm{i} \times (x\, y' + x'\, y).$$

Tout z a un opposé, $-z = -x + \mathrm{i} \times (-y)$. On peut donc définir la soustraction entre complexes par $z - z' := z + (-z')$ [*4]. En particulier, $-z = -x - \mathrm{i}\, y$.

De même, tout $z \in \mathbb{C}^*$ a un inverse,

$$z^{-1} = \frac{x - \mathrm{i}\, y}{x^2 + y^2}.$$

On peut donc définir la division de $z \in \mathbb{C}$ par $z' \in \mathbb{C}^*$: $z/z' := z\, z'^{-1}$. En particulier, $z^{-1} = 1/z$.

En étendant à $\mathbb{C}$ les définitions usuelles de l'exponentielle dans $\mathbb{R}$, tout $z \in \mathbb{C}$ peut s'écrire $z = \rho\, \mathrm{e}^{\mathrm{i}\,\theta}$, où $\rho \in \mathbb{R}^+$, $\theta \in \mathbb{R}$ et θ est défini à un multiple entier de $2\,\pi$ près. Le réel ρ est le *module* de z, noté $|z|$, et l'unique θ dans $[0, 2\,\pi[$ est son *argument*, $\arg z$. On a $x = \rho \cos\theta$ et $y = \rho \sin\theta$ (cf. fig. I.1).

3. La multiplication entre deux nombres peut également être notée $x \cdot x'$ ou $x\, x'$, sauf si x et x' sont des valeurs numériques.
4. Le « – » à gauche du « = » est le symbole de soustraction, tandis qu'il note l'opposé à droite.

Formules d'Euler :

$$\begin{aligned} \mathrm{e}^{\mathrm{i}\,\theta} &= \cos\theta + \mathrm{i}\,\sin\theta, \\ \cos\theta &= \frac{\mathrm{e}^{\mathrm{i}\,\theta}+\mathrm{e}^{-\mathrm{i}\,\theta}}{2}, \\ \sin\theta &= \frac{\mathrm{e}^{\mathrm{i}\,\theta}-\mathrm{e}^{-\mathrm{i}\,\theta}}{2\,\mathrm{i}}. \end{aligned}$$

Produit de $z_1 = \rho_1\,\mathrm{e}^{\mathrm{i}\,\theta_1}$ et $z_2 = \rho_2\,\mathrm{e}^{\mathrm{i}\,\theta_2}$:

$$z = z_1\,z_2 = \rho_1\,\rho_2\,\mathrm{e}^{i\cdot(\theta_1+\theta_2)},$$

donc le module de z est $\rho = \rho_1\,\rho_2$ et son argument est $\theta = \theta_1 + \theta_2 \pmod{2\,\pi}$.

Complexe conjugué de z, noté $\overline{z}$ ou z^* :

$$\overline{z} = x - \mathrm{i}\,y = \rho\,\mathrm{e}^{-\mathrm{i}\,\theta}.$$

Module carré de z :

$$|z|^2 = z\,\overline{z} = x^2 + y^2 = \rho^2.$$

Inverse de z :

$$\frac{1}{z} = \frac{\overline{z}}{|z|^2} = \rho^{-1}\,\mathrm{e}^{-\mathrm{i}\,\theta}.$$

(On obtient bien que $z\,z^{-1} = 1$.) En particulier,

$$\mathrm{i} = \mathrm{e}^{\mathrm{i}\,\pi/2} \quad \text{et} \quad \overline{\mathrm{i}} = \mathrm{i}^{-1} = \mathrm{e}^{-\mathrm{i}\,\pi/2}.$$

Inégalités triangulaires : $\big||a| - |b|\big| \leqslant |a+b| \leqslant |a| + |b|$ pour a et b réels ou complexes.

Si $n \in \mathbb{N}$ est pair et $x \in \mathbb{R}^+$, la racine principale n-ième de x (celle dont il s'agit quand on dit « *la* racine n-ième », et que l'on note $\sqrt[n]{x}$ ou $x^{1/n}$) est l'unique $r \in \mathbb{R}^+$ tel que $r^n = x$. *Une* autre racine réelle de cette équation est $r' = -r$. Si $x \in \mathbb{R}^{-*}$, cette équation n'a pas de solution réelle.

Si $n \in \mathbb{N}$ est impair et $x \in \mathbb{R}$, l'équation $r^n = x$ admet toujours une solution réelle r unique, positive ou négative, que l'on note encore $\sqrt[n]{x}$ ou $x^{1/n}$.

Si $n \in \mathbb{N}$ et $z \in \mathbb{C}$, l'équation $r^n = z$ admet n solutions distinctes dans $\mathbb{C}$. Le plus simple est d'écrire z sous la forme $z = |z|\,\mathrm{e}^{\mathrm{i}\,\theta}$. Les solutions sont $r_k = |z|^{1/n}\,\mathrm{e}^{\mathrm{i}\cdot(\theta/n + 2\,k\,\pi/n)}$, avec $k \in [\![0, n-1]\!]$.

I. . Polynômes

Un polynôme de degré $n \in \mathbb{N}$ en la variable $x \in \mathbb{K}$ est une fonction de la forme $P(x) = a_0 + a_1\,x + \cdots + a_n\,x^n$, avec $(a_0, \ldots, a_n) \in \mathbb{K}^n$ et $a_n \neq 0$.

Par exemple, la puissance n-ième de $(x+a)$ est un polynôme de degré n. En effet, par la formule du binôme, on obtient

$$\begin{aligned} (a+x)^n &= a^n + n\,a^{n-1}\,x + \frac{n\cdot(n-1)}{2}\,a^{n-2}\,x^2 + \cdots + x^n \\ &= \sum_{p=0}^{n} \binom{n}{p}\,a^{n-p}\,x^p, \end{aligned}$$

où le coefficient binomial $\binom{n}{p}$ (parfois noté C_n^p (*5)) vaut (*6)

$$\binom{n}{p} = \frac{n \times (n-1) \times \cdots \times (n-p+1)}{1 \times 2 \times \cdots \times (p-1) \times p} = \frac{n!}{p!\cdot(n-p)!}.$$

5. C_n^p est le nombre de *combinaisons* différentes de p éléments choisis dans un ensemble en contenant n. Si l'on tient compte de l'ordre dans lequel les p éléments sont choisis (c.-à-d. en distinguant par exemple (a, b, c) de (a, c, b)), on obtient le nombre d'*arrangements* différents de p éléments choisis parmi n, $\mathrm{A}_n^p = n!/(n-p)!$.
6. $p! = 1 \times 2 \times 3 \times \cdots \times (p-1) \times p$ est la factorielle de p. Par convention, $0! = 1$.

Les coefficients binomiaux obéissent à la relation de récurrence [*7]

$$\binom{n}{p} = \binom{n-1}{p-1} + \binom{n-1}{p}.$$

Cette relation traduit le fait que $(a+x)^n = (a+x)\cdot(a+x)^{n-1}$.

Si $P(x_1) = 0$, on dit que x_1 est un *zéro* de P (ou une *racine* de l'équation $P(x) = 0$), et on peut factoriser P de façon *unique* sous la forme $P(x) = (x - x_1)\,Q(x)$, où Q est un polynôme de degré $n-1$. Si $\mathbb{K} = \mathbb{R}$, un polynôme de degré n en x a au plus n zéros réels.

D'après le « théorème fondamental de l'algèbre », si $\mathbb{K} = \mathbb{C}$, un polynôme P de degré n a exactement n zéros $x_i \in \mathbb{C}$ ($i \in [\![1, n]\!]$), dont certains peuvent être identiques (zéros multiples). On peut donc écrire P de façon unique sous la forme

$$P(x) = a_n \prod_{i=1}^{n} (x - x_i).$$

I. . Structures algébriques

Les ensembles que nous considérerons dans ce cours sont généralement des espaces vectoriels normés, c.-à-d. des espaces vectoriels munis d'une norme. Avant de donner les définitions d'un espace vectoriel et d'une norme, nous devons rappeler celles d'un groupe et d'un corps.

Définition 1 (groupe)
Soient G un ensemble et $+$ une loi de composition interne (c.-à-d. une application de $G \times G$ dans G).

$(G, +)$ est un groupe si les propriétés suivantes sont vérifiées :

- $\forall\,(x, y, z) \in G^3$, $(x + y) + z = x + (y + z)$ (associativité) ;
- La loi « + » admet un élément neutre dans G, que nous noterons « 0 », c.-à-d. tel que $\forall\,x \in G$, $x + 0 = 0 + x = x$.
- Tout élément x de G admet un opposé, noté $(-x)$, c.-à-d. tel que $x + (-x) = (-x) + x = 0$.

$(G, +)$ est un groupe commutatif si, en outre, $\forall\,(x, y) \in G^2$, $x + y = y + x$. ⌟

Remarques
Nous avons noté « + » la loi interne, « 0 » l'élément neutre et « $(-x)$ » l'opposé car l'addition usuelle dans $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$ (mais pas $\mathbb{N}$!) fait de chacun de ces ensembles un groupe, mais une autre notation est parfois préférable. Ainsi, les matrices inversibles constituent un groupe pour la multiplication matricielle ; l'élément neutre pour cette opération est la matrice identité, pas la matrice nulle ; enfin, l'opposé de x, que l'on nomme « inverse » dans le cas d'une notation multiplicative, est noté « x^{-1} ». ⌟

Définition 2 (corps)
Soient K un ensemble et $+$ et $\times$ deux lois de compositions internes.

$(K, +, \times)$ est un corps si les propriétés suivantes sont vérifiées :

- $(K, +)$ est un groupe commutatif ;
- $\forall\,(x, y, z) \in K^3$, $x \times (y \times z) = (x \times y) \times z$ (associativité de « × ») ;
- $\forall\,(x, y, z) \in K^3$, $x \times (y + z) = x \times y + x \times z$ et $(x + y) \times z = x \times y + y \times z$ (distributivité de « × » par rapport à « + ») ;
- La loi « × » admet un élément neutre, u, c.-à-d. tel que $\forall\,x \in K$, $x \times u = u \times x = x$;
- $\forall\,x \in K \setminus \{0\}$, où 0 est l'élement neutre pour l'addition, il existe un élément, noté x^{-1} et appelé « inverse », tel que $x \times x^{-1} = x^{-1} \times x = u$. ⌟

Remarques
Nous avons noté « × » la loi interne et « x^{-1} » l'inverse car la multiplication usuelle dans les groupes $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ ou $(\mathbb{C}, +)$ (mais pas $\mathbb{N}$ et $\mathbb{Z}$!) fait de chacun d'eux un corps ; l'élément neutre de la multiplication dans ces ensembles est le nombre 1. ⌟

Définition 3 (espace vectoriel)
Soient $(K, +, \times)$ un corps commutatif et E un ensemble muni d'une loi interne (c.-à-d. de $E \times E$ dans E), $\oplus$, et d'une loi externe sur K (c.-à-d. de $K \times E$ dans E), $\otimes$.

E est un K-espace vectoriel (ou « espace vectoriel sur K ») si les propriétés suivantes sont vérifiées :

- $(E, \oplus)$ est un groupe commutatif ;

7. On peut la retrouver en construisant le triangle de Pascal.

- $\forall\,(\lambda, \mu) \in K^2, \forall\, x \in E, (\lambda + \mu) \otimes x = \lambda \otimes x \oplus \mu \otimes x$;
- $\forall\, \lambda \in K, \forall\,(x, y) \in E^2, \lambda \otimes (x \oplus y) = \lambda \otimes x \oplus \lambda \otimes y$;
- $\forall\, x \in E$, $u \otimes x = x$, où u est l'élément neutre de la loi $\times$;
- $\forall\,(\lambda, \mu) \in K^2, \forall\, x \in E, \lambda \otimes (\mu \otimes x) = (\lambda \times \mu) \otimes x$. ⌟

Remarques

- Les éléments de E sont appelés « vecteurs » et ceux de K « scalaires ».
- Pour la clarté de la définition, nous avons distingué ici $+$ et $\oplus$ ainsi que $\times$ et $\otimes$. En pratique, il n'y a aucune ambiguïté et nous abandonnerons cette distinction par la suite. Dans certains cas, K et E sont d'ailleurs confondus : $\mathbb{R}$ (resp. $\mathbb{C}$) est ainsi un $\mathbb{R}$-espace vectoriel (resp. un $\mathbb{C}$-espace vectoriel).

  $\mathbb{R}^p$ est un $\mathbb{R}$-espace vectoriel pour l'addition et la multiplication externe définies par

$$(x_1, \ldots, x_p) + (y_1, \ldots, y_p) = (x_1 + y_1, \ldots, x_p + y_p)$$

  et

$$\lambda \times (x_1, \ldots, x_p) = (\lambda \times x_1, \ldots, \lambda \times x_p).$$ ⌟

I. . Norme

Définition 4 (norme, espace vectoriel normé)

Soit E un K-espace vectoriel avec $K = \mathbb{K}$.

Une norme $\|\ldots\|$ est une application de E dans $\mathbb{R}^+$ telle que

$$\begin{aligned} \forall\, x \in E, &\quad \|x\| = 0 \Leftrightarrow x = 0, \\ \forall\,(\lambda, x) \in K \times E, &\quad \|\lambda\, x\| = |\lambda|\, \|x\|, \\ \text{et } \forall\,(x, y) \in E^2, &\quad \|x + y\| \leqslant \|x\| + \|y\|. \end{aligned}$$

Un espace vectoriel normé est un espace vectoriel muni d'une norme. ⌟

Remarques

La valeur absolue et le module sont des normes dans $\mathbb{R}$ et $\mathbb{C}$. Dans $\mathbb{R}^p$, on peut utiliser la norme euclidienne, $\|(x_1, \ldots, x_p)\| = (x_1^2 + \cdots + x_p^2)^{1/2}$. D'autres normes peuvent d'ailleurs être définies sur ces espaces, mais les théorèmes énoncés dans ce cours resteront valables : toutes les normes sont en effet équivalentes dans un espace vectoriel de dimension finie. ⌟

Chapitre II

S ´

Une suite est une application u de $\mathbb{N}$ dans un ensemble E. Pour mettre en évidence le fait que l'ensemble de départ est $\mathbb{N}$, le n-ième terme est noté u_n plutôt que $u(n)$; la suite elle-même est notée $n \mapsto u_n$, $(u_n)_{n\in\mathbb{N}}$, (u_n), voire, de manière abusive, u_n tout court.

Les ensembles E auxquels nous nous intéresserons sont des espaces vectoriels normés, c.-à-d. des espaces vectoriels dotés d'une norme, en particulier $\mathbb{R}$.

Une suite peut être définie par l'expression explicite de son terme général u_n en fonction de n. Elle peut également l'être par une relation de récurrence du genre $u_n = f(u_{n-p}, \ldots, u_{n-1})$, complétée par la donnée des p premiers termes pour permettre de « démarrer » la récurrence.

Exemple 1
Suites données explicitement :
1. $u_n = a\, r^n$ (suite géométrique de raison r) ;
2. $u_n = \sin(n\, \pi/3)$;
3. $u_n = 1/n^2$. ⌟

Exemple 2
Suites données par une récurrence à un terme :
1. $u_{n+1} = u_n + r$ (suite arithmétique de raison r) ;
2. $u_{n+1} = r\, u_n$ (suite géométrique de raison r) ;
3. $u_{n+1} = a\, u_n \cdot (1 - u_n)$ (application logistique). ⌟

Exemple 3
Suite donnée par une récurrence à deux termes : la *suite de Fibonacci*,

$$u_n = u_{n-1} + u_{n-2}, \tag{1}$$

avec les conditions initiales $u_1 = 1$ et $u_2 = 1$. Explicitement,

$$(u_n) = (1, 1, 2, 3, 5, 8, 13, \ldots). \tag{2}$$

Cette suite donne, entre autres, le nombre de spirales concentriques observées dans certains végétaux (marguerites, ananas, pommes de pin. . .). ⌟

En physique, les suites peuvent apparaître dans des contextes variés :
- Approximations de taille croissante (nombre de particules, dimensions géométriques, . . .) d'un système ;
- Résolution numérique par itération d'une équation algébrique ; les u_n sont les approximations successives d'une racine d'un polynôme, par exemple (méthode de Newton. . .) ;
- Résolution numérique par itération d'une équation différentielle ; les u_n décrivent par exemple les positions successives à des intervalles de temps réguliers, etc.

Dans tous ces cas, on souhaite savoir si la suite a une limite quand l'indice $n \to \infty$ et quelle est cette limite.

II. . Limite d'une suite

Définition 5 (limite finie)
La suite $(u_n) \in E$ tend vers une limite finie ℓ, ou *converge* vers la limite ℓ, ce que l'on écrit $\lim_{n\to\infty} u_n = \ell$ ou $u_n \xrightarrow[n\to\infty]{} \ell$, si la différence $u_n - \ell$ tend vers zéro quand n tend vers l'infini, c.-à-d. si et seulement si

$$\forall\, \epsilon (\in \mathbb{R}) > 0, \exists\, N \in \mathbb{N}, \forall\, n \in \mathbb{N}, n \geqslant N \Rightarrow \|u_n - \ell\| \leqslant \epsilon. \tag{3}$$

⌟

Autrement dit, *tous* les u_n sont arbitrairement près de ℓ, dès que n est suffisamment grand. La définition (3) se lit mot à mot « pour tout ϵ *strictement positif* (sous-entendu aussi petit qu'on veut), il existe un nombre N (sous-entendu suffisamment grand, et qui a le droit de dépendre de ϵ) tel que pour tout n plus grand que N, la distance entre u_n et ℓ soit inférieure en norme à ϵ ».

Si $E = \mathbb{R}^p$, montrer que la suite $n \mapsto u_n = ({u_n}^{(1)}, \ldots, {u_n}^{(p)}) \in \mathbb{R}^p$ converge vers la limite $\ell = (\ell^{(1)}, \ldots, \ell^{(p)})$ équivaut à démontrer que les suites $n \mapsto ({u_n}^{(j)}) \in \mathbb{R}$ convergent vers $\ell^{(j)}$ pour tout $j \in [\![1, p]\!]$.

Exemple 4
La suite $n \mapsto u_n = 1/n^2$ tend vers zéro. En effet, pour $\epsilon > 0$ donné, posons $N = \mathrm{E}(1/\sqrt{\epsilon}) + 1$, où « E » désigne la partie entière. En mettant $\ell = 0$, on a bien $|u_n - \ell| = 1/n^2 \leqslant 1/N^2 \leqslant \epsilon$ pour tout $n \geqslant N$. ⌟

Une suite peut ne pas converger, soit qu'elle ait un comportement oscillant, telle $n \mapsto \sin(n\,\pi/3)$, soit qu'elle tende vers l'infini, comme $n \mapsto \sqrt{n}$, soit qu'elle combine ces deux comportements, comme $n \mapsto \tan(n\,\pi/4 + 1/n)$.

Définition 6 (limite infinie)
- $(u_n) \in \mathbb{R}$ tend vers $+\infty$ si $\forall\, A \in \mathbb{R}, \exists\, N \in \mathbb{N}, \forall\, n \in \mathbb{N}, n \geqslant N \Rightarrow u_n \geqslant A$.
- $(u_n) \to -\infty \Leftrightarrow (-u_n) \to +\infty$. ⌟

Remarque
On réservera dans ce cours le terme de « convergence » au cas où la limite existe et est *finie*. On appliquera indifféremment les termes de « divergence » et de « non-convergence » au cas où il n'y a pas de limite et au cas où la limite existe mais est infinie. (Certains auteurs réservent le terme de « divergence » à ce dernier cas.) ⌟

Théorème 1
Si les suites $(u_n) \in E$ et $(v_n) \in E$ ont des limites finies, respectivement ℓ et ℓ', alors
- $$\forall\, \lambda \in K, \lambda\, u_n \to \lambda\, \ell\ ; \tag{4}$$
- $$u_n + v_n \to \ell + \ell'\ ; \tag{5}$$
- $$\text{si } E = \mathbb{K}, u_n\, v_n \to \ell\, \ell'\ ; \tag{6}$$
- $$\text{si } E = \mathbb{K} \text{ et } \ell \neq 0, \frac{1}{u_n} \to \frac{1}{\ell}\ ; \tag{7}$$
- $$\text{si } u_n \in \mathbb{R}^+ \text{ et } \ell = 0, \frac{1}{u_n} \to +\infty\ ; \tag{8}$$
- $$\text{si } v_n = f(u_n)\ ^{(*1)} \text{ et que } f \text{ est continue en } \ell, v_n \to f(\ell). \tag{9}$$

⌟

Théorème 2
Si $(u_n) \in \mathbb{R} \to +\infty$,
- $$\forall\, \lambda \in \mathbb{R}^*, \lambda\, u_n \to \operatorname{sgn} \lambda \times \infty\ ^{(*2)}\ ; \tag{10}$$
- $$\text{si } v_n \to v \in \mathbb{R} \text{ ou } v_n \to +\infty, u_n + v_n \to +\infty\ ; \tag{11}$$
- $$\text{si } v_n \to v \in \mathbb{R}^* \text{ ou } v_n \to +\infty, u_n\, v_n \to \operatorname{sgn}(v) \times \infty\ ; \tag{12}$$
- $$\frac{1}{u_n} \to 0^+\ ; \tag{13}$$
- $$\text{si } v_n = f(u_n)\ ^{(*3)}, \lim_{n \to +\infty} v_n = \lim_{x \to +\infty} f(x) \tag{14}$$

  (il s'agit d'une limite de suite à gauche et d'une limite de fonction à droite). ⌟

Les expressions du genre « $0 \times \infty$ », « $\infty - \infty$ », « $0/0$ », « ∞/∞ » sont en revanche des formes indéterminées et doivent être étudiées au cas par cas.

---

1. La suite v_n peut appartenir ici à un espace vectoriel F différent de E.
2. La fonction $\operatorname{sgn} x$ (« signe » de x) est définie par $\operatorname{sgn} 0 = 0$, $\operatorname{sgn} x = x/|x|$ si $x \in \mathbb{C}^*$ (donc $\operatorname{sgn} x = +1$ si $x \in \mathbb{R}^{*+}$ et $\operatorname{sgn} x = -1$ si $x \in \mathbb{R}^{*-}$), $\operatorname{sgn}(+\infty) = +1$ et $\operatorname{sgn}(-\infty) = -1$.
3. Même remarque qu'à la note du théorème précédent.

II. . Critères de convergence

Définition 7
Soit F une partie non vide de $\mathbb{R}$.

- $M \in \mathbb{R}$ est un majorant de F si, $\forall\, x \in F$, $x \leqslant M$.
- $m \in \mathbb{R}$ est un minorant de F si, $\forall\, x \in F$, $m \leqslant x$.
- S'il existe, le plus petit majorant de F s'appelle la borne supérieure de F ($\sup F$).
- S'il existe, le plus grand minorant de F s'appelle la borne inférieure de F ($\inf F$). ⌟

Théorème 3

- Toute partie majorée non vide de $\mathbb{R}$ possède une borne supérieure.
- Toute partie minorée non vide de $\mathbb{R}$ possède une borne inférieure. ⌟

Remarque
Cette propriété n'est pas vraie dans $\mathbb{Q}$: l'ensemble $\{x \in \mathbb{Q} \mid x^2 \leqslant 2\}$ n'a pas de borne supérieure dans $\mathbb{Q}$. Elle n'a pas de sens dans $\mathbb{C}$ ou $\mathbb{R}^p$ puisque ces ensembles ne sont pas ordonnés (c.-à-d. qu'on ne peut pas y définir de relation d'ordre). ⌟

Théorème 4

- Si une suite dans $\mathbb{R}$ est *croissante* et *majorée*, c.-à-d. $u_{n-1} \leqslant u_n \leqslant M$ pour un certain nombre M et pour tout n suffisamment grand, elle converge vers une limite $\ell \leqslant M$.
- De même si elle est décroissante et minorée. ⌟

Preuve
L'ensemble majoré des nombres réels (u_n) admet une borne supérieure $B \leqslant M$, qui est le plus petit nombre tel que $u_n \leqslant B$ pour tout $n \in \mathbb{N}$. Alors $\forall\, \epsilon > 0$, $\exists\, N \in \mathbb{N}$, $u_N \geqslant B - \epsilon$ (sans quoi B ne serait pas le plus petit majorant), et, par hypothèse de monotonicité, $\forall\, n \geqslant N$, $|u_n - B| = B - u_n \leqslant B - u_N < \epsilon$, donc $u_n \to \ell = B$. ⌟

Exemple 5
$u_n = 2^{1-1/n}$. On a $u_{n-1} < u_n < 2$. La suite converge vers une limite (qui est 2). ⌟

Inversement, si une suite monotone croissante n'est pas bornée (c.-à-d. si ses éléments sont arbitrairement grands), elle tend vers $+\infty$.

Théorème 5 (théorème des gendarmes)
Si $w_n \leqslant u_n \leqslant v_n\ \forall\, n \in \mathbb{N}$ et que les deux suites (v_n) et (w_n) tendent vers la même limite ℓ, (u_n) converge aussi vers ℓ. ⌟

Exemple 6
Quelles sont les limites de $n \mapsto \sin(\sqrt{2}\, n)/n^2$ et de $n \mapsto 1/(n^2 + \sin(n\,\pi/4))$? ⌟

Théorème 6 (suites adjacentes)
Si (u_n) est croissante, que (v_n) est décroissante, et que $\lim(u_n - v_n) = 0$, alors (u_n) et (v_n) convergent vers une limite commune ℓ, et $\forall\, (n, p) \in \mathbb{N}^2$, $u_n \leqslant \ell \leqslant v_p$. ⌟

Théorème 7
(u_n) converge si et seulement si la série (cf. § III) $\sum v_n$, où $v_n = u_{n+1} - u_n$, converge. ⌟

On démontre ainsi que $n \mapsto u_n = 1 + 1/2 + \cdots + 1/n - \ln n$ converge (vers un nombre appelé « constante d'Euler », habituellement noté γ et valant $\gamma \approx 0{,}577$).

Théorème 8
Soit une suite $(u_n) > 0$.

- Si $u_{n+1}/u_n \to k \in \mathbb{R}$, $u_n^{1/n} \to k$.
- Si $u_n^{1/n} \to k \in [0, 1[$, $u_n \to 0$. ⌟

Étude asymptotique

Définition 8 (notations de Landau)
Soient (u_n) et (v_n) des suites de $\mathbb{K}$, et $(v_n) \neq 0$.

- $u_n = o(v_n)$ (« u_n est un petit o de v_n » ou « u_n est négligeable devant v_n ») si $\lim(u_n/v_n) = 0$.

- $u_n = O(v_n)$ (« u_n est un grand o de v_n » ou « u_n est dominée par v_n ») si $|u_n/v_n|$ est majorée.
- $u_n \simeq v_n$ (« u_n est asymptotiquement égale (ou "équivalente") à v_n ») si $\lim(u_n/v_n) = 1$. ⌟

Remarques

- Les notations $u_n = o(v_n)$ et $u_n = O(v_n)$ sont commodes mais abusives : en effet, $u_n = o(v_n)$ et $w_n = o(v_n) \not\Rightarrow u_n = w_n$. On note parfois plus correctement $u_n \in o(v_n)$ et $u_n \in O(v_n)$.
- On note parfois $u_n = \varepsilon(v_n)$ au lieu de $u_n = o(v_n)$.
- L'équivalence entre suites est notée ici $\simeq$, conformément à la norme ISO 31-11 ; la notation $\sim$ est souvent utilisée.
- $u_n = o(v_n) \Rightarrow u_n = O(v_n)$.
- $u_n \simeq v_n \Rightarrow u_n - v_n = o(v_n)$.
- $\simeq$ est une relation d'équivalence ; elle est en particulier transitive : $u_n \simeq v_n$ et $v_n \simeq w_n \Rightarrow u_n \simeq w_n$.
- Ne pas additionner d'équivalents : ($u_n \simeq v_n$ et $u'_n \simeq v'_n$) $\not\Rightarrow u_n + u'_n \simeq v_n + v'_n$. ⌟

Théorème 9

- Si $u_n \simeq v_n$ et si $v_n \to \ell \in \mathbb{K}$, alors $u_n \to \ell$.
- Si $u_n \simeq v_n$ et si $v_n \to \pm\infty$, alors $u_n \to \pm\infty$.
- Si $u_n = o(v_n)$ et $v_n \to \ell \in \mathbb{K}$, alors $u_n \to 0$. ⌟

Exemple 7

La suite $n \mapsto v_n = 1/n^2$ converge vers 0 ; la suite $n \mapsto u_n = 1/(n^2 + \sin(n\,\pi/4))$ est asymptotiquement égale à (v_n) (pourquoi ?), donc converge aussi vers 0. Noter que l'exemple précédent, $n \mapsto u_n = 2^{1-1/n}$, relève aussi de ce cas, pour la suite constante $n \mapsto v_n = 2$. ⌟

Exemple 8

Un autre exemple important est $u_n = (1 + a/n)^n$, qui tend vers e^a. En effet, pour n grand, $u_n = \mathrm{e}^{n\ln(1+a/n)} \simeq \mathrm{e}^a$ en développant le logarithme au premier ordre. ⌟

Les notations de Landau serviront notamment pour les séries.

Suites récurrentes

Théorème 10 (points fixes)

Soit $f : \mathbb{R}^p \to \mathbb{R}$ une fonction et (u_n) une suite définie par la relation de récurrence $u_n = f(u_{n-p}, \ldots, u_{n-1})$.

Si (u_n) admet une limite l et si f est continue au voisinage du point $(\ell, \ldots, \ell)$, alors $\ell = f(\ell, \ldots, \ell)$. On dit que ℓ est un *point fixe* de f. ⌟

Exemple 9

La suite définie par $v_n = u_n/u_{n-1}$, où (u_n) est la suite de Fibonacci (2), satisfait $v_n = 1 + 1/v_{n-1}$. Son éventuelle limite est donc l'une des racines de l'équation $x^2 = x + 1$, soit $x' = (1 + \sqrt{5})/2$ ou $x'' = (1 - \sqrt{5})/2$, et ce ne peut être que celle des deux qui est positive, soit x' [*4]. Il reste à démontrer qu'effectivement $v_n \to v = x'$. ⌟

Noter que pour une suite définie par $u_n = f(u_{n-p}, \ldots, u_{n-1})$, l'existence d'une racine de l'équation $\ell = f(\ell, \ldots, \ell)$ ne garantit en rien la convergence de u_n vers ℓ. C'était le cas ci-dessus pour la deuxième racine x''.

Exemple 10 (suite arithmético-géométrique)

La recherche de poins fixes peut être utile même lorsque la suite ne converge pas. Considérons la suite définie par $u_{n+1} = a\,u_n + b$ (avec $a \neq 1$). Un point fixe est défini par $\ell = a\,\ell + b$, soit $\ell = b/(1-a)$. La suite $v_n = u_n - \ell$ est une suite géometrique de raison a (on a $v_{n+1} = a\,v_n$). Si $|a| \geqslant 1$, v_n est une suite divergente, donc u_n aussi. L'expression $u_n = a^n \cdot (u_0 - \ell) + \ell$ reste néanmoins vraie et permet de calculer directement n'importe quel terme de la suite, sans utiliser la relation de récurrence. ⌟

Pour une suite définie par une récurrence à un terme, la méthode suivante permet souvent de déterminer son sens de variation, ce qui est utile pour étudier sa convergence :

Si $u_n = f(u_{n-1})$, que f est monotone sur un intervalle I tel que $f(I) \subset I$ et que $u_0 \in I$, on peut distinguer deux cas :

f **croissante.** Si $u_1 \geqslant u_0$, (u_n) est croissante ; si $u_1 \leqslant u_0$, (u_n) est décroissante ;

f **décroissante.** $f \circ f$ est croissante : on revient donc au cas précédent en considérant séparément les suites $(u_{2\,n})$ et $(u_{2\,n+1})$.

4. Le nombre x', le fameux « nombre d'or », était connu des Grecs pour ses propriétés remarquables.

Calcul numérique

L'usage de la calculette est évidemment la solution la plus tentante pour déterminer rapidement si une suite a ou non une limite. Dans certains cas, cette méthode peut conduire à des résultats incorrects en raison des erreurs d'arrondi inhérentes au calcul numérique.

Exemple 11
Soit la suite définie par $u_{n+1} = 100\, u_n - 9$ et $u_1 = 1/11$. On vérifie immédiatement que la suite est constante : $\forall\, n, u_n = 1/11$, tandis qu'une calculette, après un certain nombre d'itérations, conclut à une divergence. ⌟

II. . Suites de Cauchy

Définition 9
Une suite qui satisfait la condition

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}, \forall\, (n, m) \in \mathbb{N}^2, (n, m) \geqslant N \Longrightarrow \|u_n - u_m\| \leqslant \epsilon \tag{15}$$

est appellée *suite de Cauchy*. ⌟

Théorème 11
Toute suite convergente est une suite de Cauchy. ⌟

Preuve
Si (u_n) converge vers ℓ, on a

$$\forall\, \epsilon' > 0, \exists\, N, \forall\, (n, m) \geqslant N, \|u_n - \ell\| \leqslant \epsilon' \text{ et } \|u_m - \ell\| \leqslant \epsilon'.$$

Or, d'après l'inégalité triangulaire,

$$\|u_n - u_m\| = \|(u_n - \ell) + (\ell - u_m)\| \leqslant \|u_n - \ell\| + \|\ell - u_m\| \leqslant 2\, \epsilon'.$$

Il suffit donc de prendre $\epsilon' = \epsilon/2$. ⌟

En particulier, si (u_n) converge, $\|u_n - u_{n+1}\|$ peut être rendu aussi petit qu'on veut. Cette conséquence de la condition de convergence est donc une condition *nécessaire* de convergence. Si elle est satisfaite, nous ne sommes pas sûrs que la suite converge. Mais si elle n'est pas satisfaite, nous sommes sûrs que la suite (u_n) ne converge pas !

Théorème 12
Si $\|u_n - u_{n+1}\|$ ne tend pas vers zéro, la suite (u_n) ne converge pas. ⌟

La condition de convergence (3) suppose connue la limite ℓ de la suite. Dans certains espaces, toute suite de Cauchy est convergente, ce qui permet de montrer leur convergence même si leur limite n'est pas connue.

Définition 10 (espace complet)
Un espace vectoriel normé E est dit complet si et seulement si toute suite de Cauchy dans E converge dans E. ⌟

Théorème 13
Les espaces vectoriels normés $\mathbb{R}$, $\mathbb{C}$ et $\mathbb{R}^p$ sont complets. ⌟

En revanche, l'ensemble $\mathbb{Q}$ des rationnels n'est pas complet : la suite définie par $u_n = 1 + 1/(1 + u_{n-1})$ et $u_1 = 1$ est une suite de Cauchy de rationnels. Elle ne converge toutefois pas dans $\mathbb{Q}$, puisque sa limite dans $\mathbb{R}$ est $\sqrt{2}$, un nombre irrationnel.

Les suites de Cauchy permettent d'ailleurs de définir l'ensemble $\mathbb{R}$ à partir de $\mathbb{Q}$. On dit que deux suites de Cauchy dans $\mathbb{Q}$, (u_n) et (v_n), sont équivalentes, ce que l'on note $(u_n)\ \mathcal{R}\ (v_n)$, si $\lim(u_n - v_n) = 0$ (*5). La relation $\mathcal{R}$ est ce qu'on appelle une relation d'équivalence. La classe d'équivalence de (u_n) est l'ensemble des suites équivalentes à (u_n). L'ensemble de ces classes d'équivalence constitue $\mathbb{R}$. À titre d'exemple, la classe d'équivalence de $\sqrt{2}$ contient entre autres la suite $(u_1 = 1, \ldots, u_n = 1 + 1/(1 + u_{n-1}), \ldots)$ et la suite $(v_1 = 1, v_2 = 1{,}4, v_3 = 1{,}41, v_4 = 1{,}414, \ldots)$ définie par le développement décimal de $\sqrt{2}$.

5. Il faut prendre ϵ dans $\mathbb{Q}$ dans la définition 5.

Chapitre III

S´ ´

III. . Introduction

Soit (u_n) une suite de $\mathbb{N}$ dans un espace vectoriel E. La *somme partielle*

$$S_n = u_0 + u_1 + u_2 + \cdots + u_n \tag{1}$$

définit une nouvelle suite, (S_n), notée $\sum u_n$ et appelée *série* de terme général u_n, qui à tout n de $\mathbb{N}$ associe $S_n \in E$.

Si cette série converge, on note

$$S_\infty = \sum_{n=0}^{\infty} u_n := \lim_{n\to\infty} S_n \tag{2}$$

sa somme. Le reste $R_n := S_\infty - S_n = \sum_{k=n+1}^{\infty}$ est alors défini et tend vers 0 quand $n \to \infty$.

On notera parfois de manière abusive $\sum_{n=0}^{\infty} u_n = \pm\infty$ quand $\lim_{n\to\infty} S_n = \pm\infty$.

Exemple 12 (série géométrique)
La série de terme général $u_n = a^n$ ($a \in \mathbb{C}$) a des sommes partielles données par

$$S_n = 1 + a + a^2 + \cdots a^n = \frac{1 - a^{n+1}}{1 - a} \text{ si } a \neq 1. \tag{3}$$

La somme partielle S_n a pour limite $a/(1-a)$ si $|a| < 1$ (puisqu'alors $a^n \to 0$), et qu'elle ne converge si $|a| > 1$ (puisque $|a|^n \to \infty$). Si $a = 1$, $S_n = n$, série évidemment divergente. Si $|a| = 1$, mais que $a \neq 1$, on a $a = \mathrm{e}^{\mathrm{i}\,\theta}$ avec $\theta \neq 2\,k\,\pi\ \forall\, k \in \mathbb{Z}$, soit $a^{n+1} = \mathrm{e}^{\mathrm{i}\cdot(n+1)\,\theta}$: a^{n+1} ne converge pas (le point représentatif de a^{n+1} tourne sur le cercle de rayon 1 du plan complexe), donc S_n non plus. ⌟

Théorème 14
Si $\sum u_n$ et $\sum v_n$ sont convergentes,

- $\sum(u_n + v_n)$ converge vers $\sum_{n=0}^{\infty} u_n + \sum_{n=0}^{\infty} v_n$;
- $\forall\, \lambda \in \mathbb{K}$, $\sum \lambda\, u_n$ converge vers $\lambda \sum_{n=0}^{\infty} u_n$. ⌟

Théorème 15
Si $\sum u_n$ est convergente et $\sum v_n$ est divergente, $\sum(u_n + v_n)$ est divergente. ⌟

Par contre, si $\sum u_n$ et $\sum v_n$ sont toutes deux divergentes, $\sum(u_n + v_n)$ peut être convergente ou divergente selon les cas.

Notons que prouver la convergence d'une série est une chose, déterminer quelle est sa somme en est une autre. Il s'agit souvent d'une tâche ardue. On verra au chapitre VII et en TD des exemples de séries convergeant vers des valeurs comme $\pi^2/6$, $\pi^2/4$, etc. En voici un autre :

Exemple 13
La série

$$\sum_{k=0}^{\infty} \left(\frac{1}{16}\right)^{k} \cdot \left(\frac{4}{8\,k+1} - \frac{2}{8\,k+4} - \frac{1}{8\,k+5} - \frac{1}{8\,k+6}\right) \tag{4}$$

converge (démonstration facile !) très vite (comme $1/(16^k\, k^2)$) vers. . . π. Vérifiez-le avec votre calculette ! ⌟

III. .1. Modification de la série

Théorème 16 (regroupement de termes consécutifs)
Soit $\sum u_n$ une série convergente et $\phi\colon \mathbb{N} \to \mathbb{N}$ une application strictement croissante telle que $\phi(0) = 0$. Notons $v_n = \sum_{k=\phi(n)}^{\phi(n+1)-1} u_k$ un regroupement de termes u_k consécutifs.

La série $\sum v_n$ converge et a la même limite que $\sum u_n$. ⌟

Remarque
La convergence de $\sum u_n$ entraîne celle de $\sum v_n$, mais la réciproque est fausse. Par exemple, la série de terme général $u_n = (-1)^n$ est divergente, puisque les sommes partielles oscillent entre $S_{2\,n-1} = -1$ et $S_{2\,n} = 0$, tandis que la série de terme général $v_n = u_{2\,n-1} + u_{2\,n} = 0$, où l'on a regroupé chaque terme d'ordre impair avec le terme consécutif, est évidemment convergente ! ⌟

En général, si on change l'ordre des termes (permutation) ou qu'on les regroupe, on ne peut plus rien affirmer sur la nature (convergente ou divergente) de la série originelle.

Par contre, si on ne modifie, permute ou regroupe qu'un nombre *fini* de termes, on ne change pas la nature (convergente ou divergente) de la série, ni, dans le cas d'une permutation ou d'un regroupement, sa limite.

Théorème 17 (supression de termes nuls)
Soit $\sum u_n$ une série et (v_n) la suite (u_n) à laquelle on a retranché les termes nuls.

La série $\sum v_n$ converge si et seulement si $\sum u_n$ converge. Si $\sum u_n$ a une limite (éventuellement infinie), $\sum v_n$ a la même limite, et inversement. ⌟

III. .2. Divergence grossière

On a vu que toute suite convergente est de Cauchy (cf. § II, théor. 11). Appliqué à une série, ceci permet de dire que si $\sum u_n$ est convergente, alors

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}, \forall\, (m, n) \in \mathbb{N}^2, n > m \geqslant N \Longrightarrow \underbrace{\left\| \sum_{k=m+1}^{n} u_k \right\|}_{S_n - S_m} \leqslant \epsilon.$$

On a en particulier le théorème suivant :

Théorème 18
Si $\sum u_n$ est convergente, $u_n = S_n - S_{n-1}$ tend nécessairement vers 0. Inversement, si $\lim u_n \neq 0$, la série $\sum u_n$ est divergente ; on dit alors qu'elle est *grossièrement divergente*. ⌟

La condition $(u_n \to 0)$ est nécessaire mais *pas suffisante* pour que $\sum u_n$ converge.Nous allons en effet rencontrer de nombreux exemples de séries dont le terme général u_n tend vers zéro quand $n \to \infty$, mais pas assez vite pour assurer la convergence de la série $\sum u_n$. Exemple : la *série harmonique* $\sum 1/n$, dont on va montrer qu'elle diverge.

III. . Séries à termes positifs

On suppose dans cette section que u_n est une suite de réels *positifs* (éventuellement à partir d'un ordre fini, ce qui ne modifie rien à la convergence). La suite $S_n = \sum_{k=0}^{n} u_k$ (somme partielle) est croissante, donc on a le théorème suivant :

Théorème 19
S_n converge si et seulement si S_n est majorée. Sinon, $S_n \to \infty$. ⌟

Théorème 20 (comparaison série-intégrale)
Soit $f\colon [a, +\infty[\to \mathbb{R}^+$ une fonction *décroissante* et $u_n = f(n)$:

- $\sum u_n$ converge $\Longleftrightarrow$ l'intégrale impropre (cf. § V.D.2) $\displaystyle\int_a^{+\infty} f$ existe ;
- $\sum u_n$ diverge $\Longrightarrow \forall\, n_0 \in \mathbb{N} \geqslant a,\ \displaystyle\lim_{n\to+\infty} \left(\sum_{k=n_0}^{n} u_k - \int_a^n f \right)$ existe. ⌟

Exemple 14 (séries de Riemann)
La série $\sum_{n\geqslant 1}(1/n^\alpha)$ converge si et seulement si l'intégrale $\int_{t=1}^{\infty} t^{-\alpha}\,\mathrm{d}t$ existe.

$$\int_{t=1}^{x} t^{-\alpha}\,\mathrm{d}t = \begin{cases} \frac{x^{1-\alpha}-1}{1-\alpha} & \text{si } \alpha \neq 1, \\ \ln x & \text{si } \alpha = 1. \end{cases}$$

En faisant tendre x vers l'infini, on constate que l'intégrale ne converge que si $\alpha > 1$, donc

$$\sum_{n\geqslant 1} \frac{1}{n^\alpha} \text{ converge si et seulement si } \alpha > 1.$$

⌟

Théorème 21
Pour la *comparaison* de séries à termes positifs, nous avons les résultats suivants (les inégalités doivent être valables à partir d'un certain rang) :

1. Si $0 \leqslant u_n \leqslant v_n$,
   - $\sum v_n$ converge $\Longrightarrow$ $\sum u_n$ converge ;
   - $\sum u_n$ diverge $\Longrightarrow$ $\sum v_n$ diverge ;
2. Si $0 < a \leqslant u_n/v_n \leqslant b < \infty$, $\sum u_n$ et $\sum v_n$ sont de même nature ;
3. Si $u_n/v_n \to c$ (c.-à-d. $u_n \simeq c\, v_n$) avec c fini et $c \neq 0$, alors $\sum u_n$ et $\sum v_n$ sont de même nature ;
4.
   - Si $u_{n+1}/u_n \leqslant k < 1$, $\sum u_n$ converge ;
   - Si $u_{n+1}/u_n \geqslant 1$, $\sum u_n$ diverge ;
5. Si $u_{n+1}/u_n \leqslant v_{n+1}/v_n$, $\sum v_n$ converge $\Longrightarrow$ $\sum u_n$ converge. ⌟

Preuve
Les preuves s'établissent par des majorations ou minorations. Intéressons-nous par exemple à la règle 4.

Dans le premier cas, à partir d'un certain rang p, $u_n < k^{n-p}\, u_p = v_n$; v_n est une suite géométrique de raison $k < 1$, donc $\sum v_n$ est convergente et, par application de la règle 1, $\sum u_n$ aussi.

Dans le deuxième cas, $\forall\, n \geqslant p$, $u_n \geqslant u_p$: u_n ne tend pas vers 0, donc la série $\sum u_n$ diverge. ⌟

Exemple 15
On prend comme série de référence $v_n = 1/n^\alpha$: si $0 < a \leqslant n^\alpha\, u_n \leqslant b < \infty$ (théor. 21, cas 2) ou si $\lim n^\alpha\, u_n = c \neq 0$ (même théor., cas 3), $\sum u_n$ est convergente si $\alpha > 1$ et divergente si $\alpha \leqslant 1$.

Si $\lim(n^\alpha\, u_n) = 0$, $\sum u_n$ converge si $\alpha > 1$, mais on ne peut pas conclure si $\alpha \leqslant 1$. De même, si $\lim(n^\alpha\, u_n) = \infty$, $\sum u_n$ diverge si $\alpha \leqslant 1$, mais on ne peut pas conclure si $\alpha > 1$.

Exemple 16 (critère de Cauchy)
On prend comme série de référence $v_n = a^n$ qui converge si $a < 1$, diverge si $a \geqslant 1$. Donc

- si $u_n^{1/n} \geqslant 1$, la série $\sum u_n$ est divergente.
  En effet, $u_n \geqslant v_n = 1^n$; $\sum v_n$ diverge, donc $\sum u_n$ aussi, d'après le théor. 21, cas 1, 1er point ;
- si $u_n^{1/n} \leqslant k < 1$, la série $\sum u_n$ est convergente.
  En effet, $u_n \leqslant v_n = k^n$; $k < 1$, donc $\sum v_n$ converge, et $\sum u_n$ aussi d'après le théor. 21, cas 1, 2e point ;
- si $\lim u_n^{1/n} = \ell$, $\sum u_n$ est divergente si $\ell > 1$ et convergente si $\ell < 1$.
  Si $\ell = 1$, on ne peut pas conclure si $u_n^{1/n} \to 1^-$; dans les autres cas ($u_n^{1/n} \to 1^+$ ou $u_n^{1/n} \to 1$ en oscillant), $\sum u_n$ diverge.

Voici quelques exemples d'application de la règle de Cauchy :

1. $u_n = \left(\dfrac{n}{n+a}\right)^{n^2}$.
   $u_n^{1/n} = (1 + a/n)^{-n}$, suite qui qui tend vers e^{-a} (cf. l'exemple 8 du chapitre II). Si $a > 0$, $\sum u_n$ est convergente ; si $a \leqslant 0$, $\sum u_n$ est divergente (puisque u_n ne tend pas vers zéro pour $a = 0$).
2. $u_n = a^n/n!$, avec a un réel donné.
   Pour un entier m supérieur à $|a|$, on a
   $$|u_n| = \frac{1}{m!}\,\frac{|a|^n}{(m+1)\cdots n} < \frac{m^m}{m!}\,\frac{|a|^n}{m^n}$$
   dès que $n > m$, d'où
   $$u_n^{1/n} < \frac{|a|}{m}\cdot\left(\frac{m^m}{m!}\right)^{1/n}$$
   qui tend vers $|a|/m < 1$. Il y a donc convergence quel que soit a. On peut en outre montrer que $\ell := \lim u_n^{1/n} = 0$.
   Cette série a pour somme le nombre e^a (cf. le développement de la fonction e^x).

Exemple 17 (critère de d'Alembert)
Si $u_{n+1}/u_n \to \ell$, on obtient les mêmes résultats que pour la règle de Cauchy puisque $u_{n+1}/u_n \to \ell \Longrightarrow u_n^{1/n} \to \ell$ (cf. théor. 8). ⌟

Théorème 22 (règle de Riemann)
Soit $(u_n) \geqslant 0$, $\alpha \in \mathbb{R}$ et $\ell \in \mathbb{R}^+ \cup \{+\infty\}$ tels que $n^\alpha\, u_n \to \ell$:

- si $\alpha > 1$ et $\ell \neq +\infty$, $\sum u_n$ converge ;
- si $\alpha \leqslant 1$ et $\ell \neq 0$, $\sum u_n$ diverge. ⌟

Théorème 23 (étude asymptotique)
Soit $v_n > 0$ et $u_n \in \mathbb{R}$ (u_n n'est pas forcément une série à termes positifs).

- Si $\sum v_n$ converge,
$$u_n = O(v_n) \Longrightarrow \sum u_n \text{ converge et } \sum_{k=n}^{\infty} u_k = O\Big(\sum_{k=n}^{\infty} v_k\Big) \text{ quand } n \to \infty.$$
On obtient la même chose en remplaçant « $= O(\ldots)$ » par « $= o(\ldots)$ » ou « $\simeq$ » à gauche et à droite de « $\Longrightarrow$ » [(*1)].
- Si $\sum v_n$ diverge,
$$u_n = O(v_n) \Longrightarrow \sum_{k=0}^{n} u_k = O\Big(\sum_{k=0}^{n} v_k\Big) \text{ quand } n \to \infty.$$
On obtient la même chose en remplaçant « $= O(\ldots)$ » par « $= o(\ldots)$ » ou « $\simeq$ » à gauche et à droite de « $\Longrightarrow$ » [(*2)].
En outre, si $u_n \simeq v_n$, la série $\sum u_n$ diverge. ⌟

Récapitulatif. — Pour décider de la convergence ou de la divergence d'une série **à termes positifs** :

- vérifier que $u_n \to 0$; sinon, $\sum u_n$ est grossièrement divergente !
- chercher un équivalent plus simple de u_n ;
- majorer par une série convergente ou minorer par une série divergente ;
- comparer à une série connue (critères $n^{-\alpha}$, $a^n \ldots$).

III. . Convergence absolue, semi-convergence

Définition 11
Une série $\sum u_n$ est dite absolument convergente si $\sum \|u_n\|$ converge. ⌟

En utilisant le critère de Cauchy et l'inégalité triangulaire, on montre qu'une série absolument convergente est aussi convergente tout court :

Théorème 24
Soit $\sum u_n$ une série dans un espace vectoriel normé complet E.
$$\sum \|u_n\| \text{ convergente} \Longrightarrow \sum u_n \text{ convergente et } \Big\|\sum^{\infty} u_n\Big\| \leqslant \sum^{\infty} \|u_n\|. \tag{5}$$
⌟

La réciproque est fausse.

Définition 12
Une série convergente mais non absolument convergente est dite semi-convergente. ⌟

Exemple 18
$\sum_{n=1}^{\infty}\big((-1)^{n+1}/n\big)$ est finie (et vaut $\ln 2$; voir TD), tandis que la série harmonique $\sum 1/n$ est divergente. ⌟

Evidemment, les séries à termes positifs convergentes sont aussi absolument convergentes.

1. Noter que le résultat est sur le *reste* de la série, $\sum_{k=n}^{\infty}$, pas sur la somme des n premiers termes.
2. Noter que le résultat est cette fois-ci sur la somme des n premiers termes, pas sur le reste (qui n'est d'ailleurs pas défini pour v_n puisque que $\sum v_n$ diverge).

III. .1. Convergence commutative

Une propriété remarquable des séries absolument convergentes est qu'on a maintenant le droit de changer l'ordre des termes :

Théorème 25 (convergence commutative)
Soit $\sum u_n$ une série absolument convergente.

Quelle que soit la permutation de $\mathbb{N}$ (c.-à-d. la bijection de $\mathbb{N}$ dans $\mathbb{N}$), $\sum u_{\sigma(n)}$ est absolument convergente et $\sum_{n=0}^{\infty} u_{\sigma(n)} = \sum_{n=0}^{\infty} u_n$. ⌟

Preuve

Absolue convergence de $\sum u_{\sigma(n)}$. Soit $N = \max_{k\in[\![0,\, n]\!]} \sigma(k)$. Puisque σ est une bijection et que $\forall\, k \in [\![0, n]\!]$, $\sigma(k) \in [\![0, N]\!]$, on a

$$\sum_{k=0}^{n} \|u_{\sigma(k)}\| \leqslant \sum_{k=0}^{N} \|u_k\| \leqslant \sum_{k=0}^{\infty} \|u_k\|.$$

Cette dernière somme existe car $\sum u_n$ est absolument convergente. La série $\sum \|u_{\sigma(n)}\|$ est une série à termes positifs majorée, donc convergente. La série $\sum u_{\sigma(n)}$ est donc absolument convergente.

Égalité des limites. Posons $p = \max_{k\in[\![0,\, n]\!]} \sigma^{\langle -1\rangle}(k)$. Pour tout $k \in [\![0, n]\!]$, $\sigma^{\langle -1\rangle}(k) \in [\![0, p]\!]$, donc $[\![0, n]\!] \subset \sigma([\![0, p]\!])$. Comme $\operatorname{card}([\![0, n]\!]) = n + 1$ et $\operatorname{card}(\sigma([\![0, p]\!])) = p + 1$ (puisque σ est une bijection), on a $p \geqslant n$.

Soit $A = \sigma([\![0, p]\!]) \setminus [\![0, n]\!]$. On a

$$\sum_{k=0}^{n} u_k - \sum_{k=0}^{p} u_{\sigma(k)} = \sum_{k\in A} u_k$$

et

$$\left\|\sum_{k=0}^{n} u_k - \sum_{k=0}^{p} u_{\sigma(k)}\right\| \leqslant \sum_{k\in A} \|u_k\| \leqslant \sum_{k=n+1}^{\infty} \|u_k\|,$$

puisque $A \subset [\![n + 1, +\infty[\![$ et que $\sum_{k=n+1}^{\infty} \|u_k\|$ existe.

Quand $n \to \infty$, $\sum_{k=0}^{n} u_k \to S_\infty$, $\sum_{k=n+1}^{\infty} \|u_k\| \to 0$; p tend également vers ∞, car $p \geqslant n$, donc $\sum_{k=0}^{p} u_{\sigma(k)} \to S_\infty$. ⌟

Il est en général incorrect de modifier l'ordre des termes d'une série non absolument convergente. Cela risque de modifier non seulement la somme mais même la nature de la série.

Exemple 19
Soit $u_n = (-1)^{n+1}/\sqrt{n}$. La série $\sum u_n$ est convergente (cf. § III.D) mais pas absolument convergente (cf. théor. 14). Réarrangeons les termes selon l'ordre $\ldots, u_{4\,p-3}, u_{4\,p-1}, u_{2\,p}, \ldots$ et supposons la nouvelle série convergente. On a alors le droit de regrouper les termes consécutifs. La série de terme général $u'_p = u_{4\,p-3} + u_{4\,p-1} + u_{2\,p}$ est équivalente à la série à termes positifs $p \mapsto 2/\sqrt{4\,p} - 1/\sqrt{2\,p} = (1 - 1/\sqrt{2})/\sqrt{p}$. D'après le critère de Riemann avec $\alpha = 1/2$, cette dernière diverge, donc $\sum u'_p$ aussi, contrairement à l'hypothèse. ⌟

On peut même montrer que si une série de réels est semi-convergente, on peut toujours trouver une permutation $\sigma\colon \mathbb{N} \to \mathbb{N}$ de ses termes telle que $\sum u_{\sigma(n)}$ converge vers n'importe quel réel ou tende vers $\pm\infty$ (théorème de réarrangement de Riemann).

Il est également incorrect de remplacer le terme général de la série par un terme équivalent. Si $u_n \simeq v_n$ et que $\sum \|v_n\|$ ne converge pas, $\sum u_n$ n'est pas nécessairement de même nature que $\sum v_n$!

III. .2. Produit de Cauchy

Soient (u_n) et (v_n) des suites de réels ou de complexes. Pour tout $p \in \mathbb{N}$ (donc fini), le produit des sommes partielles jusqu'à l'ordre p vaut

$$\left(\sum_{n=0}^{p} u_n\right)\cdot\left(\sum_{n=0}^{p} v_n\right) = \sum_{i=0}^{p}\left(\sum_{j=0}^{p} u_i\, v_j\right).$$

Numérotons de $i = 0$ à p (resp. de $j = 0$ à p) les lignes (resp. les colonnes) d'un tableau carré de taille $p + 1$ et disposons chaque $u_i\, v_j$ sur la ligne n° i et dans la colonne n° j : $(\sum_{n=0}^{p} u_n)\cdot(\sum_{n=0}^{p} v_n)$ est la somme de toutes les cases du tableau.

On obtient le même total en sommant sur les diagonales ascendantes du tableau. Notons de $n = 0$ à $2\,p$ ces diagonales. Pour $n \leqslant p$, la somme des termes sur la diagonale n° n vaut $w_n = \sum_{k=0}^{n} u_k\, v_{n-k}$, donc la somme sur les $p + 1$ premières diagonales vaut $\sum_{n=0}^{p} w_n$.

Le théorème suivant étend ce résultat aux sommes infinies sous certaines conditions :

Théorème 26 (produit de Cauchy)
Soient (u_n) et (v_n) deux suites dans $E = \mathbb{K}$.

Si $\sum u_n$ et $\sum v_n$ convergent absolument, la série de terme général $w_n = \sum_{k=0}^{n} u_k\, v_{n-k} = u_0\, v_n + \cdots + u_n\, v_0$ converge absolument et

$$\sum_{n=0}^{\infty} w_n = \left(\sum_{n=0}^{\infty} u_n\right)\cdot\left(\sum_{n=0}^{\infty} v_n\right).$$

La série $\sum w_n$ porte le nom de produit de Cauchy des séries $\sum u_n$ et $\sum v_n$. ⌟

Remarque
On peut même montrer (théorème de Mertens) que si une seule des deux séries $\sum u_n$ et $\sum v_n$ est absolument convergente (l'autre étant simplement convergente), alors $\sum w_n$ est convergente (mais pas absolument a priori) et a pour limite $(\sum_{n=0}^{\infty} u_n)\cdot(\sum_{n=0}^{\infty} v_n)$. ⌟

III. . Séries alternées. Théorème d'Abel

Définition 13
Une série de réels est *alternée* si son terme général a un signe qui change selon que n est pair ou impair. ⌟

Exemple 20
Si les suites (u_n) et (v_n) sont telles que $u_n = (-1)^n\, |u_n|$ et $v_n = (-1)^{n+1}\, |v_n|$, les séries $\sum u_n$ et $\sum v_n$ sont alternées. ⌟

Théorème 27 (Critère de Leibniz)
Soit $\sum u_n$ une série alternée. Si la suite $(|u_n|)$ est décroissante et tend vers 0, alors $\sum u_n$ est convergente, $|S_\infty - S_n| \leqslant |u_{n+1}|$ et $S_\infty - S_n$ est du signe de u_{n+1}. ⌟

Preuve
Supposons que $\forall\, n \in \mathbb{N}$, $u_{2\,n} \geqslant 0$ et $u_{2\,n+1} \leqslant 0$ (sinon, raisonner sur $v_n = -u_n$).

On montre sans difficulté que

$$S_{2\,n-1} \leqslant S_{2\,n+1} \leqslant S_{2\,n} \leqslant S_{2\,n-2}.$$

(Par exemple,

$$S_{2\,n} = S_{2\,n-2} + \underbrace{u_{2\,n} + u_{2\,n-1}}_{\leqslant 0},$$

donc $S_{2\,n} \leqslant S_{2\,n-2}$.)

La suite de terme général $v_n = S_{2\,n+1}$ est croissante, celle de terme général $w_n = S_{2\,n}$ est décroissante. Comme $v_n - w_n = S_{2\,n+1} - S_{2\,n} = u_{2\,n+1} \to 0$, les suites (v_n) et (w_n) convergent et ont même limite, S_∞, (cf. théor. des suites adjascentes).

Pour les mêmes raisons, on a $\forall\, (p, n) \in \mathbb{N}^2$, $v_p \leqslant S_\infty \leqslant w_n$, soit en particulier $S_{2\,n-1} \leqslant S_{2\,n+1} \leqslant S_\infty \leqslant S_{2\,n}$. En retranchant soit $S_{2\,n-1}$, soit $S_{2\,n}$, on obtient $0 \leqslant S_\infty - S_{2\,n-1} \leqslant u_{2\,n}$ et $u_{2\,n+1} \leqslant S_\infty - S_{2\,n} \leqslant 0$: pour tout n, $|S_\infty - S_n| \leqslant |u_{n+1}|$ et $S_\infty - S_n$ est du même signe que u_{n+1}. ⌟

Exemple 21
Les séries de terme général $u_n = (-1)^n/\sqrt{n}$ et $v_n = (-1)^n/n$ sont convergentes (mais pas absolument convergentes). ⌟

Théorème 28 (Abel)
Soient $b_n \geqslant 0$ une suite décroissante et tendant vers 0 et a_n une suite de signe quelconque, telle que $|a_0 + \cdots + a_n| \leqslant A$ pour tout n. Alors, la série $\sum a_n\, b_n$ est convergente et $|\sum a_n\, b_n| \leqslant A\, b_0$. ⌟

Preuve
Soit $A_n = a_0 + \cdots + a_n$.

$$\sum_{k=0}^{n} a_k\, b_k = a_0\, b_0 + \sum_{k=1}^{n} b_k\cdot(A_k - A_{k-1}) = A_0\, b_0 + \sum_{k=1}^{n} b_k\, A_k - \sum_{k=0}^{n-1} b_{k+1}\, A_k = A_n\, b_n + \sum_{k=0}^{n-1} A_k\cdot(b_k - b_{k+1}).$$

D'après les hypothèses du théorème, $|A_k\cdot(b_k - b_{k+1})| \leqslant A\cdot(b_k - b_{k+1})$. Or $\sum(b_k - b_{k+1})$ converge puisque b_k converge (cf. théor. 7). La série $\sum A_k\cdot(b_k - b_{k+1})$ est donc (absolument) convergente. Comme $A_n\, b_n$ tend vers zéro, la série $\sum a_k\, b_k$ converge.

Par ailleurs,

$$\left|\sum_{k=0}^{n} a_k\, b_k\right| \leqslant |A_n\, b_n| + \sum_{k=0}^{n-1} \left|A_n\cdot(b_k - b_{k+1})\right| \leqslant A\, b_n + A \sum_{k=0}^{n-1}(b_k - b_{k+1}) = A\cdot(b_n + (b_0 - b_n)) = A\, b_0.$$

┘

On retrouve aisément à l'aide de ce théorème le premier résultat du critère de Leibniz. Il suffit de poser $b_n = |u_n|$, $a_n = (-1)^n$ (ou $(-1)^{n+1}$). (b_n) est bien une suite de réels positifs décroissante et tendant vers 0 et $|A_n| \leqslant 1$, donc $\sum(-1)^n\, |u_n|$ converge.

On verra une application du théorème d'Abel aux séries de Fourier (cf. chapitre VII).

Chapitre IV

F

Une *fonction* $f: A \to F$ est une relation qui associe à chaque élément x de l'ensemble de départ A au plus un élément, noté $f(x)$, de l'ensemble d'arrivée F ; l'ensemble D_f des $x \in A$ pour lesquels $f(x)$ existe est le *domaine de définition* de f.

Une *application* est une fonction qui associe à chaque élément de l'ensemble de départ exactement un élément dans l'ensemble d'arrivée.

Définition 14 (restriction)
Soient f une application de A dans F et B une partie de A. On appelle *restriction* de f à B et on note $f|_B$ l'application

$$\begin{aligned} f|_B : B &\to F \\ x &\mapsto f(x). \end{aligned}$$

⌟

La restriction d'une fonction à son domaine de définition en fait ainsi une application.

Définition 15 (prolongement)
Soient f une application de A dans F et g une application de B dans F. g est un prolongement de f à B si et seulement si $A \subset B$ et $\forall\, x \in A$, $g(x) = f(x)$. ⌟

g est évidemment un prolongement de $g|_A$.

Une application $f: A \to F$ est *bijective* si $\forall\, y \in F, \exists!\, x \in A, f(x) = y$. f admet alors une application *réciproque* [(*1)], $f^{\langle -1 \rangle}: F \to A$, définie par $f^{\langle -1 \rangle}(y) = x$ si $f(x) = y$.

La situation que nous allons rencontrer le plus souvent est quand $F = \mathbb{R}$ et A est un intervalle (fini ou infini) de $\mathbb{R}$ ou une union de tels intervalles.

Rappelons d'abord la définition et les propriétés de quelques fonctions élémentaires.

IV. . Fonctions élémentaires

IV. .1. Fonctions trigonométriques

L'interprétation géométrique de $\cos\theta$ et $\sin\theta$ est donnée sur la figure I.1. Le rapport $\tan\theta = \sin\theta/\cos\theta$, noté $\operatorname{tg}\theta$ dans les « anciennes » notations françaises, représente la *pente* de la droite OM. On utilise parfois $\cot\theta = 1/\tan\theta$.

Les *identités trigonométriques* découlent toutes de

$$\sin(-a) = -\sin a, \tag{1a}$$
$$\cos(-a) = \cos a, \tag{1b}$$
$$\sin^2 a + \cos^2 a = 1, \tag{1c}$$
$$\sin(\pi/2 - a) = \cos a, \tag{1d}$$
$$\cos(a - b) = \cos a \cos b + \sin a \sin b, \tag{1e}$$

1. L'appellation *fonction inverse* et la notation f^{-1}, toutes deux usuelles, seront évitées dans ce cours en raison du risque de confusion avec $1/f$. En particulier, par cohérence avec la notation $\sin^2 x = (\sin x)^2 \neq \sin(\sin x)$, on ne notera pas $\sin^{-1} x$ la fonction réciproque de $\sin x$, mais $\arcsin x$. De même pour les autres fonctions trigonométriques et hyperboliques (notation $\operatorname{argsh} x$ et non $\operatorname{sh}^{-1} x$ pour ces dernières).

dont l'interprétation géométrique doit être claire (la dernière, par exemple, est obtenue en calculant le produit scalaire de deux vecteurs unitaires faisant respectivement un angle a et un angle b avec l'axe des abscisses). À partir de ces identités, on peut retrouver les identités de l'appendice IV.1.

IV. .2. Fonctions trigonométriques réciproques

Dans un intervalle où elles sont monotones, on peut « inverser » les fonctions $\sin x$, $\cos x$ et $\tan x$.

On définit la fonction $\arcsin x$ comme la fonction réciproque de $\sin x$ dans l'intervalle $[-\pi/2, \pi/2]$: $\arcsin x$ est définie pour $x \in [-1, 1]$; elle y est croissante et prend ses valeurs dans $[-\pi/2, \pi/2]$:

x	-1		0		1
$\arcsin x$	$-\frac{\pi}{2}$	$\nearrow$	0	$\nearrow$	$\frac{\pi}{2}$

La solution générale de $x = \sin y$ est $y = \arcsin x + 2\,k\,\pi$ ou $y = \pi - \arcsin x + 2\,k\,\pi$, k un entier arbitraire.

De même, $\arccos x$ est définie pour $x \in [-1, 1]$; elle y est décroissante et prend ses valeurs dans $[0, \pi]$:

x	-1		0		1
$\arccos x$	π	$\searrow$	$\frac{\pi}{2}$	$\searrow$	0

La solution générale de $x = \cos y$ est $y = \pm \arccos x + 2\,k\,\pi$. On a $\arcsin x + \arccos x = \pi/2$, ce que l'on peut déduire soit de l'inversion de $x = \cos y = \sin(\pi/2 - y)$, soit du calcul de la dérivée de cette somme.

Enfin, la fonction réciproque $\arctan x$ est définie pour tout x réel ; elle est croissante et prend ses valeurs dans $[-\pi/2, \pi/2]$:

x	$-\infty$		0		$+\infty$
$\arctan x$	$-\frac{\pi}{2}$	$\nearrow$	0	$\nearrow$	$\frac{\pi}{2}$

La solution générale de $x = \tan y$ est $y = \arctan x + k\,\pi$.

Les dérivées des fonctions trigonométriques réciproques sont

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}, \quad (\arccos x)' = \frac{-1}{\sqrt{1-x^2}}, \quad (\arctan x)' = \frac{1}{1+x^2}.$$

IV. .3. Fonctions exponentielles et logarithmes

Il s'agit de e^x, ou plus généralement a^x pour $a > 0$. Leur propriété caractéristique est

$$a^x\,a^y = a^{x+y} \tag{2}$$

(et donc aussi $(a^x)^y = a^{x\,y}$).

La fonction réciproque de a^x est le logarithme de base a, $\log_a x$. La fonction $\log_{\mathrm{e}} x$ est notée $\ln x$.

$$y = \mathrm{e}^x \Leftrightarrow x = \ln y, \tag{3a}$$
$$y = a^x \Leftrightarrow x = \log_a y, \tag{3b}$$
$$a^x = \mathrm{e}^{x \ln a}, \tag{3c}$$
$$\log_a x = \frac{\ln x}{\ln a}. \tag{3d}$$

La relation avec les fonctions trigonométriques est

$$\mathrm{e}^{\mathrm{i}\,x} = \cos x + \mathrm{i}\,\sin x. \tag{4}$$

IV. .4. Fonctions hyperboliques

Les cosinus, sinus et tangente hyperboliques, notés $\operatorname{ch} x$ ou $\cosh x$, $\operatorname{sh} x$ ou $\sinh x$, et $\operatorname{th} x$ ou $\tanh x$, sont définis de la manière suivante :

$$\operatorname{ch} x = \frac{1}{2}\,(\mathrm{e}^x + \mathrm{e}^{-x}), \qquad \operatorname{sh} x = \frac{1}{2}\,(\mathrm{e}^x - \mathrm{e}^{-x}), \qquad \operatorname{th} x = \frac{\operatorname{sh} x}{\operatorname{ch} x}. \tag{5}$$

Ces fonctions obéissent aux relations

$$\operatorname{ch}(-x) = \operatorname{ch} x, \qquad \operatorname{sh}(-x) = -\operatorname{sh} x, \qquad \operatorname{ch}^2 x - \operatorname{sh}^2 x = 1, \qquad (\operatorname{ch} x)' = \operatorname{sh} x \quad \text{et} \quad (\operatorname{sh} x)' = \operatorname{ch} x. \tag{6}$$

Les fonctions hyperboliques sont reliées aux foncions trigonométriques par

$$\operatorname{ch}(\mathrm{i}\,x) = \cos x \quad \text{et} \quad \operatorname{sh}(\mathrm{i}\,x) = \mathrm{i}\,\sin x. \tag{7}$$

Les propriétés de symétrie et le comportement à l'infini des fonctions hyperboliques sont illustrées sur la figure IV.1.

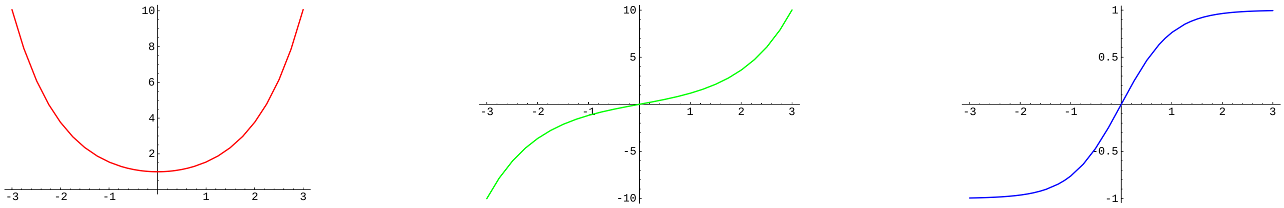


Figure IV.1 – Graphes des fonctions hyperboliques $\operatorname{ch} x$, $\operatorname{sh} x$ et $\operatorname{th} x$.

Les fonctions réciproques de $\operatorname{sh} x$, $\operatorname{ch} x$ et $\operatorname{th} x$ sont respectivement $\operatorname{argsh} x$, $\operatorname{argch} x$ et $\operatorname{argth} x$. Comme

$$x = \operatorname{sh}(\operatorname{argsh} x) = \frac{\mathrm{e}^{\operatorname{argsh} x} - \mathrm{e}^{-\operatorname{argsh} x}}{2} \Longrightarrow y^2 - 2\,x\,y - 1 = 0, \text{ avec } y = \mathrm{e}^{\operatorname{argsh} x},$$

on a

$$\operatorname{argsh} x = \ln(x + \sqrt{x^2+1}).$$

On obtient de même que

$$\operatorname{argch} x = \ln(x + \sqrt{x^2-1}) \quad \text{et} \quad \operatorname{argth} x = \frac{1}{2}\ln\Big(\frac{1+x}{1-x}\Big).$$

On a

$$(\operatorname{argsh} x)' = \frac{1}{\sqrt{x^2+1}}, \qquad (\operatorname{argch} x)' = \frac{1}{\sqrt{x^2-1}}, \qquad (\operatorname{argth} x)' = \frac{1}{1-x^2}.$$

IV. . Limite d'une fonction

Définition 16 (adhérence)
Soit E un espace vectoriel normé et A une partie de E.

On appelle adhérence de A (dans E) l'ensemble noté $\overline{A}$ des points x de E tels que

$$\forall\, \epsilon > 0, \exists\, y \in A, \|x - y\| < \epsilon.$$

$\lrcorner$

Remarque
On a évidemment $A \subset \overline{A}$. $\lrcorner$

Exemple 22
Si $(a, b) \in \mathbb{R}^2$, $\overline{]a, b[} = \overline{]a, b]} = \overline{[a, b[} = \overline{[a, b]} = [a, b]$. $\lrcorner$

Exemple 23
Soit $E = \mathbb{R}$. On a $\overline{\mathbb{Q}} = \overline{\mathbb{R} \setminus \mathbb{Q}} = \overline{\mathbb{R}} = \mathbb{R}$ [*2]. $\lrcorner$

2. On note parfois aussi $\overline{\mathbb{R}}$ l'ensemble $\mathbb{R} \cup \{-\infty, +\infty\} \supsetneq \mathbb{R}$. Il ne s'agit donc pas de l'adhérence de $\mathbb{R}$ dans l'espace vectoriel normé $\mathbb{R}$, puisque, par définition, l'adhérence *dans* E d'une partie A de E est incluse dans E.
   L'ensemble $\mathbb{R} \cup \{-\infty, +\infty\}$ n'est pas un espace vectoriel normé ($+\infty + (-\infty)$ n'est par exemple pas défini). En revanche, c'est un espace métrique, c.-à-d. qu'on peut le munir d'une distance. En généralisant la définition de l'adhérence au cas d'un espace métrique, on a bien $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$, où $\overline{\mathbb{R}}$ est l'adhérence dans $\mathbb{R} \cup \{-\infty, +\infty\}$.

Limite d'une fonction en un point à distance finie

Définition 17 (limite finie en un point à distance finie)
Soient E et F des espaces vectoriels normés, A une partie de E et f une application de A dans F.

f admet une limite (finie) $\ell \in F$ en un point a de $\overline{A}$ si

$$\forall\, \epsilon > 0, \exists\, \eta > 0, \forall\, x \in A, \|a - x\| < \eta \Rightarrow \|f(x) - \ell\| < \epsilon. \tag{8}$$

On note ceci $\lim_{x\to a} f(x) = \ell$ ou $f(x) \xrightarrow[x\to a]{} \ell$. ⌟

Exemple 24
La fonction x^2 a pour limite 4 quand $x \to 2$. En effet, $|x^2 - 4| = \left|(x-2)\,(x+2)\right|$, qui est inférieur à $5\,|x-2|$ dès que $|x+2| \leqslant 5$ (comme $x \to 2$, il suffit de prendre x tel que $|x-2| \leqslant 1$) : $5\,|x-2|$ peut donc être rendu aussi petit que l'on veut quand $x \to 2$. ⌟

Exemple 25
$f(x) = \sin(1/x)$ *n'a pas* de limite quand $x \to 0$: $f(x)$ oscille de plus en plus vite entre -1 et 1. ⌟

Remarques

- Il n'est pas nécessaire que f soit définie en a : il faut seulement que a appartienne à l'adhérence $\overline{A}$ d'une partie A sur laquelle f est définie. Par exemple, bien que la fonction $x \mapsto \sin x/x$ ne soit pas définie en 0, sa limite en 0 existe ($\lim_{x\to 0}(\sin x/x) = 1$).
  En revanche, si $a \in A$ et que f admet une limite finie en a, on a nécessairement $\lim_{x\to a} f(x) = f(a)$.
- Des notations particulières sont utilisées dans les cas suivants :
  - Si $B = A \setminus \{a\}$, $\lim_{x\to a,\, x\neq a} f(x) := \lim_{x\to a} f|_B(x)$;
  - Si $A \subset \mathbb{R}$ et $B = A \cap\,]-\infty, a[$, $\lim_{x\to a^-} f(x) := \lim_{x\to a} f|_B(x)$ (limite à gauche) ;
  - Si $A \subset \mathbb{R}$ et $B = A \cap\,]a, +\infty[$, $\lim_{x\to a^+} f(x) := \lim_{x\to a} f|_B(x)$ (limite à droite).

  Il faut bien sûr que $a \in \overline{B}$. Ce n'est par exemple pas le cas si $A = \{0\} \cup\,]1, +\infty[$, $a = 0$ et $B = A \setminus \{a\} = \,]1, +\infty[$: $a \notin \overline{B} = [1, +\infty[$.
- La limite est unique, mais elle dépend de l'ensemble de départ de l'application. ⌟

Exemple 26
Soit l'application

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} 0 & \text{si } x \in \mathbb{Q}, \\ 1 & \text{si } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases} \end{aligned}$$

On a $0 \in \overline{\mathbb{Q}} = \overline{\mathbb{R} \setminus \mathbb{Q}} = \mathbb{R}$. La limite de f en 0 n'est pas définie et

$$\lim_0 f|_{\mathbb{Q}} = 0 \neq \lim_0 f|_{\mathbb{R}\setminus\mathbb{Q}} = 1.$$

f, $f|_{\mathbb{Q}}$ et $f|_{\mathbb{R}\setminus\mathbb{Q}}$ sont des applications différentes. ⌟

Définition 18 (limite infinie en un point à distance finie)
Soient E un espace vectoriel normé, A une partie de E et f une application de A dans $\mathbb{R}$.

- f tend vers $+\infty$ en un point a de $\overline{A}$ si

$$\forall\, L \in \mathbb{R}, \exists\, \eta > 0, \forall\, x \in A, \|a - x\| < \eta \Rightarrow f(x) > L. \tag{9}$$

  On note ceci $\lim_{x\to a} f(x) = +\infty$ ou $f(x) \xrightarrow[x\to a]{} +\infty$.
- f tend vers $-\infty$ en un point a de $\overline{A}$ si

$$\forall\, M \in \mathbb{R}, \exists\, \eta > 0, \forall\, x \in A, \|a - x\| < \eta \Rightarrow f(x) < M. \tag{10}$$

⌟

Remarque
Bien que la notation soit la même que pour une limite finie, on ne dira pas que « f converge vers $+\infty$ ». ⌟

Exemple 27
La fonction tangente est définie dans les intervalles ouverts $](2\,n-1)\,\pi/2, (2\,n+1)\,\pi/2[$; $\tan x \to +\infty$ quand $x \to (2\,n+1)\,\pi/2$ par valeurs inférieures ; $\tan x \to -\infty$ quand $x \to (2\,n+1)\,\pi/2$ par valeurs supérieures. ⌟

Limite d'une fonction à l'infini

Définition 19 (limite finie en un point à l'infini)
Soient F un espace vectoriel normé, A une partie non majorée de $\mathbb{R}$ et f une application de A dans F.

f admet une limite (finie) $\ell \in F$ quand x tend vers $+\infty$ si

$$\forall\, \epsilon > 0, \exists\, X \in \mathbb{R}, \forall\, x \in A, x > X \Rightarrow \|f(x) - \ell\| < \epsilon. \tag{11}$$

On note ceci $\lim_{x\to+\infty} f(x) = \ell$ ou $f(x) \xrightarrow[x\to+\infty]{} \ell$. ⌟

On définit de manière analogue les limites infinies quand x tend vers $\pm\infty$.

Exemple 28
f tend vers $+\infty$ quand x tend vers $-\infty$ si et seulement si

$$\forall\, L \in \mathbb{R}, \exists\, X \in \mathbb{R}, \forall\, x \in A, x < X \Rightarrow f(x) > +\infty.$$ ⌟

Exemple 29
Nous avons $\lim_{x\to\infty}(2 + \mathrm{e}^{-x}) = 2$, tandis que $\lim_{x\to-\infty}(2 + \mathrm{e}^{-x}) = +\infty$. ⌟

Exemple 30
Le comportement quand $x \to +\infty$ d'une fraction rationnelle $P(x)/Q(x)$, rapport de deux polynômes P et Q de degrés respectifs p et q, est donné par

$$\frac{P(x)}{Q(x)} \to \begin{cases} 0 & \text{si } p < q, \\ a_p/b_p & \text{si } p = q, \\ \operatorname{sgn}(a_p/b_q) \times \infty & \text{si } p > q, \end{cases} \tag{12}$$

où a_p et b_q sont les coefficients de degré maximal de P et Q. ⌟

Si f et g sont deux applications de E dans F et que $f(x) \to \ell_f$ et $g(x) \to \ell_g$ quand $x \to a$, et $\lambda \in K$ (où K est le corps sur lequel l'espace vectoriel F est défini), nous avons les règles suivantes :

Théorème 29
Soient f et g deux applications de A dans F, de limites *finies* ℓ_f et ℓ_g quand $x \to a$, et $\lambda \in K$ (où K est le corps sur lequel l'espace vectoriel F est défini).

Quand $x \to a$,

$$\lambda\, f(x) \to \lambda\, \ell_f\ ; \tag{13a}$$
$$f(x) + g(x) \to \ell_f + \ell_g\ ; \tag{13b}$$
$$\text{si } F = \mathbb{K},\ f(x)\, g(x) \to \ell_f\, \ell_g\ ; \tag{13c}$$
$$\text{si } F = \mathbb{K} \text{ et si } \ell_g \neq 0,\ \frac{f(x)}{g(x)} \to \frac{\ell_f}{\ell_g}. \tag{13d}$$

Ces propriétés sont également valables quand $x \to \pm\infty$. ⌟

Par contre, si l'une des limites de f ou g est *infinie*, on rencontre parfois une « forme indéterminée » (voir section IV.F). Par exemple,

$$\lim_{x\to0}\left(\sqrt{x+1} + \sin\left(\frac{\pi\, x}{2}\right)\right) = 1, \tag{14a}$$
$$\lim_{x\to1}\left(\sqrt{x+1}\ \tan\left(\frac{\pi\, x}{2}\right)\right) = \infty, \tag{14b}$$
$$\lim_{x\to0}\left(\frac{\sqrt{x}}{\tan(\pi\, x/2)}\right) = \frac{0}{0}\ ??? \tag{14c}$$

Encadrement

Théorème 30 (« théorème des gendarmes »)
Si la fonction $f(x)$ est encadrée par les fonctions g et h en ce sens que $g(x) \leqslant f(x) \leqslant h(x)$ pour tout $x \in [a, b[$ (b pouvant être ∞), et si $g(x)$ et $h(x)$ ont la *même* limite (finie ou infinie) quand $x \to b$, alors $f(x)$ a elle aussi cette même limite. ⌟

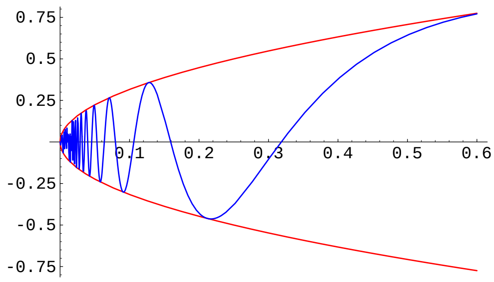


Figure IV.2 – La fonction $\sqrt{x}\,\sin(1/x)$, encadrée par les fonctions $\pm\sqrt{x}$, tend vers 0 quand $x \to 0$.

Exemple 31
Soit $f(x) = \sqrt{x}\,\sin(1/x)$, $g(x) = -\sqrt{x}$, $h(x) = \sqrt{x}$ pour $x \in\,]0, 1]$. Le théorème nous dit que $\lim_{x\to 0} f(x) = 0$. Voir figure IV.2. ⌟

Exemple 32
De même pour $x > 0$, $f(x) = \sin x/x$, $g(x) = -1/x$, $h(x) = 1/x$. Le théorème nous dit que $\lim_{x\to\infty} f(x) = 0$. ⌟

IV. . Notations de Landau

Une variable x qui tend vers 0 est appelée un *infiniment petit*. Une fonction $f(x)$ qui tend vers zéro avec x est un *infiniment petit équivalent à* x si $f(x)/x$ a une limite finie non nulle quand $x \to 0$; $f(x)$ est dite un *infiniment petit d'ordre* p (par rapport à x) si $f(x)/x^p \to \lambda$ fini quand $x \to 0$, où $p > 0$ est un rationnel. On écrit alors

$$f(x) \simeq \lambda\, x^p \tag{15}$$

et l'on dit que $\lambda\, x^p$ est la *partie principale* de $f(x)$ quand $x \to 0$.

IV. . Continuité d'une fonction

Comparons la limite et la valeur d'une fonction en un point.

Définition 20 (continuité)
Soit $f\colon A \subset E \to F$ une application.
f est *continue* en $a \in A$ si $\lim_{x\to A} f(x) = f(a)$. Elle est continue sur A si elle l'est en chaque point a de A. ⌟

Exemple 33
Les fonctions usuelles, puissances, polynômes, trigonométriques, exponentielle, logarithme, etc., sont continues en tout point de leur ensemble de définition. ⌟

Exemple 34
La fonction *partie entière* $\mathrm{E}(x)$,

$$\mathrm{E}(x) := n \text{ si } n \leqslant x < n+1, \text{ avec } n \in \mathbb{Z}, \tag{16}$$

a une limite en tout point non entier et y est continue, mais est *discontinue* en tout point entier. En effet, la limite de $\mathrm{E}(x)$ quand $x \to n^-$ (c.-à-d. tend vers l'entier n par valeurs inférieures) est $n-1$, mais elle est n si $x \to n^+$. ⌟

Définition 21 (prolongement par continuité)
Soit $f\colon A \subset E \to F$ une application continue sur A et B une partie de E telle que $A \subset B \subset \overline{A}$.
Si, $\forall\, b \in B \setminus A$, f admet une limite *finie* en b, alors la fonction

$$\begin{aligned} g\colon B &\longrightarrow F \\ x &\longmapsto \begin{cases} f(x) & \text{si } x \in A, \\ \lim_{x\to b} f(x) & \text{si } x \in B \setminus A \end{cases} \end{aligned}$$

est continue sur B. On l'appelle le prolongement par continuité à B de f. ⌟

Exemple 35
La fonction $f\colon x \mapsto \sin x/x$ est définie et continue sur $A = \mathbb{R}^*$. Sa limite en 0 est finie et vaut 1. La fonction

$$\begin{aligned} g\colon \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} f(x) & \text{si } x \neq 0, \\ 1 & \text{si } x = 0 \end{cases} \end{aligned}$$

est donc le prolongement par continuité de f à $\mathbb{R}$. ⌟

IV. . Dérivation d'une fonction

Définition 22
Une fonction $f(x)$ est dite dérivable en un point x_0 si la limite

$$f'(x_0) = \lim_{x \to x_0,\, x \neq x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{17}$$

existe. On appelle alors $f'(x_0)$ la dérivée de $f(x)$ en x_0, également notée

$$f'(x_0) = \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x=x_0}. \tag{18}$$

⌟

Les dérivées successives (si elles existent) se notent

$$f''(x_0) = \left.\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right|_{x=x_0}, \quad \ldots \quad, \quad f^{(k)}(x_0) = \left.\frac{\mathrm{d}^k f}{\mathrm{d}x^k}\right|_{x=x_0}, \quad \ldots. \tag{19}$$

En physique, la notation f' est souvent réservée à une dérivation par rapport à une coordonnée d'espace, tandis qu'une dérivée par rapport au temps est notée par un point suscrit :

$$\dot{f}(t) = \frac{\mathrm{d}f}{\mathrm{d}t}, \quad \ddot{f}(t) = \frac{\mathrm{d}^2 f}{\mathrm{d}t^2}, \quad \text{etc.} \tag{20}$$

S'il y a plusieurs variables (par exemple des coordonnées d'espace et de temps, ou des variables de pression et de température), on utilise les notations de dérivée partielle. Par exemple, pour une fonction f dépendant des variables u et v,

$$f'_u(u,v) \equiv \frac{\partial f}{\partial u} := \lim_{h \to 0} \frac{f(u+h,v) - f(u,v)}{h}, \quad f'_v(u,v) \equiv \frac{\partial f}{\partial v}, \quad f''_{u,v}(u,v) \equiv \frac{\partial^2 f}{\partial u\, \partial v} \equiv \frac{\partial}{\partial u}\left(\frac{\partial f}{\partial v}\right), \quad \text{etc.} \tag{21}$$

Si $f(x)$ et $g(x)$ sont toutes deux dérivables en x, et si $\lambda \in \mathbb{R}$, on a les règles suivantes :

$$(\lambda\, f(x))' = \lambda\, f'(x)\,; \tag{22a}$$

$$(f(x) + g(x))' = f'(x) + g'(x)\,; \tag{22b}$$

$$(f(x)\, g(x))' = f'(x)\, g(x) + f(x)\, g'(x)\,; \tag{22c}$$

$$\left(\frac{f}{g}\right)' = \frac{f'\, g - f\, g'}{g^2}. \tag{22d}$$

Si $g(x)$ est dérivable en x et $f(x)$ est dérivable en $x_0 = g(x)$, on peut dériver la fonction composée

$$(f \circ g)' = g' \times f' \circ g, \quad \text{c.-à-d.} \quad (f[g(x)])' = g'(x)\, f'[g(x)]. \tag{23}$$

On a enfin

$$\left(f^{\langle -1 \rangle}\right)' = \frac{1}{f' \circ f^{\langle -1 \rangle}}, \quad \text{c.-à-d.} \quad \left(f^{\langle -1 \rangle}(x)\right)' = \frac{1}{f'(f^{\langle -1 \rangle}(x))}. \tag{24}$$

Ces règles découlent de la définition (17).
Les dérivées de quelques fonctions élémentaires sont données à l'appendice IV.2.

Définition 23 (fonction C^k)
Une fonction de $\mathbb{R}$ dans E est dite (de classe) C^k sur un intervalle I (on note $f \in C^k(I \to E)$) si elle est continue sur I et si toutes ses dérivées jusqu'à l'ordre k inclus existent et sont continues sur I.
Si elle est indéfiniment dérivable, on dit qu'elle est C^∞. ⌟

Une fonction C^0 est donc simplement une fonction continue.

Définition 24 (fonction C^k par morceaux)
Une fonction de $\mathbb{R}$ dans E est dite C^k par morceaux sur un intervalle I (on note $f \in^k (I \to E)$) si, pour tout intervalle $[a, b]$ inclus dans I, il existe une subdivision *finie* $x_0 = a < x_1 < \ldots < x_n = b$ de $[a, b]$ telle que, $\forall\, i \in [\![0, n-1]\!]$, $f\big|_{]x_i,\, x_{i+1}[}$ (c.-à-d. la restriction de f à l'intervalle $]x_i, x_{i+1}[$) soit prolongeable en une fonction C^k sur $[x_i, x_{i+1}]$ [*3]. ⌟

Une fonction ^{0} (continue par morceaux) sur un intervalle $[a, b]$ est donc une fonction continue sur $[a, b]$, sauf en un nombre fini de points, et dont les discontinuités en ces points sont finies.

IV. . Formes indéterminées

On rencontre souvent des situations où la limite se présente sous une *forme indéterminée*. On appelle ainsi des expressions de la forme « $0/0$ », « ∞/∞ », « $0 \times \infty$ », « $\infty - \infty$ », « 1^∞ », « 0^0 », « ∞^0 », etc.

Exemple 36
Quelles sont les limites pour $x \to 0$ de $\sin x/x$, $x \ln x$, $\ln x + 1/x$, $(1+x)^{1/x}$, x^x ? ⌟

On va montrer plus bas qu'on peut se ramener au cas « $0/0$ ». Considérons donc un rapport $f(x)/g(x)$ où $f(x)$ et $g(x)$ tendent vers 0 quand $x \to x_0$. Il existe plusieurs procédés d'analyse d'une telle forme indéterminée.

Tout d'abord, d'après la définition de la dérivée d'une fonction $f(x)$ en $x = x_0$ comme $\lim_{x\to x_0}(f(x) - f(x_0))/(x - x_0)$, on a le théorème suivant :

Théorème 31 (règle de l'Hospital)
Si $f(x_0) = g(x_0) = 0$, si les dérivées de f et g existent en x_0 et si leur rapport est fini, on a

$$\lim_{x_0} \frac{f(x)}{g(x)} = \frac{f'(x_0)}{g'(x_0)}. \tag{25}$$

⌟

Preuve
En effet,

$$\lim_{x_0} \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \lim_{x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} \, \frac{x - x_0}{g(x) - g(x_0)} \right). \tag{26}$$

⌟

Si le rapport $f'(x_0)/g'(x_0)$ est toujours indéterminé et si les dérivées d'ordre supérieur existent, on a le droit d'itérer la règle de l'Hospital, c.-à-d. examiner $f^{(n)}(x_0)/g^{(n)}(x_0)$ pour $n = 2, 3, \ldots$, jusqu'à l'obtention d'une forme déterminée.

Exemple 37
Que vaut $\lim_{x\to 0}(\sqrt{2+x} - \sqrt{2})/x$?

Première méthode : on multiplie et divise par la « quantité conjuguée », c.-à-d. on écrit

$$\frac{\sqrt{2+x} - \sqrt{2}}{x} = \frac{2 + x - 2}{(\sqrt{2+x} + \sqrt{2})\, x} = \frac{1}{\sqrt{2+x} + \sqrt{2}} \to \frac{1}{2\sqrt{2}}. \tag{27}$$

Seconde méthode : on se rappelle la définition de la dérivée :

$$\lim_{x\to 0} \frac{\sqrt{2+x} - \sqrt{2}}{x} = \left(\frac{\mathrm{d}}{\mathrm{d}x} \sqrt{x} \right)\bigg|_{x=2} = \frac{1}{2\sqrt{x}}\bigg|_{x=2} = \frac{1}{2\sqrt{2}}. \tag{28}$$

⌟

3. $f\big|_{]x_i,\, x_{i+1}[}$ est prolongeable sur $[x_i, x_{i+1}]$ si les limites de f en x_i^+ et x_{i+1}^- sont finies. Le prolongement de $f\big|_{]x_i,\, x_{i+1}[}$ est alors la fonction f_i définie par
   - $\forall\, x \in\,]x_i, x_{i+1}[$, $f_i(x) = f(x)$;
   - $f_i(x_i) = \lim_{x_i^+} f$;
   - $f_i(x_{i+1}) = \lim_{x_{i+1}^-} f$.

   f est donc ^{k} si $\forall\, i \in [\![0, n-1]\!]$, f_i est C^k sur $[x_i, x_{i+1}]$.

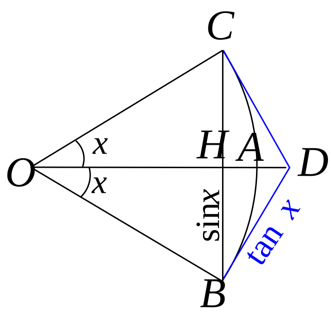


Figure IV.3 – Interprétation géométrique de (29) : $OB = OC = 1$; $BH = HC = \sin x$; $BD = DC = \tan x$. On a $BC \leqslant \text{arc}(BC) \leqslant BD + DC$, donc $2 \sin x \leqslant 2\,x \leqslant 2 \tan x$.

Exemple 38
Un résultat important est que

$$\lim_{x\to 0} \frac{\sin x}{x} = 1. \tag{29}$$

En effet, la limite est par définition la dérivée de $\sin x$ en 0, soit $\cos x$ évalué en 0, c.-à-d. 1. L'interprétation géométrique de ce résultat est aussi importante à garder en tête : $2 \sin x$ mesure la longueur de la corde dessinée sur un cercle de rayon unité par l'angle mesuré (en radians) par $2\,x$. Pour de petits angles, la longueur de la corde vaut approximativement celle de l'arc : $2 \sin x \approx 2\,x$ (voir figure IV.3). ⌟

Une autre idée est d'utiliser des majorations ou minorations (comme dans le théorème des gendarmes) pour se ramener à d'autres formes indéterminées dont la limite est plus aisément calculable. Ainsi, dans l'exemple précédent de $\sin x/x$, on a $2 \sin x \leqslant 2\,x \leqslant 2 \tan x$ (la corde BC est plus courte que celle de l'arc BAC, qui est plus court que la ligne brisée BDC), d'où l'encadrement $\cos x \leqslant \sin x/x \leqslant 1$.

Exemple 39
Quand $x \to \infty$, $\ln x$ tend vers ∞ *plus lentement* que x^α, pour tout $\alpha > 0$. Quand $x \to 0$, $|\ln x|$ tend vers ∞ *plus lentement* que $x^{-\alpha}$, pour tout $\alpha > 0$.

Pour voir cela, partons de l'inégalité $\ln x < \sqrt{x}$.

Pour démontrer cette inégalité, étudions la fonction $f(x) = \ln x - \sqrt{x}$. Sa dérivée, $f'(x) = (2 - \sqrt{x})/(2\,x)$, est négative pour $x > 4$, positive pour $0 < x < 4$. La fonction f est donc maximale en $x = 4$ et son maximum vaut $f(4) = 2\cdot(\ln 2 - 1) = 2 \ln(2/\mathrm{e}) < 0$. La fonction f est donc toujours négative.

On en déduit que $0 < \ln x/x < \sqrt{x}/x$, donc que $\lim_{+\infty} \ln x/x = 0$ et, plus généralement, par changement de variable, que

$$\forall\, \alpha > 0, \quad \lim_{x\to\infty} \frac{\ln x}{x^\alpha} = 0 \quad \text{et} \quad \lim_{x\to 0} x^\alpha \ln x = 0, \tag{30}$$

où l'on passe de la première assertion à la seconde par $x \mapsto 1/x$. ⌟

Quand toutes les méthodes précédentes ont échoué, parce que la dérivée de f ou de g n'existe pas, ou que f'/g' est encore du type 0/0, et qu'aucun encadrement n'a pu être trouvé, une discussion plus serrée s'impose. Si on a deux fonctions f et g dont le rapport $f(x)/g(x) \to 0/0$ quand $x \to x_0$, on peut remplacer chacune d'elles par sa partie principale au sens des infiniment petits : si $f(x) \simeq \lambda\cdot(x - x_0)^p$ et $g(x) \simeq \mu\cdot(x - x_0)^q$, le rapport f/g a une limite nulle, finie ou infinie selon que $p > q$, $p = q$ ou $p < q$. Obtenir ces parties principales de f et g peut lui-même nécessiter l'emploi de développements limités. . .

Exemple 40
Trouver la limite de $(x - \sin x)/(x - \tan x)$ quand $x \to 0$. En utilisant les développements limités (cf. § IV.G), on calcule $x - \sin x \simeq x^3/3!$, $x - \tan x \simeq -x^3/3$, donc le rapport tend vers $-1/2$. ⌟

Exercice
Calculer de même la limite de $(1 - x + \ln x)/(1 - \cos(1 - x))$ quand $x \to 1$. ⌟

Finalement, des formes indéterminées de la forme ∞/∞, 1^∞, etc., peuvent se ramener au cas 0/0 précédent par passage à l'inverse, exponentiation, etc. Ainsi, si $f \to 0$ et $g \to \infty$, le produit $f\,g$ a une forme indéterminée $0 \times \infty$ mais $f/(1/g)$ est de la forme 0/0. De même, si $f \to 1$, $g \to \infty$, $f^g = \exp(\ln f/(1/g))$ et l'argument de l'exponentielle est à nouveau de la forme 0/0. Exemples : quand $x \to 0$, $(1 + x)^{1/x} = \exp(\ln(1 + x)/x)$ a donc pour limite e ; $\lim_{|x|\to 0} x^{\ln(1+x)}$, de la forme 0^0, vaut 1, etc.

IV. . Développements limités

Un développement limité de $f(x)$ au voisinage de $x = 0$ est un développement dans lequel les termes successifs sont des infiniment petits d'ordre entier croissant :

$$f(x) = a_0 + a_1\,x + a_2\,x^2 + \cdots + a_p\,x^p + o(x^p), \tag{31}$$

où le dernier terme est une notation signifiant que $\lim_{x\to 0} o(x^p)/x^p = 0$ (on trouve aussi dans les livres la notation $\varepsilon(x^p)$). Plus généralement, un développement limité de $f(x)$ au voisinage d'un point $x = a$ est un développement de la forme

$$f(x) = a_0 + a_1{\cdot}(x-a) + a_2{\cdot}(x-a)^2 + \cdots + a_p{\cdot}(x-a)^p + o((x-a)^p). \tag{32}$$

Développements de Taylor

Une façon pratique de construire de tels développements limités est d'utiliser la formule de Taylor. Celle-ci est une généralisation d'une propriété connue des fonctions continues et dérivables sur un intervalle $[a, b]$: on peut écrire, pour $x \in [a, b]$, $f(x) - f(a) = (x-a)\,f'(a) + o(x-a)$ (et on dit que « $f(a) + (x-a)\,f'(a)$ est la meilleure approximation affine à $f(x)$ »). Le théorème suivant généralise ceci :

Théorème 32 (Taylor-Young)
Si la fonction $f(x)$ est continue ainsi que ses $n-1$ premières dérivées sur l'intervalle $[a, b]$, et si $f^{(n)}(a)$ existe en a, on peut, pour tout $x \in [a, b]$, écrire le développement de Taylor-Young

$$f(x) = f(a) + (x-a)\,f'(a) + \cdots + \frac{(x-a)^n}{n!}\,f^{(n)}(a) + o((x-a)^n). \tag{33}$$

⌟

Preuve (ébauche)
La formule de Taylor-Young se démontre par récurrence, en l'appliquant à l'ordre $n-1$ à la fonction f' et en intégrant le développement limité obtenu, ce qu'on montre être légitime. ⌟

Il existe une autre formule de Taylor (Taylor-Lagrange) qui donne une autre forme du reste :

Théorème 33 (Taylor-Lagrange)
Si la fonction $f(x)$ est continue ainsi que ses n premières dérivées sur l'intervalle $[a, b]$, et si $f^{(n+1)}$ existe dans $]a, b[$, pour tout $x \in [a, b]$, il existe $\xi \in\,]a, x[$ tel que

$$f(x) = f(a) + (x-a)\,f'(a) + \cdots + \frac{(x-a)^n}{n!}\,f^{(n)}(a) + \frac{(x-a)^{n+1}}{(n+1)!}\,f^{(n+1)}(\xi). \tag{34}$$

⌟

En pratique, la formule de Taylor-Lagrange avec son reste explicite est utile pour trouver des majorations/minorations d'une fonction (si on connaît le signe du reste). La formule (33) est utile pour fournir des développements limités.

Développements limités de fonctions classiques

Les fonctions classiques qu'on va considérer sont indéfiniment dérivables en 0, donc de la formule de Taylor on déduit aisément

$$\mathrm{e}^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + o(x^n), \tag{35a}$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + (-1)^n\,\frac{x^{2\,n+1}}{(2\,n+1)!} + o(x^{2\,n+1}), \tag{35b}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + (-1)^n\,\frac{x^{2\,n}}{(2\,n)!} + o(x^{2\,n}), \tag{35c}$$

$$\tan x = x + \frac{x^3}{3} + o(x^3), \tag{35d}$$

$$(1+x)^\alpha = 1 + \alpha\,x + \frac{\alpha{\cdot}(\alpha-1)}{2!}\,x^2 + \cdots + \frac{\alpha{\cdot}(\alpha-1)\cdots(\alpha-n+1)}{n!}\,x^n + o(x^n), \tag{35e}$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + (-1)^{n-1}\,\frac{x^n}{n} + o(x^n). \tag{35f}$$

Exercice
Calculer en particulier les développements de $\sqrt{1+x}$ et de $(1-x)^{-1}$. Ce dernier est particulièment simple et pouvait être deviné a priori. Pourquoi ? ⌟

Opérations sur les développements limités

On peut effectuer des opérations d'addition, de multiplication et de division, de composition et d'intégration de développements limités, à condition d'effectuer ces développements à des ordres cohérents avec l'ordre désiré pour le résultat final. Ces manipulations sont d'usage très courant pour le physicien !

Exemple 41
Calculons le développement de $\tan x$ à l'ordre 3 en 0. On écrit

$$\sin x = x - \frac{x^3}{6} + o(x^3), \tag{36a}$$

$$\cos x = 1 - \frac{x^2}{2} + o(x^2), \tag{36b}$$

qu'on inverse en utilisant le développement de $1/(1-u)$, d'où

$$\begin{aligned}\frac{\sin x}{\cos x} = \frac{x - x^3/6 + o(x^3)}{1 - x^2/2 + o(x^2)} &= x\cdot\left(1 - \frac{x^2}{6} + o(x^2)\right)\cdot\left(1 + \frac{x^2}{2} + o(x^2)\right)\\ &= x\cdot\left(1 + \frac{x^2}{3} + o(x^2)\right) = x + \frac{x^3}{3} + o(x^3).\end{aligned}$$ ⌟

Développements limités en $a \neq 0$ ou à l'infini

On obtient un développement limité d'une fonction f au voisinage d'un point a grâce à la formule de Taylor (33), au prix du calcul des dérivées successives en a de la fonction f. Pour les fonctions usuelles, on peut aussi souvent se ramener au développement au voisinage de l'origine $x = 0$, en effectuant un changement de variable.

Exemple 42
Développons $\sin x$ au voisinage de a. On écrit $\sin x = \sin(a + x - a) = \sin a \cos(x - a) + \cos a \sin(x - a)$ et on utilise les développements précédents de $\sin(x - a)$ et $\cos(x - a)$ au voisinage de $x - a = 0$. ⌟

Exemple 43
De même, pour $x \approx a$, $\ln x = \ln(a + x - a) = \ln(a\cdot[1 + (x - a)/a]) = \ln a + \ln(1 + (x - a)/a)$ et on applique la dernière formule (35a) à la variable $x' = (x - a)/a$. ⌟

Mais on peut aussi aussi transposer la discussion au cas d'infiniment grands pour une variable $x \to \pm\infty$: on se ramène au cas précédent par un changement de variable $x = 1/u$. On sait donc aussi écrire des développements limités au voisinage de l'infini.

Exemple 44
Développement limité de $(1 + x^2)^{1/2}$ quand $x \to \infty$. On écrit $(1 + x^2)^{1/2} = |x|\cdot(1 + x^{-2})^{1/2}$ et on applique le développement de $(1 + u)^{1/2}$ au voisinage de $u = x^{-2} = 0$. ⌟

Annexe IV.1. Identités trigonométriques

Elles découlent toutes de cinq identités fondamentales :

$$\sin(-x) = -\sin x, \tag{37a}$$

$$\cos(-x) = \cos x, \tag{37b}$$

$$\sin^2 x + \cos^2 x = 1, \tag{37c}$$

$$\cos x = \sin\left(\frac{\pi}{2} - x\right), \tag{37d}$$

$$\cos(x - y) = \cos x \cos y + \sin x \sin y. \tag{37e}$$

Par exemple, (37c) implique pour $\tan x = \sin x/\cos x$ et $\cot x = 1/\tan x$ que

$$1 + \tan^2 x = \frac{\sin^2 x + \cos^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} \tag{38a}$$

$$\text{et} \quad 1 + \cot^2 x = \frac{\sin^2 x + \cos^2 x}{\sin^2 x} = \frac{1}{\sin^2 x}. \tag{38b}$$

Décalage de l'argument

Les propriétés fondamentales de périodicité sont

$$\sin(x + 2\,\pi) = \sin x, \tag{39a}$$

$$\cos(x + 2\,\pi) = \cos x, \tag{39b}$$

$$\tan(x + \pi) = \tan x. \tag{39c}$$

On a également

$$\sin(x + \pi) = \cos\left(x + \frac{\pi}{2}\right) = -\sin x, \tag{40a}$$

$$\cos(x + \pi) = -\sin\left(x + \frac{\pi}{2}\right) = -\cos x \tag{40b}$$

et

$$\sin\left(x + \frac{\pi}{2}\right) = \sin\left(\frac{\pi}{2} - (-x)\right) = \cos(-x) = \cos x, \tag{41a}$$

$$\cos\left(x + \frac{\pi}{2}\right) = \cos\left(\frac{\pi}{2} - (-x)\right) = \sin(-x) = -\sin x, \tag{41b}$$

$$\tan\left(x \pm \frac{\pi}{2}\right) = -\cot x. \tag{41c}$$

Formules d'addition

$$\cos(x + y) = \cos x \cos y - \sin x \sin y. \tag{42a}$$

$$\sin(x + y) = \cos\left(\frac{\pi}{2} - x - y\right) = \sin x \cos y + \cos x \sin y. \tag{42b}$$

$$\tan(x + y) = \frac{\sin x \cos y + \cos x \sin y}{\cos x \cos y - \sin x \sin y} = \frac{\tan x + \tan y}{1 - \tan x \tan y}. \tag{42c}$$

Inversement,

$$\cos x \cos y = \frac{1}{2}\,(\cos(x + y) + \cos(x - y)), \tag{43a}$$

$$\sin x \sin y = \frac{1}{2}\,(\cos(x - y) - \cos(x + y)), \tag{43b}$$

$$\sin x \cos y = \frac{1}{2}\,(\sin(x + y) + \sin(x - y)), \tag{43c}$$

ou encore

$$\cos a + \cos b = 2\,\cos\left(\frac{a + b}{2}\right)\cos\left(\frac{a - b}{2}\right), \text{ etc.} \tag{44}$$

Formules de doublement

$$\begin{aligned} \cos(2\,x) &= \cos^2 x - \sin^2 x = 2\,\cos^2 x - 1 \\ &= 1 - 2\,\sin^2 x = \frac{2}{1 + \tan^2 x} - 1 = \frac{1 - \tan^2 x}{1 + \tan^2 x} \end{aligned} \tag{45a}$$

$$\sin(2\,x) = 2\,\sin x \cos x = 2\,\tan x \cos^2 x = \frac{2\,\tan x}{1 + \tan^2 x} \tag{45b}$$

$$\tan(2\,x) = \frac{2\,\tan x}{1 - \tan^2 x}. \tag{45c}$$

Noter que $\cos x$, $\sin x$ et $\tan x$ sont des fractions rationnelles de $\tan(x/2)$.

Annexe IV.2. Dérivées des fonctions élémentaires

$$\frac{\mathrm{d}}{\mathrm{d}x} x^{\alpha} = \alpha\, x^{\alpha-1}. \tag{46a}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \mathrm{e}^{x} = \mathrm{e}^{x}. \tag{46b}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} a^{x} = \ln a\, a^{x}. \tag{46c}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \ln x = \frac{1}{x}. \tag{46d}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \sin x = \cos x. \tag{46e}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \cos x = -\sin x. \tag{46f}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \tan x = 1 + \tan^2 x. \tag{46g}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \arcsin x = \frac{1}{\sqrt{1-x^2}}. \tag{46h}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \arccos x = -\frac{1}{\sqrt{1-x^2}}. \tag{46i}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \arctan x = \frac{1}{1+x^2}. \tag{46j}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \operatorname{sh} x = \operatorname{ch} x. \tag{46k}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \operatorname{ch} x = \operatorname{sh} x. \tag{46l}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \operatorname{th} x = 1 - \operatorname{th}^2 x. \tag{46m}$$

Chapitre V

I ´

V. . Intégrale de Riemann

Définition 25 (intégrale d'une fonction en escalier)
Une fonction f de $[a, b]$ dans un espace vectoriel normé E est une fonction en escalier, ce que l'on note $f \in \mathcal{E}([a, b] \to E)$, s'il existe une subdivision finie $(x_0, \ldots, x_n)$ de $[a, b]$ *adaptée* à f, c.-à-d. telle que $\forall\, i \in [\![0, n-1]\!]$, $\forall\, x \in \,]x_i, x_{i+1}[$, $f(x) = f_i$, où $f_i \in E$ est une constante.

On appelle *intégrale* de f et l'on note $\int_a^b f$ la quantité $\sum_{i=0}^{n-1} f_i \cdot (x_{i+1} - x_i)$. ⌟

Remarque
Cette définition a un sens car toute subdivision adaptée à f donne le même résultat. ⌟

Définition 26 (intégrale de Riemann)
Une fonction $f : [a, b] \to E$ est *intégrable au sens de Riemann* s'il existe une suite de fonctions (cf. § VI.A) $(\phi_n) \in \mathcal{E}([a, b] \to E)$ et une suite de fonctions $(\psi_n) \in \mathcal{E}([a, b] \to \mathbb{R})$ telle que

1. $\forall\, n \in \mathbb{N}, \forall\, x \in [a, b]$, on ait $\|f(x) - \phi_n(x)\| \leqslant \psi_n(x)$

et

2. $\lim_{n\to\infty} \int_a^b \psi_n = 0$.

On appelle *intégrale de Riemann* de f et on note $\int_a^b f$ la quantité $\lim_{n\to\infty} \int_a^b \phi_n$. ⌟

Remarques

1. Pour mettre en évidence la variable d'intégration, on écrira souvent $\int_{x=a}^b f(x)\,\mathrm{d}x$ au lieu de $\int_a^b f$.
   Les notations $\int_{[a,\, b]} f$ et $\int_{x \in [a,\, b]} f(x)\,\mathrm{d}x$ sont aussi utilisés ; elles présentent l'intérêt d'être généralisables aux intégrales multiples.
2. Le nombre $\int_a^b f$ (avec $(a, b) \in \mathbb{R}^2$ et f définie sur $[a, b]$) est parfois appelé *intégrale définie*, par opposition, d'une part, à l'*intégrale indéfinie* ou *primitive* de f, qui est la fonction $x \mapsto \int_a^x f$, et, d'autre part, à l'*intégrale impropre*, où f n'est pas définie en a ou b.
3. Il existe d'autres définitions plus générales (mais plus compliquées) de l'intégrale que celle de Riemann, notamment celle de Lebesgue. Celle-ci permet ainsi d'attribuer une valeur (0 en l'occurrence) à $\int_0^1 f$ pour la fonction valant $f(x) = 1$ si $x \in \mathbb{Q}$ et $f(x) = 0$ si $x \in \mathbb{R} \setminus \mathbb{Q}$. En revanche, cette fonction ne peut être approchée par une suite de fonctions en escalier et n'est donc pas intégrable au sens de Riemann.
4. Si la fonction f est en escalier, la définition 26 redonne évidemment le même résultat que la définition 25. Toutes les définitions de l'intégrale (Riemann, Lebesgue, Stieltjes) donnent le même résultat pour les fonctions en escalier et, plus généralement, pour les fonctions continues par morceaux.
5. La valeur de f est indépendante des suites de fonctions ϕ_n et ψ_n. ⌟

Théorème 34
Si $f : [a, b] \to E$ est intégrable sur $[a, b]$, alors $\|f\| : [a, b] \to \mathbb{R}$ est intégrable sur $[a, b]$ et $\left\|\int_a^b f\right\| \leqslant \int_a^b \|f\|$. ⌟

Théorème 35
Si f est continue par morceaux sur $[a, b]$, f est intégrable sur $[a, b]$. ⌟

Théorème 36
Si f est une fonction à valeurs dans $\mathbb{R}$ et est intégrable sur $[a, b]$,

$$\int_a^b f = \lim_{n\to\infty} \sum_{i=0}^{n-1} S_i, \quad \text{où} \quad S_i = \frac{f(x_i) + f(x_{i+1})}{2} \frac{b-a}{n} \quad \text{et} \quad x_i = a + i\,\frac{b-a}{n}.$$

⌟

Pour comprendre la signification du terme de droite, supposons f positive et découpons l'intervalle $[a, b]$ en n intervalles de largeur identique, $(b-a)/n$: S_i est l'aire du trapèze compris entre les droites $y = 0$, $x = x_i$, $x = x_{i+1}$ et la corde reliant les points $(x_i, f(x_i))$ et $(x_{i+1}, f(x_{i+1}))$. Comme $\lim_{n\to 0}((b-a)/n) = 0$, $\lim_{n\to\infty} \sum_{i=0}^{n-1} S_i$ est l'aire de la surface comprise entre le graphe de f et les droites $y = 0$, $x = a$ et $x = b$.

$\sum_{i=0}^{n-1} S_i$ est un cas particulier de *somme de Riemann*. Il en existe d'autres : par exemple

$$\sum_{i=0}^{n-1} f(x_i)\cdot\frac{b-a}{n} \quad \text{ou} \quad \sum_{i=0}^{n-1} f\Big(\frac{x_i + x_{i+1}}{2}\Big)\cdot\frac{b-a}{n}\ ;$$

si f est intégrable, toutes convergent vers $\int_a^b f$.

Définition 27
Soit f une fonction définie sur $[a, b]$. Une fonction F est une *primitive* de f sur $[a, b]$ si $\forall\, x \in [a, b]$, F est dérivable en x et $F'(x) = f(x)$. ⌟

Théorème 37 (théorème fondamental de l'analyse)
Si f est une fonction continue sur $[a, b]$, la fonction $F\colon x \mapsto F(x) = \int_{t=a}^{x} f(t)\,\mathrm{d}t$ est une primitive de f.

En outre, si F est une primitive de f, on a $\int_{x=a}^{b} f(x)\,\mathrm{d}x = F(b) - F(a) =: \Big[F(x)\Big]_{x=a}^{b}$. ⌟

Remarques
1. Si F est une primitive de f, les primitives de f sont toutes et seulement les fonctions de la forme $\Phi(x) = F(x) + \text{constante}$. Quelle que soit la primitive Φ de f, $\int_a^b f(x)\,\mathrm{d}x = \Big[\Phi(x)\Big]_a^b$.
2. La *primitive* d'une fonction f est aussi appelée *intégrale indéfinie* et est souvent notée $\int^x f(t)\,\mathrm{d}t$, sans préciser la borne inférieure de l'intégrale, puisque changer celle-ci revient à ajouter une constante. ⌟

V. . Propriétés de l'intégrale définie

L'intégration est une opération linéaire :

Théorème 38
Soient E un K-espace vectoriel, $\lambda \in K$, et f et g deux applications à valeurs dans E et intégrables sur $[a, b]$.

Alors les intégrales ci-dessous existent,

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g$$

et

$$\int_a^b (\lambda\, f) = \lambda \int_a^b f.$$

⌟

Par extension, on peut donc intégrer une somme *finie* de fonctions terme à terme.

Théorème 39
Soient f et g deux applications à valeurs dans $\mathbb{R}$ et intégrables sur $[a, b]$.

Si $\forall\, x \in [a, b]$, $f(x) \leqslant g(x)$, alors $\int_a^b f \leqslant \int_a^b g$. ⌟

L'inverse est évidemment faux.

Théorème 40 (relation de Chasles)
Si $c \in [a, b]$, $\int_a^b f = \int_a^c f + \int_c^b f$. ⌟

Si $b < a$, on pose $\int_a^b f := -\int_b^a f$. Cette définition est cohérente avec la relation de Chasles, puisque $\int_a^a f = \int_a^b f + \int_b^a f = -\int_b^a f + \int_b^a f = 0$.

Théorème 41 (première formule de la moyenne)
Si f est continue sur $[a, b]$ et si g est intégrable et de signe constant sur $[a, b]$, il existe $c \in [a, b]$ tel que

$$\int_a^b f(x)\, g(x)\, \mathrm{d}x = f(c) \int_a^b g(x)\, \mathrm{d}x. \tag{1}$$

⌟

Preuve
Supposons $g \geqslant 0$ et notons m et M les bornes inférieure et supérieure de f sur $[a, b]$. Pour tout $x \in [a, b]$, on a $m\, g(x) \leqslant f(x)\, g(x) \leqslant M\, g(x)$.

Si $\int_a^b g = 0$, la démonstration est triviale.

Sinon,

$$m \leqslant y := \frac{\int_a^b f\, g}{\int_a^b g} \leqslant M.$$

La fonction f étant continue, $\forall\, y \in f([a, b]) = [m, M]$, $\exists\, c \in [a, b]$ tel que $y = f(c)$ d'après le théorème des valeurs intermédiaires. ⌟

Corollaire 1
En prenant $g = 1$, on obtient en particulier qu'il existe $c \in [a, b]$ tel que

$$\int_a^b f(x)\, \mathrm{d}x = f(c)\cdot(b - a). \tag{2}$$

⌟

Autrement dit, la moyenne $\int_a^b f(x)\, \mathrm{d}x/(b - a)$ est égale à la valeur prise par la fonction f en (au moins) un point de l'intervalle $[a, b]$.

Inégalité de (Cauchy-)Schwarz

Soient f et g deux fonctions intégrables sur $[a, b]$. Pour tout réel λ, on peut écrire l'inégalité triviale $\int_a^b (f(x) + \lambda\, g(x))^2\, \mathrm{d}x \geqslant 0$, soit

$$\lambda^2 \int_a^b (g(x))^2\, \mathrm{d}x + 2\, \lambda \int_a^b f(x)\, g(x)\, \mathrm{d}x + \int_a^b (f(x))^2\, \mathrm{d}x \geqslant 0. \tag{3}$$

Le trinôme du second degré en λ est positif ou nul pour tout λ réel si et seulement si son discriminant est négatif ou nul. En effet, si le discriminant est strictement positif, le trinôme admet deux racines réelles distinctes et est négatif entre celles-ci puisque $\int g^2 \geqslant 0$. On a donc

$$\left| \int_a^b f(x)\, g(x)\, \mathrm{d}x \right| \leqslant \sqrt{\int_a^b (f(x))^2\, \mathrm{d}x}\; \sqrt{\int_a^b (g(x))^2\, \mathrm{d}x}\,, \tag{4}$$

une inégalité parfois très utile.

On peut généraliser ceci au cas de fonctions à valeurs complexes :

Théorème 42 (inégalité de (Cauchy-)Schwarz)
Soient f et g deux applications intégrables de $[a, b]$ dans $\mathbb{C}$.

On a

$$\left| \int_a^b f\, \overline{g} \right| \leqslant \sqrt{\int_a^b |f|^2}\; \sqrt{\int_a^b |g|^2}\,.$$

⌟

Cette inégalité rappelle l'inégalité sur le produit scalaire de deux vecteurs de $\mathbb{R}^n$,

$$|\vec{u} \cdot \vec{v}| \leqslant \|\vec{u}\|\, \|\vec{v}\|,$$

et il y a de bonnes raisons à cela : $\int_a^b f\,\overline{g}$ définit en effet un produit scalaire (hermitien) entre des fonctions f et g continues sur $[a, b]$, et $\int_a^b f\,\overline{f} = \int_a^b |f|^2$ est le carré de la norme associée. On a d'ailleurs

$$\sqrt{\int_a^b |f+g|^2} \leqslant \sqrt{\int_a^b |f|^2} + \sqrt{\int_a^b |g|^2} \qquad \text{(inégalité de Minkowski).}$$

V. . Méthodes d'intégration

En pratique, pour calculer effectivement une intégrale, il est indispensable de bien connaître les formules de *dérivation* des fonctions simples (voir appendice IV.2) et de les combiner avec quelques astuces.

V. .1. Intégration par parties

Théorème 43
Soient $(f, g) \in (C^1([a, b] \to \mathbb{K}))^2$.
Alors,

$$\int^x f'\,g = f\,g - \int^x f\,g'$$

et

$$\int_a^b f'\,g = \left[f\,g\right]_a^b - \int_a^b f\,g'.$$

⌟

Ceci découle de $(f\,g)' = f'\,g + f\,g'$.

Ce théorème permet donc de calculer la primitive d'une fonction h de la forme $h = f'\,g$ si l'on connaît la primitive de $f\,g'$. Ça ne présente d'intérêt que si cette dernière est plus simple à calculer !

Exemple 45

$$\int \ln x \,\mathrm{d}x = \int (x)' \ln x \,\mathrm{d}x = x \ln x - \int x{\cdot}(\ln x)' \,\mathrm{d}x = x \ln x - \int x\,\frac{1}{x}\,\mathrm{d}x = x{\cdot}(\ln x - 1). \tag{5}$$

⌟

Exercice
Calculer de même $\int x \ln x \,\mathrm{d}x$. ⌟

V. .2. Changement de variable

Théorème 44
Soient $(a, b) \in \mathbb{R}^2$ (avec $a \leqslant b$), I un intervalle de $\mathbb{R}$, $g \in C^1([a, b] \to I)$ et $f \in C^0(I \to \mathbb{R})$.
Alors,

$$\int_{t=a}^{b} f(g(t))\,g'(t)\,\mathrm{d}t = \int_{x=g(a)}^{g(b)} f(x)\,\mathrm{d}x. \tag{6}$$

⌟

Remarques
- On a cette fois utilisé la propriété de dérivation d'une fonction composée, $(F \circ g)' = g' \times F' \circ g$, et on l'a appliquée à $f = F'$.
- Pour se souvenir de cette formule, le plus simple est de remplacer x par $g(t)$ partout :
  - dans la fonction f ;
  - dans la différentielle, ce qui donne $\mathrm{d}x = (\mathrm{d}x/\mathrm{d}t){\cdot}\mathrm{d}t = g'(t)\,\mathrm{d}t$;
  - dans les bornes. Donc, si t varie de a à b, x doit varier de $g(a)$ à $g(b)$. ⌟

Exemple 46

$$\int \frac{x\,\mathrm{d}x}{1+x^2} = \frac{1}{2}\int \frac{\mathrm{d}t}{1+t} = \frac{1}{2}\ln(1+t) = \frac{1}{2}\ln(1+x^2), \tag{7}$$

où on a effectué le changement de variable $x \mapsto t = x^2$. ⌟

Exemple 47
Pour calculer $\int \mathrm{d}x/(1+x^2)$, on pose $x = \tan t$, et l'intégrale devient simplement $\int \mathrm{d}t$, donc

$$\int \frac{\mathrm{d}x}{1+x^2} = t = \arctan x, \tag{8}$$

la fonction réciproque de $\tan x$, définie pour tout x et prenant ses valeurs dans $]-\pi/2, \pi/2[$.

De même,

$$\int \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \arcsin x, \tag{9}$$

la fonction réciproque de $\sin x$, définie sur $[-1, 1]$ et prenant ses valeurs dans $[-\pi/2, \pi/2]$. ⌟

L'équation (6) peut aussi s'utiliser de droite à gauche pour calculer $\int_{x=\alpha}^{\beta} f(x)\,\mathrm{d}x$, où $(\alpha,\beta) \in I^2$ et $f \in C^0(I \to \mathbb{R})$. Il faut seulement qu'existent $(a, b) \in \mathbb{R}^2$ et $g \in C^1([a, b] \to I)$ tels que $g(a) = \alpha$ et $g(b) = \beta$. On a alors

$$\int_{x=\alpha}^{\beta} f(x)\,\mathrm{d}x = \int_{t=a}^{b} f(g(t))\, f'(t)\,\mathrm{d}t.$$

Il est essentiel que $g([a, b]) \subset I$ ou, ce qui revient au même, que $f \in C^0(g([a, b]) \to \mathbb{R})$. Or, d'une manière générale,

$$(g(a) = \alpha,\ g(b) = \beta) \not\Rightarrow g([a, b]) = [\inf(\alpha,\ \beta), \sup(\alpha, \beta)]\ (\subset I).$$

C'est le cas en revanche si g est bijective sur $[a, b]$. On a alors le théorème suivant :

Théorème 45
Soient $(\alpha, \beta, a, b) \in \mathbb{R}^4$ (avec $\alpha \leqslant \beta$), $f \in C^0([\alpha, \beta] \to \mathbb{R})$, et g une application bijective et C^1 de $[a, b]$ (ou $[b, a]$ si $b < a$) dans $[\alpha, \beta]$ telle que $g(a) = \alpha$ et $g(b) = \beta$.

Alors,

$$\int_{x=\alpha}^{\beta} f(x)\,\mathrm{d}x = \int_{t=g^{\langle -1\rangle}(\alpha)}^{g^{\langle -1\rangle}(\beta)} f(g(t))\, g'(t)\,\mathrm{d}t. \tag{10}$$

⌟

Remarque
g étant continue, elle est bijective si elle est strictement monotone. Comme elle est même C^1, il suffit que $\forall\ t \in [a, b]$, $g'(t) \neq 0$. ⌟

Si la fonction g n'est pas strictement monotone, il est plus sage de décomposer l'intervalle d'intégration en sous-intervalles où elle l'est. Faute de quoi, on risque de trouver des absurdités.

V. .3. Changement de variable pour une intégrale multiple

La théorie des intégrales multiples dépassant le cadre de ce cours, nous allons simplement expliquer comment procéder à un changement de variable, une technique très utile en physique.

Soit une bijection

$$\begin{aligned} g = (g_1, \ldots, g_n)\colon \quad & D_t \longrightarrow D_x \\ & t = (t_1, \ldots, t_n) \longmapsto x = (x_1, \ldots, x_n), \end{aligned}$$

où D_t et D_x sont les « domaines » de $\mathbb{R}^n$ balayés par t et x, et $x_i = g_i(t)$. On a

$$\int \cdots \int_{D_x} f(x)\,\mathrm{d}x_1 \cdots \mathrm{d}x_n = \int \cdots \int_{D_t} f(g[t])\,|\text{dét}\, J(t)|\,\mathrm{d}t_1 \cdots \mathrm{d}t_n,$$

où

$$J(t) = \begin{pmatrix} \partial g_1/\partial t_1 & \ldots & \partial g_1/\partial t_n \\ \vdots & \ddots & \vdots \\ \partial g_n/\partial t_1 & \ldots & \partial g_n/\partial t_n \end{pmatrix}$$

est la *matrice jacobienne* et dét $J(t)$, son déterminant, porte le nom de *jacobien*.

Dans le cas où $n = 1$, on retrouve bien le résultat obtenu précédemment : le jacobien vaut alors $\partial g/\partial t = g'(t)$ et $\int_{D_x} f(x)\,\mathrm{d}x = \int_{D_t} f(g[t])\,|g'(t)|\,\mathrm{d}t$.

Remarque
Si $D_x = [\alpha, \beta]$, $a = g^{\langle -1\rangle}(\alpha)$ et $b = g^{\langle -1\rangle}(\beta)$, on a $D_t = [\inf\{a,\ b\}, \sup\{a, b\}]$. Il ne faut donc pas oublier la valeur absolue autour du jacobien. En effet, si g est décroissante, $D_t = [b, a]$ et l'on retrouve bien

$$\int_{x=\alpha}^{\beta} f(x)\,\mathrm{d}x = \int_{t=b}^{a} f(g[t])\,|g'(t)|\,\mathrm{d}t = -\int_{t=b}^{a} f(g[t])\,g'(t)\,\mathrm{d}t = \int_{t=a}^{b} f(g[t])\,g'(t)\,\mathrm{d}t.$$
⌟

Considérons à titre d'exemple le passage des coordonnées polaires aux coordonnées cartésiennes [(*1)],

$$g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} : \begin{pmatrix} t_1 = \rho \\ t_2 = \phi \end{pmatrix} \mapsto \begin{pmatrix} x_1 = x = g_1(\rho, \phi) = \rho\cos\phi \\ x_2 = y = g_2(\rho, \phi) = \rho\sin\phi \end{pmatrix}. \tag{11}$$

La matrice jacobienne est

$$J = \begin{pmatrix} \dfrac{\partial(\rho\cos\phi)}{\partial\rho} = \cos\phi & \dfrac{\partial(\rho\cos\phi)}{\partial\phi} = -\rho\sin\phi \\ \dfrac{\partial(\rho\sin\phi)}{\partial\rho} = \sin\phi & \dfrac{\partial(\rho\sin\phi)}{\partial\phi} = \rho\cos\phi \end{pmatrix},$$

donc dét $J = \rho$ et

$$\iint_{D_{(x,\,y)}} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \iint_{D_{(\rho,\,\phi)}} f(\rho\cos\phi, \rho\sin\phi)\,\rho\,\mathrm{d}\rho\,\mathrm{d}\phi.$$

Exemple 48
Calculons $\int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y$ (cf. § V.D.2 pour la définition de l'intégrale quand les bornes sont infinies). Au domaine $D_{(x,\,y)} =]-\infty, \infty[\times]-\infty, \infty[$ correspond le domaine $D_{(\rho,\,\phi)} = [0, \infty[\times[0, 2\pi[$. Par ailleurs, $x^2 + y^2 = \rho^2$, donc

$$\int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y = \int_{\rho=0}^{\infty}\int_{\phi=0}^{2\pi} \mathrm{e}^{-\rho^2}\,\rho\,\mathrm{d}\rho\,\mathrm{d}\phi = 2\pi\int_{\rho=0}^{\infty} \mathrm{e}^{-\rho^2}\,\rho\,\mathrm{d}\rho.$$

Par application du changement de variable $\rho \mapsto \lambda = \rho^2$, on trouve que

$$\int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y = 2\pi\int_{\lambda=0}^{\infty} \mathrm{e}^{-\lambda}\,\frac{\mathrm{d}\lambda}{2} = \pi.$$

Comme

$$\int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y = \int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} \mathrm{e}^{-x^2}\,\mathrm{e}^{-y^2}\,\mathrm{d}x\,\mathrm{d}y = \left(\int_{x=-\infty}^{\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x\right)\cdot\left(\int_{y=-\infty}^{\infty} \mathrm{e}^{-y^2}\,\mathrm{d}y\right) = \left(\int_{x=-\infty}^{\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x\right)^2,$$

on obtient que

$$\int_{x=-\infty}^{\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x = \sqrt{\pi}$$

Par changement de variable $x \mapsto (x-\mu)/\sigma$, on a encore

$$\int_{-\infty}^{\infty} \mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}\,\mathrm{d}x = \sqrt{2\pi}\,\sigma, \tag{12}$$

une intégrale fort importante en théorie des probabilités. En effet,

$$\frac{1}{\sqrt{2\pi}\,\sigma}\,\mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}$$

est la densité de probabilité d'un processus gaussien de moyenne μ et de variance σ^2. L'intégrale (probabilité totale) est bien égale à 1. ⌟

1. x représente ici l'abscisse x_1 et non (x_1, x_2).

V. . Intégrales impropres

On vient de voir que si la fonction $f(x)$ est définie et continue par morceaux en tout point d'un intervalle $[a, b]$ de la droite réelle, elle est intégrable au sens de Riemann sur cet intervalle : l'intégrale définie $\int_a^b f(x)\,\mathrm{d}x$ existe et est finie.

Si $b = \infty$ ou si b est fini mais est une singularité de f (c.-à-d. que f tend vers l'infini en b ou oscille de plus en plus vite au voisinage de b), f n'est pas intégrable au sens de Riemann sur $[a, b[$. On peut néanmoins dans certains cas définir une intégrale dite impropre sur cet intervalle.

Définition 28
Soient I un intervalle de $\mathbb{R}$, E un espace vectoriel normé et $f : I \to E$ une application.

f est *localement intégrable* sur I si et seulement si $\forall\,[a, b] \subset I$, $\int_a^b f$ existe. ⌟

Définition 29
Soit $f : [a, b[\to E$, avec $b \in \mathbb{R} \cup \{+\infty\}$ et E un espace vectoriel normé.

- Si f est localement intégrable sur $[a, b[$ et
- si $\lim_{x\to b^-} \int_a^x f(t)\,\mathrm{d}t$ existe et est finie,

on dit que l'*intégrale impropre* (ou *généralisée*) $\int_a^b f$ converge et l'on note $\int_a^b f = \lim_{x\to b^-} \int_a^x f(t)\,\mathrm{d}t$. ⌟

On définit de la même manière l'intégrale impropre sur $]a, b]$ par $\lim_{x\to a^+} \int_x^b f(t)\,\mathrm{d}t$ si $a = -\infty$ ou est une singularité de f.

Si l'intégrale est impropre en ses deux bornes, elle n'est définie que si, pour un point $c \in\,]a, b[$ quelconque, les intégrales $\int_a^c f$ et $\int_c^b f$ convergent *séparément* ; on a alors $\int_a^b f := \int_a^c f + \int_c^b f$ (on obtient la même valeur quel que soit c). Ainsi, $\int_{-\infty}^{\infty} f = \lim_{x\to\infty} \int_{-x}^0 f + \lim_{x\to\infty} \int_0^x f$. En revanche, $\lim_{x\to\infty} \int_{-x}^x f$ peut très bien exister sans que $\int_{-\infty}^{\infty} f$ existe.

Théorème 46
Soit $f : [a, b[\to E$ une application. Si il existe une application $g : [a, b] \to E$ prolongeant f en b et telle que $\int_a^b g$ existe au sens de Riemann, alors $\int_a^b f$ existe et $\int_a^b f = \int_a^b g$. On dit alors que $\int_a^b f$ est *faussement impropre*. ⌟

En particulier, une fonction continue sur $[a, b[$ prolongeable par continuité sur $[a, b]$ est intégrable sur $[a, b]$.

Théorème 47 (critère de Cauchy)
Soient E un espace vectoriel normé complet et $f : [a, b[\to E$ une fonction localement intégrable sur $[a, b[$ dont l'intégrale est impropre en b.

$\int_a^b f$ converge si et seulement si elle vérifie le critère de Cauchy, c.-à-d. si

$$\forall\,\epsilon > 0,\ \exists\, c \in [a, b[,\ \forall\,(x, y) \in [a, b[^2,\ (x \geqslant c,\ y \geqslant c) \Longrightarrow \left\| \int_x^y f \right\| \leqslant \epsilon.$$

⌟

Définition 30
$\int_a^b f$ est dite *absolument convergente* si $\int_a^b \|f\|$ est convergente.

Une intégrale convergente mais non absolument convergente est dite *semi-convergente*. ⌟

Théorème 48
Soient E un espace vectoriel normé complet et $f : [a, b[\to E$ une fonction localement intégrable sur $[a, b[$ dont l'intégrale est impropre en b.

Si $\int_a^b f$ est absolument convergente, alors $\int_a^b f$ est convergente et $\left\|\int_a^b f\right\| \leqslant \int_a^b \|f\|$. ⌟

Théorème 49
Soit $f \geqslant 0$ une fonction localement intégrable de $[a, b[$ dans $\mathbb{R}^+$.

$\int_a^b f$ converge si et seulement si $\exists\, M \geqslant 0$, $\forall\, x \in [a, b[$, $\int_a^x f \leqslant M$. ⌟

Théorème 50
Soient f et g deux fonctions localement intégrables de $[a, b[$ dans $\mathbb{R}^+$ telles que $0 \leqslant f \leqslant g$.

$\int_a^b g$ converge $\Longrightarrow \int_a^b f$ converge.
$\int_a^b f$ diverge $\Longrightarrow \int_a^b g$ diverge. ⌟

Théorème 51
Soient f et g deux fonctions localement intégrables de $[a, b[$ dans $\mathbb{R}$ et $g \geqslant 0$.

- Si $\int_a^b g$ converge,
$$f = O(g) \text{ en } b^- \Longrightarrow \int_a^b f \text{ converge et } \int_x^b f = O\left(\int_x^b f\right) \text{ quand } x \to b^-.$$
On obtient la même chose en remplaçant « $= O(\ldots)$ » par « $= o(\ldots)$ » ou « $\simeq$ » à gauche et à droite de « $\Longrightarrow$ ».
- Si $\int_a^b g$ diverge,
$$f = O(g) \text{ en } b^- \Longrightarrow \int_a^x f = O\left(\int_a^x f\right) \text{ quand } x \to b^-.$$
On obtient la même chose en remplaçant « $= O(\ldots)$ » par « $= o(\ldots)$ » ou « $\simeq$ » à gauche et à droite de « $\Longrightarrow$ ».
En outre, si $f \simeq g$ en b^-, $\int_a^b f$ diverge. ⌟

Théorème 52 (intégration par parties)
Soient $(f, g) \in (C^1([a, b[\to \mathbb{R}))^2$.
Si $\lim_{b^-}(f\, g)$ existe et que $\int_a^b f\, g'$ converge, alors $\int_a^b f'\, g$ converge et
$$\int_a^b f'\, g = \left[f\, g\right]_a^b - \int_a^b f\, g'.$$ ⌟

Théorème 53 (changement de variable)
Soient $(\alpha, a) \in \mathbb{R}^2$, $\beta \in \mathbb{R} \cup \{+\infty\}$, $b \in \mathbb{R} \cup \{\pm\infty\}$, $f \in C^0([\alpha, \beta[\to \mathbb{R})$, et g une fonction bijective et C^1 de $[a, b[$ (ou $]b, a]$) dans $[\alpha, \beta[$ telle que $g(a) = \alpha$ et $\lim_{t\to b} g(t) = \beta$.
$\int_\alpha^\beta f(x)\, \mathrm{d}x$ converge si et seulement si $\int_a^b f[g(t)]\, g'(t)\, \mathrm{d}t$ converge ; si c'est le cas,
$$\int_\alpha^\beta f(x)\, \mathrm{d}x = \int_a^b f[g(t)]\, g'(t)\, \mathrm{d}t.$$ ⌟

V. .1. Intégrales impropres avec une singularité sur un intervalle fini

Supposons maintenant que la fonction soit définie sur l'intervalle ouvert $[a, b[$, mais qu'elle ait un point singulier en b, c.-à-d. un point où elle n'est pas définie, soit qu'elle tende vers l'infini, soit qu'elle oscille de plus en plus vite au voisinage de ce point. (Exemple du premier cas, $1/(x - b)$; du deuxième, $\sin(1/(x - b))$).

Exemple 49
$\int_0^1 \sin(1/x)\, \mathrm{d}x$ est absolument convergente. En effet, $0 \leqslant |\sin(1/x)| \leqslant 1$. Or, $\forall\, x \in]0, 1]$, $\int_x^1 |\sin(1/x')|\, \mathrm{d}x' \leqslant 1 - x \leqslant M = 1$, donc $\int_0^1 |\sin(1/x)|\, \mathrm{d}x$ converge par application du théor. 49, ainsi que $\int_0^1 \sin(1/x)\, \mathrm{d}x$ d'après le théor. 48. Le même raisonnement s'applique à $\int_0^1 \cos(1/x)\, \mathrm{d}x$.

$\int_0^1 \sin(1/x)\, \mathrm{d}x/x$ est également convergente.En effet, on obtient en intégrant par parties que
$$\int_0^1 \sin\frac{1}{x}\,\frac{\mathrm{d}x}{x} = \left[x \cos\frac{1}{x}\right]_0^1 - \int_0^1 \cos\frac{1}{x}\, \mathrm{d}x,$$
qui est bien finie. La non-convergence absolue de l'intégrale $\int_0^1 \sin(1/x)\, \mathrm{d}x/x$ est en revanche plus délicate à établir et nécessite de la minorer par une série divergente. ⌟

Exemple 50
Considérons la fonction $g_\alpha(x) = (b - x)^{-\alpha}$ $(x < b)$. Comme sa primitive est
$$\int^x (b - x')^{-\alpha}\, \mathrm{d}x' = \begin{cases} \frac{1}{\alpha-1} (b - x)^{-\alpha+1} & \text{si } \alpha \neq 1, \\ -\ln(b - x) & \text{si } \alpha = 1, \end{cases} \tag{13}$$

cette fonction est intégrable en b si $\alpha < 1$, et non intégrable si $\alpha \geqslant 1$. ⌟

Comme dans les cas précédemment étudiés de limites, on peut trouver des conditions suffisantes d'intégrabilité (c.-à-d. de convergence de l'intégrale) pour une fonction **positive** en la minorant ou en la majorant par des fonctions connues comme les $g_\alpha(x)$ ci-dessus, ou en en faisant une étude asymptotique. Ainsi, si A est une constante et

1. si $0 \leqslant f(x) \leqslant A/(b-x)^\alpha$ pour $\alpha < 1$, $\int_a^b f(x)\,\mathrm{d}x$ converge ;
2. si $f(x) \geqslant A/(b-x)^\alpha > 0$ pour $\alpha \geqslant 1$, l'intégrale diverge ;
3. si $f(x) \simeq A/(b-x)^\alpha$ quand x tend vers b^-, son intégrale de a à b est convergente si $\alpha < 1$, et divergente si $\alpha \geqslant 1$.

Remarque

Il faut bien noter qu'il est donc possible que $f(x) \to \infty$ quand $x \to b$ mais que $\int_a^b f(x)\,\mathrm{d}x$ soit *finie* ! Il suffit de considérer une fonction $g_\alpha(x)$ avec $0 < \alpha < 1$, par exemple $1/\sqrt{x}$ en 0. ⌟

Exemple 51

Soit $f(x) = 1/(x\cdot(x^2+1))$, à intégrer de 0 à 1.

On montre que $f(x) = 1/x - x/(x^2+1)$. Seule l'intégration du premier terme pose problème : dans le cas présent, elle est impossible en 0. Plus simplement, $f(x) \simeq 1/x$ quand $x \to 0$, donc n'est pas intégrable.

Même discussion (avec la conclusion opposée !) pour $g(x) = \sqrt{x}\, f(x)$. ⌟

Exercice

Vérifier que $\int \mathrm{d}x/\sin(x-b) = \ln|\tan((b-x)/2)|$ en utilisant la formule de trigonométrie $\sin u = 2\tan(u/2)/(1+\tan^2(u/2))$. En déduire que l'intégrale $\int_0^b \mathrm{d}x/\sin(x-b)$ n'est pas définie. Donner un argument plus simple reposant sur un équivalent. ⌟

Exercice

Discuter l'intégrabilité d'une fraction rationnelle $\int_a^b P(x)/Q(x)\,\mathrm{d}x$ si Q a des zéros sur l'intervalle $[a, b]$. ⌟

Singularité logarithmique

Un cas important en pratique, qui n'est pas couvert par la comparaison avec les fonctions g_α, est celui d'une singularité logarithmique. Supposons que, quand $x \to b$, $f(x) \simeq \ln(b-x)$. Se rappelant la primitive de $\ln x$,

$$\int \ln x\,\mathrm{d}x = \int (x)' \ln x\,\mathrm{d}x = x\cdot(\ln x - 1), \tag{14}$$

on vérifie que $\int^b \ln(b-x)\,\mathrm{d}x$ converge et donc, par comparaison, que c'est aussi le cas de l'intégrale de la fonction f considérée : *une singularité logarithmique à distance finie est intégrable.*

Exercice

Montrer que $\int^b \ln^n(b-x)\,\mathrm{d}x$ converge (intégration par parties et récurrence), que $\int^b \mathrm{d}x/\ln(b-x)$ est absolument convergente, et que $\int^b \ln(\ln(b-x))\,\mathrm{d}x$ converge (intégration par parties). ⌟

V. .2. Intégrales impropres avec intervalle d'intégration infini

Comme dans le cas d'une singularité à distance finie examiné au paragraphe précédent, dans le cas de fonctions positives, la condition d'intégrabilité à $\pm\infty$ se ramène à l'étude du comportement de la fonction $f(x)$ quand $x \to \pm\infty$. En utilisant à nouveau les fonctions puissances $x^{-\alpha}$ comme étalons de comparaison, $\int^b \mathrm{d}x/x^\alpha$ a une limite si $\alpha > 1$, et n'en a pas si $\alpha \leqslant 1$. Donc, si A est une constante et

1. si $0 \leqslant f(x) \leqslant A/x^\alpha$ pour $\alpha > 1$, $\int_a^\infty f$ existe ;
2. si $f(x) \geqslant A/x^\alpha > 0$ pour $\alpha \leqslant 1$, $A > 0$, l'intégrale diverge ;
3. si $f(x) \simeq A/x^\alpha$ quand $x \to \infty$, son intégrale de a à ∞ est convergente si $\alpha > 1$, divergente si $\alpha \leqslant 1$.

Exemple 52

L'intégrale $\int_0^\infty P(x)/Q(x)\,\mathrm{d}x$, où P et Q sont des polynômes de degrés respectifs p et q, et où l'on suppose toutes les racines de Q strictement négatives ou complexes, converge pour $p \leqslant q-2$. ⌟

Attention : le fait que $f(x) \to 0$ quand $x \to \infty$ n'est pas une condition *suffisante* de convergence de $\int^{\infty} f$.

On a en effet avec les fonctions $1/x^{\alpha}$, $\alpha \leqslant 1$, des exemples de fonctions qui tendent vers zéro à l'infini, mais pas assez vite pour assurer la convergence de l'intégrale.

Exercice

Étudier la convergence de $\int^{\infty} \mathrm{d}x/(x \ln^{\alpha} x)$ selon la valeur de α. ⌟

Le fait que $f(x) \to 0$ quand $x \to \infty$ n'est pas non plus une condition *nécessaire* de convergence de $\int^{\infty} f$.

Exemple 53

Un exemple est fourni par l'intégrale de Fresnel, $\int_1^{\infty} \sin(x^2)\,\mathrm{d}x$, qu'on rencontre en optique dans la théorie de la diffraction de la lumière et dont on démontre qu'elle converge, bien que $|\sin(x^2)|$ n'ait pas de limite en $+\infty$.

Pour montrer la convergence, on peut effectuer un changement de variable $x \mapsto u = x^2$, suivi d'une intégration par parties, qui amène à une intégrale absolument convergente :

$$\int_1^{\infty} \sin(x^2)\,\mathrm{d}x = \frac{1}{2}\int_1^{\infty} \sin u\,\frac{\mathrm{d}u}{\sqrt{u}} = -\frac{1}{2}\left[\frac{\cos u}{\sqrt{u}}\right]_1^{\infty} - \frac{1}{4}\int_1^{\infty} \cos u\,\frac{\mathrm{d}u}{u^{3/2}}. \tag{15}$$

⌟

En revanche, si f admet une limite $+\infty$ et que celle-ci est différente de zéro, $\int^{\infty} f$ diverge.

Chapitre VI

S ´

Dans ce chapitre, nous allons combiner plusieurs des notions que nous avons traitées jusqu'à maintenant. Nous considérerons d'abord des suites et des séries de fonctions, en particulier des séries dites entières. Ensuite, nous examinerons les difficultés rencontrées lorsqu'on manipule des expressions contenant plusieurs opérations de limite, de sommation, d'intégration et de dérivation.

VI. . Suites de fonctions

On considère maintenant des *suites de fonctions* $(x \mapsto f_n(x))_{n\in\mathbb{N}}$ où $f_n : A \to F$ est une application d'un ensemble A dans un espace vectoriel normé F.

Exemples : $x \mapsto f_n(x) = n \sin(x/n)$; $x \mapsto f_n(x) = n \sin x/(1 + n\, x)$.

Définition 31 (convergence simple)
Soient A un ensemble, F un espace vectoriel normé, $(f_n : A \to F)_{n\in\mathbb{N}}$ une suite de fonctions définies sur A, et f une application de A dans F.

La suite de fonctions $(f_n)_{n\in\mathbb{N}}$ converge (simplement) vers f sur A si et seulement si la suite $(f_n(x))_{n\in\mathbb{N}}$ converge vers $f(x)$ pour tout x de A. ⌟

$(f_n)_{n\in\mathbb{N}}$ converge donc vers f sur A ssi

$$\forall\, \epsilon > 0, \forall\, x \in A, \exists\, N(x) \in \mathbb{N}, \forall\, n \in \mathbb{N}, n \geqslant N(x) \Longrightarrow \|f_n(x) - f(x)\| \leqslant \epsilon, \tag{1}$$

où l'on a bien souligné que l'entier N dépend a priori de x.

Cette condition de convergence est parfois trop faible. Par exemple, il n'est pas vrai en général que si les fonctions f_n sont continues et convergent simplement, leur limite f est continue.

Exemple 54
Soit la suite de fonctions définies par $f_n(x) = x^n$ sur l'intervalle $[0, 1]$. Pour tout $x < 1$, $x^n \to 0$, tandis que pour $x = 1$, la valeur limite est 1. La fonction limite $f(x)$ est donc la fonction discontinue qui vaut 0 sur $[0, 1[$ et 1 en 1 (c'est la fonction $\mathrm{E}(x)$ définie p. 29, restreinte à l'intervalle $[0, 1]$). ⌟

En fait, la dépendance de N en x dans (1) est souvent gênante et on a besoin d'une condition de convergence plus forte (c.-à-d. plus contraignante), où l'on impose à N d'être indépendant de x :

Définition 32 (convergence uniforme)
Une suite de fonctions $(f_n)_{n\in\mathbb{N}}$ est dite *uniformément convergente* vers la fonction f sur A si et seulement si

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}, \forall\, n \in \mathbb{N}, \forall\, x \in A, n \geqslant N \Longrightarrow \|f_n(x) - f(x)\| \leqslant \epsilon. \tag{2}$$

On dit alors que f est la limite uniforme des f_n. ⌟

Remarque
On a donc convergence uniforme de (f_n) vers f sur A si et seulement si

$$\lim_{n\to\infty}\left(\sup_{x\in A} \|f_n(x) - f(x)\|\right) = 0.$$

⌟

Théorème 54
Si (f_n) converge uniformément vers f sur A, (f_n) converge simplement vers f sur A. ⌟

L'inverse est en revanche faux : la convergence simple *n'implique pas* la convergence uniforme ; il faut bien noter et comprendre l'ordre distinct des quantificateurs $\forall$ et $\exists$ dans les deux conditions (1) et (2) !

Exemple 55
Dans le cas des fonctions $f_n(x) = x^n$ définies sur l'intervalle $[0, 1]$ de l'exemple ci-dessus, la convergence des f_n n'est pas uniforme vers $f(x) = \mathrm{E}(x)$, comme on va le montrer plus bas par l'absurde.

Mais il est instructif de le comprendre autrement. Montrons que $\forall\, a$ fixé, avec $0 \leqslant a < 1$, la suite $f_n(x) = x^n$ converge uniformément vers 0 sur l'intervalle $[0, a]$. En effet,

$$\forall\, n \geqslant N, \forall\, x \in [0, a], |x^n - 0| = x^n \leqslant a^n \leqslant a^N, \tag{3}$$

⌟

par des inégalités triviales. Pour assurer que $|x^n - 0| \leqslant \epsilon$, il suffit de choisir N tel que $a^N \leqslant \epsilon$ ou encore, de façon équivalente, $N \geqslant \ln\epsilon/\ln a$, par exemple $N = \mathrm{E}(\ln\epsilon/\ln a) + 1$. Ce choix de N assure bien la convergence uniforme des f_n vers 0 dans l'intervalle $[0, a]$. Mais noter que ce N augmente sans limite quand a s'approche de 1. En conséquence, la convergence uniforme ne peut être maintenue sur tout l'intervalle $[0, 1[$, même « ouvert » à droite.

Exemple 56
Soit la suite de fonctions

$$f_n(x) = \begin{cases} (n-1)\cdot x & \text{si } x \in [0, 1/n], \\ 1 - x & \text{si } x \in [1/n, 1]. \end{cases} \tag{4}$$

Pour $x = 0$, $f_n(0) = 0$ donc $\lim f_n(0) = 0$. Pour tout $x \neq 0$, il existe un n assez grand à partir duquel $f_n(x) = 1 - x$. La suite converge donc simplement vers la fonction discontinue telle que $f(0) = 0$ et $f(x) = 1 - x$ si $x \neq 0$, mais la convergence *n'est pas* uniforme. ⌟

Exemple 57
Par contre, les fonctions $x \mapsto f_n(x) = n \sin(x/n)$ convergent uniformément vers $f(x) = x$ sur l'intervalle $[0, 1]$.

En effet,

$$\left|f_n(x) - x\right| = x - n \sin\!\left(\frac{x}{n}\right) = n\cdot\!\left(\frac{x}{n} - \sin\!\left(\frac{x}{n}\right)\right) \leqslant \frac{x^3}{6\, n^2} \leqslant \frac{1}{6\, n^2} \tag{5}$$

⌟

quel que soit $x \in [0, 1]$. On peut démontrer l'inégalité $\phi(\alpha) \coloneqq \sin\alpha - \alpha + \alpha^3/6 \geqslant 0$ pour tout $\alpha \in [0, 1/n]$ utilisée dans cet argument en étudiant les fonctions dérivées $\phi'(\alpha)$ et $\phi''(\alpha)$. On peut également utiliser la formule de Taylor-Lagrange : $\sin\alpha = \alpha - \alpha^3/6 + c^5/5!$, où $c \in]0, 1/n[$, donc $\sin\alpha - \alpha + \alpha^3/6 \geqslant 0$.

Exercice
Montrer que les fonctions $f_n(x) \coloneqq \sum_{p=1}^{n} x^p/p!$ et $g_n(x) \coloneqq (1 + x/n)^n$ convergent, pour tout x réel, vers la fonction exponentielle e^x. La convergence est-elle uniforme ? ⌟

Théorème 55 (critère de Cauchy uniforme)
Soient A un ensemble, F un espace vectoriel normé *complet* et $(f_n\colon A \to F)_{n\in\mathbb{N}}$ une suite de fonctions définies sur A.

Il existe une application $f\colon A \to F$ vers laquelle $(f_n)_{n\in\mathbb{N}}$ converge uniformément sur A si et seulement si

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}, \forall\, x \in A, \forall\, (n, p) \in \mathbb{N}^2, (n \geqslant N, p \geqslant N) \Longrightarrow \|f_n(x) - f_p(x)\| \leqslant \epsilon.$$ ⌟

Une des motivations principales de la définition de convergence uniforme est contenue dans le théorème suivant :

Théorème 56 (continuité de la limite uniforme)
Soient E et F des espaces vectoriels normés, A une partie de E, $x_0 \in A$, et $(f_n\colon A \to F)$ une suite de fonctions définies sur A convergeant uniformément vers une application $f\colon A \to F$.

Si les f_n sont continues en x_0 (resp. sur A), f est continue en x_0 (resp. sur A). ⌟

Preuve
Montrons la continuité de f en tout $x_0 \in A$. Par la convergence uniforme,

$$\forall\, \epsilon > 0, \exists\, N, \forall\, n \geqslant N, x \in A \Longrightarrow \|f_n(x) - f(x)\| \leqslant \frac{\epsilon}{3}.$$

Ceci est vrai pour tout $x \in A$, y compris pour x_0, donc

$$\forall\, n \geqslant N, \|f_n(x_0) - f(x_0)\| \leqslant \frac{\epsilon}{3}.$$

Choisissons un tel $n \geqslant N$. La continuité de f_n dit que

$$\forall\, \epsilon > 0, \exists\, \eta, \forall\, x \in A, \|x_0 - x\| \leqslant \eta \Longrightarrow \|f_n(x) - f_n(x_0)\| \leqslant \frac{\epsilon}{3}$$

et l'inégalité triangulaire dit alors que

$$\begin{aligned}\|f(x) - f(x_0)\| &\leqslant \|f(x) - f_n(x)\| + \|f_n(x) - f_n(x_0)\| + \|f_n(x_0) - f(x_0)\|\\ &\leqslant \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon,\end{aligned}$$

ce qu'il fallait démontrer. ⌟

Théorème 57 (dérivation)
Les fonctions f_n sont ici définies et dérivables sur un intervalle I. F est en outre supposé *complet*.

Si la suite $(f_n(x_0))_{\in\mathbb{N}}$ converge en un point $x_0 \in I$ et si la suite de fonctions $(f'_n)_{\in\mathbb{N}}$ converge *uniformément* sur I, alors

- la suite de fonctions $(f_n)_{\in\mathbb{N}}$ converge simplement vers une fonction f dérivable sur I ;
- cette convergence est uniforme sur tout intervalle fermé $[a, b] \subset I$;
- $$\lim_{n\to\infty} f'_n = \left(\lim_{n\to\infty} f_n\right)' \; (= f').$$ ⌟

Remarque
La condition de convergence uniforme est sur (f'_n), pas (f_n). ⌟

Théorème 58 (intégration)
Les fonctions f_n sont ici définies et continues sur un intervalle $[a, b]$. F est en outre supposé *complet*.

Soit $x_0 \in [a, b]$. Si les (f_n) convergent uniformément vers f sur $[a, b]$, alors

- f admet une primitive sur $[a, b]$;
- $$\forall\, x \in [a, b], \quad \lim_{n\to\infty}\left(\int_{x_0}^{x} f_n(t)\,\mathrm{d}t\right) = \int_{x_0}^{x}\left(\lim_{n\to\infty} f_n(t)\right)\mathrm{d}t\,;$$
- cette convergence est uniforme. ⌟

Preuve
Les (f_n) étant continues et convergeant uniformément vers f, f est continue, donc admet une primitive ; les (f_n) aussi d'après l'hypothèse de continuité.

Soit $g(x) = \int_{x_0}^{x} f(t)\,\mathrm{d}t$ et $g_n(x) = \int_{x_0}^{x} f_n(t)\,\mathrm{d}t$ ces primitives. Comme $\forall n \in \mathbb{N}$, $g_n(x_0) = 0$, la suite $(g_n(x_0))_{n\in\mathbb{N}}$ converge (vers zéro !). On peut donc appliquer le théorème sur la dérivation en remplaçant f'_n par f_n et f_n par g_n. ⌟

Remarque
Si les f_n sont seulement intégrables au sens de Riemann sur $[a, b]$ (pas nécessairement continues), f est intégrable au sens de Riemann et

$$\lim_{n\to\infty}\left(\int_a^b f_n(t)\,\mathrm{d}t\right) = \int_a^b\left(\lim_{n\to\infty} f_n(t)\right)\mathrm{d}t.$$ ⌟

VI. . Séries de fonctions

Les considérations concernant les suites de fonctions s'appliquent aux séries de fonctions $\sum f_n(x)$. On définit pour elles les notions de convergence simple et de convergence uniforme de la même façon qu'on l'a fait pour les suites de fonctions, en les appliquant aux sommes partielles $S_n(x) = \sum_{k=0}^{n} f_k(x)$ et en remplaçant $(f_n)_{n\in\mathbb{N}}$ par $\sum f_n$, f par S et $\lim_{n\to\infty}$ par $\sum^{\infty}$. On a donc les théorèmes suivants :

Théorème 59
Soit une série de fonctions $\sum f_n$ convergeant uniformément vers sa somme S sur A.

Si les f_n sont continues sur A, S l'est aussi. ⌟

Théorème 60 (dérivation)
Les fonctions f_n sont ici définies et dérivables sur un intervalle I. F est en outre supposé *complet*.

Si la série $\sum f_n(x_0)$ converge en un point $x_0 \in I$ et si la série de fonctions $\sum f_n'$ converge *uniformément* sur I, alors

- la série de fonctions $\sum f_n$ converge simplement vers une fonction f dérivable sur I ;
- cette convergence est uniforme sur tout intervalle fermé $[a, b] \subset I$;
- $$\sum_{n=0}^{\infty} f_n' = \left(\sum_{n=0}^{\infty} f_n\right)' \quad (= S').$$

⌟

Théorème 61 (intégration)
Les fonctions f_n sont ici définies et continues sur un intervalle $[a, b]$. F est en outre supposé *complet*.

Soit $x_0 \in [a, b]$. Si $\sum f_n$ converge uniformément vers S sur $[a, b]$, alors

- S admet une primitive sur $[a, b]$;
- $$\forall \ \in [a, b], \quad \sum_{n=0}^{\infty}\left(\int_{x_0}^{x} f_n(t)\,\mathrm{d}t\right) = \int_{x_0}^{x}\left(\sum_{n=0}^{\infty} f_n(t)\right)\mathrm{d}t\ ;$$
- cette convergence est uniforme.

⌟

Remarque
Si les f_n sont seulement intégrables au sens de Riemann sur $[a, b]$ (pas nécessairement continues), S est intégrable au sens de Riemann et
$$\sum_{n=0}^{\infty}\left(\int_{a}^{b} f_n(t)\,\mathrm{d}t\right) = \int_{a}^{b}\left(\sum_{n=0}^{\infty} f_n(t)\right)\mathrm{d}t.$$
⌟

Définition 33 (convergence normale)
Soit $\sum f_n$ une série de fonctions définies sur A. On dit que $\sum f_n$ est *normalement convergente* sur A si il existe une suite (u_n) de réels positifs tels que $\forall\, x \in A$, $\|f_n(x)\| \leqslant u_n$ et $\sum u_n$ soit convergente. ⌟

$\sum f_n$ est donc normalement convergente sur A si $\sum \sup_{x\in A} \|f_n(x)\|$ est convergente.

Le théorème suivant permet souvent de démontrer la convergence uniforme d'une série de fonctions.

Théorème 62 (critère de Weierstrass)
Soit F un espace vectoriel normé complet.

Si $\sum f_n$ est normalement convergente sur A, $\sum f_n$ est uniformément convergente sur A. ⌟

La réciproque est fausse.

Preuve
Il suffit de montrer que $S_n = \sum_{k=0}^{n} f_k$ est uniformément de Cauchy. $\sum u_n$ est convergente, donc de Cauchy, c.-à-d. que
$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}, \forall\, (n, p) \in \mathbb{N}^2, (p > n \geqslant N) \Longrightarrow \Big|\sum_{k=n+1}^{p} u_k\Big| = \sum_{k=n+1}^{p} u_k \leqslant \epsilon.$$

Or, $\forall\, x \in A$,
$$\|S_n(x) - S_p(x)\| = \Big\|\sum_{k=n+1}^{p} f_k(x)\Big\| \leqslant \sum_{k=n+1}^{p} \|f_k(x)\| \leqslant \sum_{k=n+1}^{p} u_k \leqslant \epsilon.$$
⌟

VI. . Séries entières

Une série de fonctions de la forme $(z \mapsto \sum a_n\, z^n)_{n\in\mathbb{N}}$, où les a_n et z sont réels ou complexes, est appelée une *série entière*. Dans cette expression, on posera $0^0 := 1$.

Définition 34
On appelle *domaine de convergence* l'ensemble $\{z \in \mathbb{C} \text{ (ou } \mathbb{R}) \,\big|\, \sum a_n\, z^n \text{ converge}\}$.

On appelle *rayon de convergence* l'élément R de $\mathbb{R} \cup \{+\infty\}$ défini par $R = \sup\{r > 0 \,\big|\, \sum(|a_n|\, r^n) \text{ converge}\}$. ⌟

Attention, la définition de R ne dit pas que $\sum(|a_n|\, R^n)$ converge.

Théorème 63
Soit $\sum a_n\, z^n$ une série entière de rayon de convergence R :

- $\forall\, z, |z| < R \Longrightarrow \sum a_n\, z^n$ converge absolument ;
- $\forall\, z, |z| > R \Longrightarrow \sum a_n\, z^n$ diverge.

Le disque *ouvert* de centre 0 et de rayon R ($\left\{z \in \mathbb{C} \text{ ou } \mathbb{R} \,\middle|\, |z| < R\right\}$) porte le nom de *disque de convergence*.

$\sum a_n\, z^n$ converge *normalement* sur tout disque *fermé* de rayon $r < R$ et de centre 0, c.-à-d. sur l'ensemble $\left\{z \in \mathbb{C} \text{ ou } \mathbb{R} \,\middle|\, |z| \leqslant r\right\}$. ⌟

Si les a_n et z sont réels, le disque de convergence se réduit à l'intervalle ouvert $]-R, R[$.

Théorème 64
Si $\lim_{n\to\infty}(|a_n|^{1/n}) = \ell \in \mathbb{R}^+ \cup \{+\infty\}$, alors $R = 1/\ell$ (règle de Cauchy).

De même, si $\lim_{n\to\infty} |a_{n+1}/a_n| = \ell \in \mathbb{R}^+ \cup \{+\infty\}$ (règle de d'Alembert). ⌟

Exemple 58
Quelques exemples de séries entières réelles :

1. On a vu ci-dessus que $\sum x^n/n!$ est convergente pour tout réel x, puisque $\ell = 0$;
2. À l'inverse, pour $\sum x^n\, n!$, on a $\ell = \infty$, donc la série ne converge qu'en 0 ;
3. $\sum(x/r)^n$, $r \neq 0$. On a $\ell = 1/r$, donc un rayon de convergence $R = r$. Pour $|x| = r$, la série ne converge en aucun point (puisque le terme d'ordre n ne s'annule pas) ;
4. Pour $\sum x^n/n$, $R = 1$. Dans ce cas, on connaît déjà la nature de la série quand $|x| = R$. En $x = 1$, la série diverge ; en $x = -1$, elle converge (cf. séries alternées) vers $\ln 2$;
5. Pour $\sum x^n/n^2$, $R = 1$. Pour $x = \pm 1$, il y a convergence absolue ;
6. Pour $\sum n\, x^n$, $R = 1$. La série diverge pour $x = \pm 1$. ⌟

Il ressort de cette série d'exemples que le comportement sur le bord de l'intervalle de convergence est une question délicate, à étudier avec soin.

Théorème 65
Une série entière est continue en tout point de son disque de convergence. ⌟

Preuve
Les fonctions $z \mapsto z^n$ sont continues sur C, donc les sommes partielles $z \mapsto \sum_{k=0}^{n} a_k\, z^k$ aussi. Pour tout z tel que $|z| < R$, il existe r tel que $|z| < r < R$. Or $\sum a_n\, z^n$ est normalement convergente, donc uniformément convergente, sur tout disque fermé de rayon $r < R$. $\sum_{k=0}^{\infty} a_k\, z^k$ est donc continue en z. ⌟

Posons $x = z$ et supposons x et les a_n réels [(*1)]. En intégrant terme à terme la série $\sum a_n\, t^n$ de 0 à x, ce qui ne pose pas de problème si $|x| < R$ puisque $\sum a_n\, t^n$ est normalement (donc uniformément) convergente sur $[-r, r] \ni x$ si $r < R$, on obtient le résultat suivant :

Théorème 66
Soient $\sum a_n\, x^n$ une série entière de rayon R et $S(x) = \sum_{n=0}^{\infty} a_n\, x^n$.

La série entière $\sum((a_n/(n+1))\, x^{n+1})$ converge pour $|x| < R$ vers l'intégrale $\int_0^x S(t)\, \mathrm{d}t$; son rayon de convergence est R. ⌟

De même, pour la dérivation :

Théorème 67
La série entière $\sum_{n\geqslant 1} n\, a_n\, x^{n-1}$ converge pour $|x| < R$ vers $S'(x)$; son rayon de convergence est R. ⌟

Exemple 59
En intégrant terme à terme entre 0 et x la série géométrique $\sum_{n=0}^{\infty}(-1)^n\, x^n$, qui converge vers $1/(1+x)$ pour $x \in\,]-1, 1[$, on obtient

$$\ln(1+x) = \sum_{n=1}^{\infty}(-1)^{n-1}\, \frac{x^n}{n}, \tag{6}$$

1. Les résultats sur l'intégration et la dérivation de séries entières réelles sont applicables aux séries complexes, mais les concepts d'intégrale et de dérivation complexes sont hors programme.

qui converge pour $|x| < 1$. Pour $x = 1$, la série converge encore (voir TD). On a donc $1 - 1/2 + 1/3 - \cdots = \ln 2$. ⌟

VI. .1. Séries de Taylor

On peut donc dériver terme à terme une série entière à l'intérieur de son disque de convergence (pour x réel, $|x| < R$), puis dériver à nouveau, etc. On obtient

$$\forall\, x \in\,]-R, R[,\ S^{(n)}(x) = \sum_{k=n}^{\infty} k!\, a_k\, x^k,$$

soit $a_n = S^{(n)}(0)/n!$. Pour tout $x \in\,]-R, R[$, une série entière s'identifie donc à son développement en *série de Taylor(- Mac-Laurin)*,

$$S(x) = \sum_{n=0}^{\infty} \frac{S^{(n)}(0)}{n!}\, x^n. \tag{7}$$

Remarquer qu'il s'agit d'un développement *illimité*.

Soit f une fonction de la variable réelle x, indéfiniment dérivable au voisinage de 0. On peut écrire *formellement* son développement en série de Taylor,

$$f(0) + x\, f'(0) + \frac{x^2}{2}\, f''(0) + \cdots + \frac{x^n}{n!}\, f^{(n)}(0) + \cdots . \tag{8}$$

Deux questions se posent alors :

1. Cette série entière converge-t-elle sur un certain domaine ?
2. Si oui, sa somme vaut-elle $f(x)$ sur tout ou partie de ce domaine ?

Nous ne poursuivrons pas plus loin la discussion de ce problème important, qui prend tout son sens avec l'étude du comportement de la série et des singularités de la fonction pour des valeurs complexes de la variable x.

Insistons seulement sur le caractère nécessaire de l'existence (et donc de la continuité) de toutes les dérivées de f en 0 (f doit être de classe C^∞ »). Par exemple, les fonctions $x \mapsto \sqrt{x}$ et $x \mapsto |x|$ ne sont pas développables en série de Taylor au voisinage de 0.

Inversement, le développement de Taylor de f peut avoir un rayon de convergence non nul mais ne pas représenter la fonction f. Considérons la fonction définie par $f(0) = 0$, $f(x) = \mathrm{e}^{-1/x^2}$ pour $x \neq 0$. Comme e^{-1/x^2} tend vers zéro plus vite que toute puissance de x, c.-à-d. $\lim \mathrm{e}^{-1/x^2}/x^p = 0$ pour tout p, la fonction f est continue et indéfiniment dérivable en 0 et toutes ses dérivées y sont nulles ! Il est clair que son développement de Taylor ($\sum 0 \times x^n$) est convergent vers 0 pour toute valeur de x, mais que ce développement ne converge nulle part vers $f(x)$ sauf en $x = 0$. Cela est dû au comportement très singulier de la fonction en 0. Pour le voir, effectuons une petite incursion dans le plan complexe : pour $x = \mathrm{i}\, t$, la fonction vaut e^{1/t^2}, qui explose au voisinage de 0 (on parle de singularité essentielle) !

VI. .2. Séries entières usuelles

- $\forall\, z \in \mathbb{C}, \quad \exp z = \sum_0^\infty \frac{z^n}{n!}$.
- $\forall\, z \in \mathbb{C}, \quad \cos z = \sum_0^\infty (-1)^n \frac{z^{2n}}{(2n)!}$.
- $\forall\, z \in \mathbb{C}, \quad \sin z = \sum_0^\infty (-1)^n \frac{z^{2n+1}}{(2n+1)!}$.
- $\forall\, z \in \mathbb{C}, \quad \mathrm{ch}\, z = \sum_0^\infty \frac{z^{2n}}{(2n)!}$.
- $\forall\, z \in \mathbb{C}, \quad \mathrm{sh}\, z = \sum_0^\infty \frac{z^{2n+1}}{(2n+1)!}$.
- $\forall\, z \in \mathbb{C}, \quad |z| < 1 \Rightarrow \frac{1}{1+z} = \sum_0^\infty (-1)^n z^n$.
- $\forall\, x \in \mathbb{R}, \quad |x| < 1 \Rightarrow \ln(1+z) = \sum_1^\infty (-1)^{n+1} \frac{x^n}{n}$; ce développement est également valable en $z = 1$.
- $\forall\, x \in \mathbb{R}, \quad |x| < 1 \Rightarrow \arctan x = \sum_0^\infty (-1)^n \frac{x^{2n+1}}{2n+1}$; ce développement est également valable en $z = 1$.
- $\forall\, x \in \mathbb{R}, \quad |x| < 1 \Rightarrow \mathrm{argth}\, x = \sum_0^\infty \frac{x^{2n+1}}{2n+1}$.

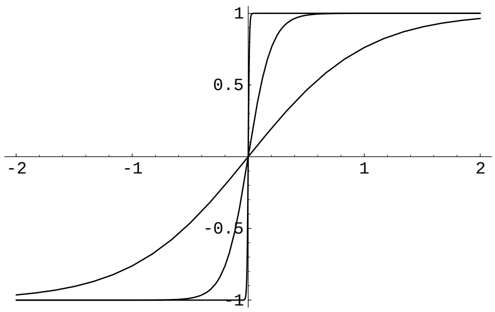


Figure VI.1 – La fonction $\operatorname{th}(n\,x)$ pour $n = 1, 5$ et 100.

- $\forall\, x \in \mathbb{R}, \forall\, \alpha \in \mathbb{R} \setminus \{-1\}, \quad |x| < 1 \Rightarrow (1+x)^\alpha = \sum_0^\infty C_\alpha^n\, x^n,$
  où $C_\alpha^n = \alpha\cdot\big(\alpha-1\big)\cdots\big(\alpha-(n-2)\big)\cdot\big(\alpha-(n-1)\big)/n!$ si $n \neq 0$ et $C_\alpha^0 = 1$.

VI. . Interversion des opérations lim, $\sum$, $\int$ et $\frac{\mathrm{d}}{\mathrm{d}x}$

Les opérations de limite, de sommation infinie, d'intégration et de dérivation ne commutent que sous certaines conditions. Nous n'allons pas ici discuter tous les cas possibles, mais plutôt nous restreindre à quelques mises en garde (sous forme de contre-exemples) et énoncer quelques théorèmes utiles pour l'interversion de deux opérations.

Les cas

$$\lim_{n\to\infty} \frac{\mathrm{d}}{\mathrm{d}x} \stackrel{?}{=} \frac{\mathrm{d}}{\mathrm{d}x} \lim_{n\to\infty},$$

$$\sum^\infty \frac{\mathrm{d}}{\mathrm{d}x} \stackrel{?}{=} \frac{\mathrm{d}}{\mathrm{d}x} \sum^\infty,$$

$$\lim_{n\to\infty} \int \stackrel{?}{=} \int \lim_{n\to\infty},$$

et

$$\sum^\infty \int \stackrel{?}{=} \int \sum^\infty,$$

les deux derniers pour des intégrales propres, ont déjà été étudiés.

VI. .1. Double limite

Exemple 60

Considérons la suite de fonctions

$$f_n(x) = \operatorname{th}(n\,x) = \frac{1 - \mathrm{e}^{-2\,n\,x}}{1 + \mathrm{e}^{-2\,n\,x}}$$

quand $x \to 0$ et $n \to \infty$. Nous avons $\lim_{x\to 0} f_n(x) = 0$, donc

$$\lim_{n\to\infty}\Big(\lim_{x\to 0} f_n(x)\Big) = 0,$$

tandis que $\lim_{n\to\infty} f_n(x) = \operatorname{sgn} x$, la « fonction saut » (voir figure VI.1) :

$$\lim_{x\to 0}\Big(\lim_{n\to\infty} f_n(x)\Big)$$

n'est donc pas définie. ⌟

Cette situation se rencontre aussi en physique. C'est ainsi qu'en mécanique statistique, où on s'intéresse aux propriétés de systèmes de grand nombre de constituants, il faut faire attention à l'ordre des limites. Considérons par exemple la question de l'*aimantation spontanée* d'un ferromagnétique. Un système de N moments magnétiques m_i est soumis à un champ magnétique h, qui tend à aligner ces moments et conduit à l'apparition d'une aimantation moyenne $M(h, N) = \sum m_i/N$. L'aimantation spontanée est la limite de cette aimantation quand

$h \to 0$. On démontre qu'à N fini, $\lim_{h\to 0} M = 0$, tandis que certains systèmes sont tels que $\lim_{N\to\infty} M(h, N)$ existe et demeure finie quand $h \to 0$.

En fait, bien sûr, un nombre d'atomes, de particules, etc., n'est jamais strictement infini, mais seulement colossalement grand, de l'ordre du nombre d'Avogadro $\mathcal{N} \approx 6 \times 10^{23}$. Dans l'exemple précédent, il faut donc comprendre que N est si grand que les quantités physiques envisagées ne diffèrent de celles pour N infini que dans des intervalles des variables (ici de champ) extrêmement faibles. La fonction de la figure VI.1 en donne un exemple : l'intervalle où la fonction diffère de la fonction saut est de l'ordre de $|\Delta x| \approx 1/N$.

Théorème 68 (double limite)
Soient E un espace vectoriel normé, A une partie de E, $a \in A$, F un espace vectoriel normé *complet*, $(f_n : A \to F)_{n\in\mathbb{N}}$ une suite de fonctions définies sur A, et $f : A \to F$ une application.

Si les $(f_n)_{n\in\mathbb{N}}$ convergent uniformément vers f sur A et si $f_n(x)$ converge vers $\phi_n \in F$ quand $x \to a$, alors

- $\lim_{n\to\infty} \phi_n$ existe (dans F) ;
- $\lim_{x\to a} f(x)$ existe (dans F) ; et
-

$$\lim_{n\to\infty}\Big(\lim_{x\to a} f_n(x)\Big) = \lim_{x\to a}\Big(\lim_{n\to\infty} f_n(x)\Big).$$

⌟

Ce théorème se transpose sans difficulté au cas d'une série de fonction.

La question de l'interversion des limites se pose également pour une fonction de plusieurs variables. Considérons une fonction de deux variables x et λ (λ peut être considéré comme un paramètre dont dépend la fonction de x), et examinons le comportement quand $x \to x_0$ et $\lambda \to \lambda_0$ *à la fois*. Plus précisément, supposons que $f(x, \lambda) \to g(x)$ quand $\lambda \to \lambda_0$, tandis que $f(x, \lambda) \to h(\lambda)$ quand $x \to x_0$. Peut-on affirmer que les fonctions $g(x)$ et $h(\lambda)$ convergent quand $x \to x_0$ ou $\lambda \to \lambda_0$? Et si oui, a-t-on

$$\lim_{x\to x_0}\Big(\lim_{\lambda\to\lambda_0} f(x,\lambda)\Big) \stackrel{?}{=} \lim_{\lambda\to\lambda_0}\Big(\lim_{x\to x_0} f(x,\lambda)\Big), \tag{9}$$

c'est-à-dire, l'échange des limites est-il autorisé ?

Il est aisé de construire des exemples mathématiques où cela n'est pas le cas, soit que l'une des limites de $g(x)$ ou de $h(\lambda)$ n'existe pas, soit que ces limites existent mais sont différentes.

Exemple 61
La fonction $f(x, \lambda) = \lambda/(x^2 + \lambda^2)$ est telle que $g(x) = \lim_{\lambda\to 0} f(x, \lambda) = 0$ pour tout x, tandis que $h(\lambda) = \lim_{x\to 0} f(x, \lambda) = 1/\lambda$. La limite de h quand $\lambda \to 0$ n'existe pas. ⌟

VI. .2. Dérivation et intégration

Dans le même ordre d'idées, peut-on dériver sous le signe d'intégration ?

On démontre le théorème suivant :

Théorème 69 (Leibniz)
Soit $f : (x, \lambda) \mapsto f(x, \lambda)$ une fonction définie pour $x \in [a, b]$ et λ dans un intervalle réel I. Si f est continue et si $\partial f/\partial\lambda$ existe et est également continue [(*2)], alors

$$\forall\, \lambda \in I, \quad \frac{\partial}{\partial\lambda}\int_{x=a}^{b} f(x,\lambda)\,\mathrm{d}x = \int_{x=a}^{b} \frac{\partial f(x,\lambda)}{\partial\lambda}\,\mathrm{d}x. \tag{10}$$

⌟

Si l'intégrale est impropre, des conditions supplémentaires sont requises pour pouvoir intervertir la dérivation et l'intégration.

VI. .3. Deux intégrations

Rappelons que pour une série dont la convergence est *absolue*, on a le droit de permuter l'ordre des termes. Vu que l'intégrale est définie comme la limite d'une somme, on peut appliquer ce principe au calcul d'une intégrale double. Plus précisément, nous avons le théorème suivant :

Théorème 70 (Fubini)
Soit $f(x, y)$ une fonction définie pour $x \in [a, b]$ et $y \in [c, d]$.

2. f et $\partial f/\partial\lambda$ doivent être continues par rapport au couple (x, λ), pas par rapport à x et λ séparément.

Si

$$\int_{x=a}^{b}\left(\int_{y=c}^{d}\left|f(x,y)\right|\cdot\mathrm{d}y\right)\cdot\mathrm{d}x \tag{11}$$

converge, alors

$$\int_{y=c}^{d}\left(\int_{x=a}^{b}f(x,y)\cdot\mathrm{d}x\right)\cdot\mathrm{d}y = \int_{x=a}^{b}\left(\int_{y=c}^{d}f(x,y)\cdot\mathrm{d}y\right)\cdot\mathrm{d}x. \tag{12}$$

⌟

Dans le même ordre d'idées, dans une intégration sur un domaine bidimensionnel, on a le droit de procéder à un changement de variables pour « balayer » ce domaine autrement, sous la même condition. Nous avons déjà utilisé ce résultat (sans le dire !) en calculant $\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x = \sqrt{\pi}$ au § V.c.3.

Chapitre VII

S´ F

VII. . Préambule sur les fonctions périodiques

Définition 35
Soient une application

$$\begin{aligned} f\colon \mathbb{R} &\longrightarrow \mathbb{R} \text{ ou } \mathbb{C} \\ x &\longmapsto f(x) \end{aligned}$$

et un réel T.

f est dite *T-périodique* si

$$\forall\, x \in \mathbb{R},\ f(x+T) = f(x). \tag{1}$$

On dit alors que T est *une période* (parmi d'autres éventuellement).

f est dite *périodique* (tout court) si il existe un $T \neq 0$ tel que f soit T-périodique ; elle est dite *apériodique* sinon. ⌟

Toute fonction définie sur $\mathbb{R}$ tout entier est bien sûr 0-périodique. L'intérêt de considérer 0 comme une période est que l'ensemble des périodes constitue un groupe pour l'addition : la somme ou la différence de deux périodes est une période, de même que tout multiple entier d'une période.

Définition 36
Soient f une fonction périodique et $\tau \coloneqq \inf\{T > 0 \mid f \text{ est } T\text{-périodique}\}$. Si $\tau > 0$ (strictement), τ porte le nom de *plus petite période* de f. ⌟

Remarques

- La plus petite période sera souvent appelée *la* période. Par exemple, la (plus petite) période de $x \mapsto \sin x$ ou de $x \mapsto \cos x$ est $2\,\pi$; celle de $x \mapsto \tan x$ est π.
- La plus petite période n'est pas définie pour une fonction constante. En effet, pour une telle fonction, tout réel est une période. Or $\inf \mathbb{R}^{+*} = 0$.
- Toute fonction périodique continue par morceaux et non constante admet une plus petite période.
- En revanche, on peut trouver des fonctions non continues par morceaux périodiques ne possédant pas de plus petite période. Par exemple,

$$\begin{aligned} f\colon \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) = \begin{cases} 1 & \text{si } x \in \mathbb{Q}, \\ 0 & \text{si } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases} \end{aligned}$$

  En effet, pour tout T rationnel, f est T-périodique (si $x \in \mathbb{Q}$ et $y \in \mathbb{Q}$, $x + y \in \mathbb{Q}$; et si $x \in \mathbb{Q}$ et $y \in \mathbb{R} \setminus \mathbb{Q}$, $x + y \in \mathbb{R} \setminus \mathbb{Q}$). Or $\inf \mathbb{Q}^{+*} = 0$. ⌟

Une fonction de période $T > 0$ est complètement définie par sa restriction à tout intervalle $[a, a+T[$ (ou $]a, a+T]$) de longueur T ; on dira qu'un tel intervalle est ***une** période*, sans risque de confusion avec une période au sens de la définition 35. Si la fonction est dérivable, sa dérivée est aussi périodique de même période.

Théorème 71
Si f est une fonction périodique et intégrable, son intégrale sur une période $T > 0$ ne dépend pas de l'intervalle choisi :

$$\forall\, (a, b) \in \mathbb{R}^2, \quad \int_a^{a+T} f(x)\,\mathrm{d}x = \int_b^{b+T} f(x)\,\mathrm{d}x. \tag{2}$$

⌟

Preuve
En effet,

$$\int_a^{a+T} f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_b^{b+T} f(x)\,\mathrm{d}x + \int_{b+T}^{a+T} f(x)\,\mathrm{d}x$$

et, dans le dernier terme, le changement de variable $x = x' + T$ permet d'écrire

$$\int_{b+T}^{a+T} f(x)\,\mathrm{d}x = \int_b^a f(x')\,\mathrm{d}x' = -\int_a^b f(x)\,\mathrm{d}x.$$

┘

On va utiliser fréquemment cette propriété, en choisissant la période la plus propice, c.-à-d. celle rendant le calcul le plus simple. Par exemple, si la fonction est périodique et impaire, $\int_{-T/2}^{T/2} f(x)\,\mathrm{d}x = 0$, et donc $\forall\, a$, $\int_a^{a+T} f(x)\,\mathrm{d}x = 0$.

Les fonctions périodiques interviennent de façon très courante dans la description des phénomènes physiques ou biologiques. La période n'est pas nécessairement un temps ; elle peut également être de nature spatiale, par exemple la longueur d'onde d'un phénomène ondulatoire ou la maille d'un réseau cristallin, etc. Bien sûr, dans un système physique réel, il n'est pas possible d'effectuer des translations arbitrairement grandes du temps ou d'une coordonnée d'espace : la notion de fonction périodique est une idéalisation de la réalité physique, quand la durée d'observation ou les dimensions de l'échantillon sont beaucoup plus grandes que les périodes observées.

VII. . Séries trigonométriques

Les fonctions trigonométriques $x \mapsto \cos(n\,x)$ et $x \mapsto \sin(n\,x)$ sont évidemment des fonctions périodiques de période $T = 2\,\pi$. C'est le cas également de tout *polynôme trigonométrique* de degré n, c.-à-d. de toute expression de la forme

$$S_n(x) = \frac{1}{2}\,a_0 + \sum_{k=1}^{n} (a_k \cos(k\,x) + b_k \sin(k\,x)) \quad \text{où } (a_n, b_n) \neq (0, 0) \tag{3}$$

(remarquer qu'il s'agit d'une combinaison linéaire de $2\,n + 1$ termes), et, plus généralement, de toute *série trigonométrique*,

$$S(x) = \frac{1}{2}\,a_0 + \sum_{n=1}^{\infty} (a_n \cos(n\,x) + b_n \sin(n\,x)), \tag{4}$$

aux points où elle converge.

Une telle série converge notamment dans les deux cas suivants :

1. Si les séries $\sum a_n$ et $\sum b_n$ sont absolument convergentes, comme $|\cos(n\,x)| \leqslant 1$ et $|\sin(n\,x)| \leqslant 1$, la série de fonctions S_n est normalement convergente sur $\mathbb{R}$. Elle est donc uniformément convergente sur $\mathbb{R}$ et sa limite est une fonction continue car les fonctions $x \mapsto \cos(n\,x)$ et $x \mapsto \sin(n\,x)$ sont continues.
   $\sum a_n$ et $\sum b_n$ sont absolument convergentes en particulier si
   $$a_n = O\!\left(\frac{1}{n^{\alpha}}\right), \qquad b_n = O\!\left(\frac{1}{n^{\beta}}\right), \qquad (\alpha, \beta) > 1. \tag{5}$$
   La somme de la série est éventuellement dérivable terme à terme si les séries $\sum n\,a_n$ et $\sum n\,b_n$ sont elles-mêmes absolument convergentes. On peut alors écrire
   $$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{2}\,a_0 + \sum_{n=1}^{\infty} (a_n \cos(n\,x) + b_n \sin(n\,x))\right) = \sum_{n=1}^{\infty} (n\,b_n \cos(n\,x) - n\,a_n \sin(n\,x)), \tag{6}$$
   avec une série convergente au membre de droite. Par exemple, si $(\alpha, \beta) > 2$ dans (5), la somme de la série (4) est continue et dérivable, et sa dérivée, donnée par (6), est elle-même continue.
2. Quelles que soient les suites (a_n) et (b_n) décroissantes et tendant vers zéro, la série trigonométrique converge pour tout $x \neq 2\,\ell\,\pi$, où $\ell \in \mathbb{Z}$. (Par contre, on ne peut rien dire pour $x = 0 \pmod{2\,\pi}$.)
   Ceci est une conséquence du théorème d'Abel (théor. 28). On peut écrire
   $$\sum_{k=1}^{n} \mathrm{e}^{\mathrm{i}\,k\,x} = \mathrm{e}^{\mathrm{i}\,x} \sum_{k=0}^{n-1} \mathrm{e}^{\mathrm{i}\,k\,x} = \mathrm{e}^{\mathrm{i}\,x}\,\frac{1 - \mathrm{e}^{\mathrm{i}\,n\,x}}{1 - \mathrm{e}^{\mathrm{i}\,x}} = \frac{\mathrm{e}^{\mathrm{i}\,x}\,\mathrm{e}^{\mathrm{i}\,n\,x/2}}{\mathrm{e}^{\mathrm{i}\,x/2}}\,\frac{\mathrm{e}^{-\mathrm{i}\,n\,x/2} - \mathrm{e}^{\mathrm{i}\,n\,x/2}}{\mathrm{e}^{-\mathrm{i}\,x/2} - \mathrm{e}^{\mathrm{i}\,x/2}} = \mathrm{e}^{\mathrm{i}\cdot(n+1)\,x/2}\,\frac{\sin(n\,x/2)}{\sin(x/2)}$$

pourvu que $\sin(x/2) \neq 0$, c.-à-d. $x \neq 2\,\ell\,\pi$. On a donc

$$\sum_{k=1}^{n} \cos(k\,x) = \mathrm{Re}\left(\sum_{k=1}^{n} \mathrm{e}^{\mathrm{i}\,k\,x}\right) = \cos\left(\frac{n+1}{2}\,x\right)\frac{\sin(n\,x/2)}{\sin(x/2)}$$

et

$$\sum_{k=1}^{n} \sin(k\,x) = \mathrm{Im}\left(\sum_{k=1}^{n} \mathrm{e}^{\mathrm{i}\,k\,x}\right) = \sin\left(\frac{n+1}{2}\,x\right)\frac{\sin(n\,x/2)}{\sin(x/2)},$$

soit

$$\left|\sum_{k=1}^{n} \cos(k\,x)\right| \leqslant \frac{1}{\left|\sin(x/2)\right|}, \qquad \left|\sum_{k=1}^{n} \sin(k\,x)\right| \leqslant \frac{1}{\left|\sin(x/2)\right|},$$

et l'hypothèse du théorème d'Abel est satisfaite.

Exemple : si $a_n = 1/n$, $\sum a_n \cos(n\,x)$ et $\sum a_n \sin(n\,x)$ convergent si $x \neq 2\,\ell\,\pi$; en particulier, pour $x = \pi$, $\sum_{n\geqslant 1}(-1)^n/n = -\ln 2$. Mais pour $x = 0$, $\sum a_n \cos 0$ diverge, $\sum a_n \sin 0 = 0$ converge.

On peut alors légitimement se demander à quelles conditions une fonction périodique f est développable en une série trigonométrique $\widetilde{f}$ convergeant vers f sur tout ou partie de $\mathbb{R}$.

VII. . Développement en série de Fourier

VII. .1. Fonction $2\,\pi$-périodique

Il est en fait plus judicieux de se poser la question suivante : quelle est la série trigonométrique la plus « proche » de f sur un intervalle de longueur $2\,\pi$? Pour répondre à cette question, il faut définir une distance entre f et l'espace (ou un sous-espace) des séries trigonométriques et chercher parmi celles-ci celle qui minimise cette distance.

Plusieurs choix de distance sont possibles, notamment les suivants :

- $d_\infty(f, g) = \sup_{x\in[0,\,2\,\pi[} |f(x) - g(x)|$;
- $d_1(f, g) = \dfrac{1}{2\,\pi}\displaystyle\int_0^{2\,\pi} |f(x) - g(x)|\,\mathrm{d}x$;
- $d_2(f, g) = \sqrt{\dfrac{1}{2\,\pi}\displaystyle\int_0^{2\,\pi} |f(x) - g(x)|^2\,\mathrm{d}x}\,.$

La distance d_∞ semble la plus naturelle. Remarquons en particulier que si $d_\infty(f, g) = 0$, alors $f = g$. Il est cependant difficile de calculer $d_\infty(f, g)$ et de trouver la série g qui la minimise pour une fonction f donnée.

La meilleure solution est en fait d'utiliser la distance euclidienne d_2, car elle est associée à un produit hermitien (la généralisation du produit scalaire au cas où le corps est $\mathbb{C}$; cf. annexe VII.1) sur l'espace des fonctions continues sur $[0, 2\,\pi[$,

$$\langle f, g\rangle = \frac{1}{2\,\pi}\int_0^{2\,\pi} \overline{f(x)}\,g(x)\,\mathrm{d}x.$$

Ce produit scalaire permet de définir une norme,

$$\|f\|_2 = \sqrt{\langle f, f\rangle} = \sqrt{\frac{1}{2\,\pi}\int_0^{2\,\pi} |f(x)|^2\,\mathrm{d}x}$$

et une distance, $d_2(f, g) = \|f - g\|_2$.

Attention, contrairement à d_∞, l'application d_2 (de même pour d_1) est une distance sur l'espace des fonctions continues, mais pas sur celui des fonctions continues par morceaux. $\|\cdot\|_2$ est seulement une semi-norme sur ce dernier. En effet, si f et g sont deux fonctions égales *presque partout* (par exemple, égales sur un intervalle sauf en un nombre dénombrable de points), $\|f - g\|_2 = 0$ bien que $f \neq g$. C'est en particulier le cas pour deux fonctions continues par morceaux ne différant qu'aux points de discontinuité.

On peut généraliser ceci à un espace plus grand que les fonctions continues mais plus petit que les fonctions continues par morceaux :

Définition 37
On note D l'espace des fonctions $\mathrm{C}^0_{\mathrm{m}}$ et $2\,\pi$-périodiques de $\mathbb{R}$ dans $\mathbb{C}$ telles que

$$\forall\, x \in \mathbb{R},\ f(x) = \frac{f(x^-) + f(x^+)}{2}.$$

Théorème 72
L'application [(*1)]

$$\begin{aligned}\langle\cdot,\cdot\rangle\colon\quad & D^2 \longrightarrow \mathbb{C}\\ (f,g) & \longmapsto \langle f,g\rangle = \frac{1}{2\,\pi}\int_0^{2\,\pi}\overline{f(x)}\,g(x)\,\mathrm{d}x\end{aligned}$$

est un produit hermitien sur D.

$\|f\|_2 := \sqrt{\langle f, f\rangle}$ est une norme sur D. ⌟

Les fonctions $e_n\colon x \mapsto \mathrm{e}^{\mathrm{i}\,n\,x}$ ($n \in \mathbb{Z}$) constituent une famille orthonormée de D :

$$\forall\,(m,n)\in\mathbb{Z}^2,\ \langle e_m, e_n\rangle = \frac{1}{2\,\pi}\int_0^{2\,\pi}\mathrm{e}^{\mathrm{i}\cdot(n-m)\,x}\,\mathrm{d}x = \delta_{m,\,n},$$

où $\delta_{m,\,n}$ est le symbole de Kronecker, qui vaut 0 si $m \neq n$ et 1 si $m = n$. En effet, $\int_0^{2\,\pi}\mathrm{d}x = 2\,\pi$ (cas $m = n$) et

$$\int_0^{2\,\pi}\mathrm{e}^{\mathrm{i}\cdot(n-m)\,x}\,\mathrm{d}x = \frac{1}{\mathrm{i}\cdot(n-m)}\cdot\left[\mathrm{e}^{\mathrm{i}\cdot(n-m)\,x}\right]_0^{2\,\pi} = 0$$

si $n \neq m$.

Les fonctions $(e_k)_{k\in[\![-n,\,n]\!]}$ engendrent un sous-espace vectoriel F_n de D de dimension $2\,n+1$, celui des polynômes trigonométriques de degré inférieur ou égal à n, et en constituent une base orthonormée. Le polynôme trigonométrique $f_n \in F_n$ le plus proche de f, au sens de la norme $\|\cdot\|_2$, est la projection orthogonale de f sur F_n (cf. annexe § VII.2). Dans cette base, on a

$$f_n(x) = \sum_{k=-n}^{n}\langle e_k, f\rangle\,e_k = \sum_{k=-n}^{n} c_k\,\mathrm{e}^{\mathrm{i}\,k\,x},$$

où $c_k := \langle e_k, f\rangle$.

Les fonctions $t_0\colon x\mapsto 1$, $t_n\colon x\mapsto\sqrt{2}\,\cos(n\,x)$ et $t_{-n}\colon x\mapsto\sqrt{2}\,\sin(n\,x)$ ($n\in\mathbb{N}^*$) constituent une autre famille orthonormée de D. Les fonctions $(t_k)_{k\in[\![-n,\,n]\!]}$ constituent également une base orthonormée de F_n. Dans cette base, la projection orthogonale f_n vaut

$$\begin{aligned}f_n(x) &= \sum_{k=-n}^{n}\langle t_k, f\rangle\,t_k = \frac{a_0}{2}\times 1 + \sum_{k=1}^{n}\left(\frac{a_k}{\sqrt{2}}\times\sqrt{2}\,\cos(k\,x) + \frac{b_k}{\sqrt{2}}\times\sqrt{2}\,\sin(k\,x)\right)\\ &= \frac{a_0}{2} + \sum_{k=1}^{n}(a_k\cos(k\,x) + b_k\sin(k\,x)),\end{aligned}$$

où $a_0 := 2\,\langle t_0, f\rangle$, $a_{k,\,k\in\mathbb{N}^*} := \sqrt{2}\,\langle t_k, f\rangle$ et $b_{k,\,k\in\mathbb{N}^*} := \sqrt{2}\,\langle t_{-k}, f\rangle$.

Définition 38
On appelle *développement en série de Fourier* de f la série trigonométrique

$$\widetilde{f}(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n\cos(n\,x) + b_n\sin(n\,x)) \tag{7}$$

$$= \sum_{n=-\infty}^{n} c_n\,\mathrm{e}^{\mathrm{i}\,n\,x}, \tag{8}$$

où les *coefficients de Fourier* a_n, b_n et c_n sont définis par

$$a_n = \frac{1}{\pi}\int_0^{2\,\pi} f(x)\cos(n\,x)\,\mathrm{d}x,\quad \forall\,n\in\mathbb{N}, \tag{9}$$

$$b_n = \frac{1}{\pi}\int_0^{2\,\pi} f(x)\sin(n\,x)\,\mathrm{d}x,\quad \forall\,n\in\mathbb{N}^* \tag{10}$$

$$\text{et}\quad c_n = \frac{1}{2\,\pi}\int_0^{2\,\pi} f(x)\,\mathrm{e}^{-\mathrm{i}\,n\,x},\quad \forall\,n\in\mathbb{Z}. \tag{11}$$

⌟

1. Pour des fonctions f et g de période T, le produit scalaire est défini par

$$\langle f, g\rangle = \frac{1}{T}\int_0^{T}\overline{f(x)}\,g(x)\,\mathrm{d}x.$$

Remarques

- Ce développement ne converge pas nécessairement vers f en tout point, ni même n'est convergent partout.
- L'expression $\sum_{n=-\infty}^{\infty} c_n\, \mathrm{e}^{\mathrm{i}\, n\, x}$ doit être comprise de la manière suivante :

$$\sum_{n=-\infty}^{\infty} c_n\, \mathrm{e}^{\mathrm{i}\, n\, x} := \lim_{n\to\infty} \sum_{k=-n}^{n} c_k\, \mathrm{e}^{\mathrm{i}\, k\, x}.$$

- Les a_k, b_k et c_k ne dépendent pas de n : si n' est un entier supérieur à n et si les coefficients a'_k, b'_k et c'_k correspondent à la projection de f sur $F_{n'}$, on a $a'_k = a_k$, $b'_k = b_k$ et $c'_k = c_k$ pour tout k tel que $|k| \leqslant n$.
- Si f est paire, tous les b_k sont nuls ; si f est impaire, tous les a_k sont nuls.
- Le développement en série de Fourier peut également être écrit

$$\widetilde{f}(x) = \sum_{n=0}^{\infty} A_n \cos(n\, x - \varphi_n).$$

  A_n est l'amplitude du n-ième mode de Fourier et φ_n sa phase.
- L'ensemble des coefficients (a_n, b_n), (A_n, φ_n) ou c_n constitue le *spectre de Fourier* de la fonction f. La fonction $x \mapsto a_0/2 = c_0 = A_0 \cos\varphi_0$, pour $n = 0$, ou

$$x \mapsto a_n \cos(n\, x) + b_n \sin(n\, x) = c_n\, \mathrm{e}^{\mathrm{i}\, n\, x} + c_{-n}\, \mathrm{e}^{-\mathrm{i}\, n\, x} = A_n \cos(n\, x - \varphi_n), \quad \text{pour } n > 0,$$

  est le n-ième *harmonique* (ou n-ième *mode de Fourier*) du développement de f. ⌟

VII. .2. Relations entre les coefficients de Fourier

À l'aide des formules d'Euler,

$$\cos(n\, x) = \frac{\mathrm{e}^{\mathrm{i}\, n\, x} + \mathrm{e}^{-\mathrm{i}\, n\, x}}{2} \quad \text{et} \quad \sin(n\, x) = \frac{\mathrm{e}^{\mathrm{i}\, n\, x} - \mathrm{e}^{-\mathrm{i}\, n\, x}}{2\,\mathrm{i}},$$

on obtient aisément les relations suivantes entre les coefficients de Fourier :

$$a_n = c_n + c_{-n} \quad \text{et} \quad b_n = \mathrm{i}\cdot(c_n - c_{-n}).$$

Supposons que la fonction f soit à valeurs réelles. C'est alors aussi le cas pour $\widetilde{f}$. En écrivant

$$\widetilde{f}(x) = \sum c_n\, \mathrm{e}^{\mathrm{i}\, n\, x} = \sum c_{-n}\, \mathrm{e}^{\mathrm{i}\cdot(-n)\, x} = \overline{\widetilde{f}(x)} = \sum \overline{c_n}\, \mathrm{e}^{-\mathrm{i}\, n\, x}$$

et en identifiant terme à terme, on voit que $\overline{c_n} = c_{-n}$. Si f est à valeur réelle, on a donc $a_n = 2\,\mathrm{Re}\, c_n$ et $b_n = -2\,\mathrm{Im}\, c_n$.

On peut poser $\varphi_0 = 0$; on a alors $A_0 = a_0/2$. Pour $n > 0$, $A_n = \sqrt{a_n^2 + b_n^2}$, $\cos\varphi_n = a_n/A_n$ et $\sin\varphi_n = b_n/A_n$.

VII. .3. Fonction T-périodique

On a raisonné sur l'intervalle $[-\pi, \pi[$, mais il est aisé de transposer toute la discussion à une fonction f périodique de période T.

Posons $u = 2\,\pi\, x/T$ et $g(u) = f(T\, u/(2\,\pi)) = f(x)$. La fonction g est de période $2\,\pi$:

$$g(u + 2\,\pi) = f(T\cdot(u + 2\,\pi)/(2\,\pi)) = f(x + T) = f(x) = g(u).$$

Son développement en série de Fourier est donc

$$\widetilde{g}(u) = \frac{1}{2}\, a_0 + \sum_{n=1}^{\infty} (a_n \cos(n\, u) + b_n \sin(n\, u)),$$

où

$$a_n = \frac{1}{\pi} \int_{u=0}^{2\,\pi} g(u) \cos(n\, u)\, \mathrm{d}u \quad \text{et} \quad b_n = \frac{1}{\pi} \int_{u=0}^{2\,\pi} g(u) \sin(n\, u)\, \mathrm{d}u.$$

Exprimons a_n en fonction de f et de $x = T\, u/(2\,\pi)$. On a

$$a_n = \frac{1}{\pi} \int_{u=0}^{2\,\pi} g(u) \cos(n\, u)\, \mathrm{d}u = \frac{1}{\pi} \int_{x=0}^{T} f(x) \cos\Bigl(n\, \frac{2\,\pi\, x}{T}\Bigr)\, \mathrm{d}\Bigl(\frac{2\,\pi\, x}{T}\Bigr) = \frac{2}{T} \int_{x=0}^{T} f(x) \cos\Bigl(\frac{2\,\pi}{T}\, n\, x\Bigr)\, \mathrm{d}x.$$

De même pour b_n et c_n.

Le développement en série de Fourier d'une fonction f de période T est donc

$$\begin{aligned}\widetilde{f}(x) &= \frac{1}{2}\, a_0 + \sum_{n=1}^{\infty}\left(a_n \cos\!\left(\frac{2\,\pi}{T}\, n\, x\right) + b_n \sin\!\left(\frac{2\,\pi}{T}\, n\, x\right)\right) \\ &= \sum_{n=-\infty}^{\infty} c_n \exp\!\left(\frac{2\,\pi}{T}\, \mathrm{i}\, n\, x\right),\end{aligned}$$

où

$$a_n = \frac{2}{T}\int_{x=0}^{T} f(x) \cos\!\left(\frac{2\,\pi}{T}\, n\, x\right) \mathrm{d}x, \quad b_n = \frac{2}{T}\int_{x=0}^{T} f(x) \sin\!\left(\frac{2\,\pi}{T}\, n\, x\right) \mathrm{d}x$$
$$\text{et} \quad c_n = \frac{1}{T}\int_{x=0}^{T} f(x) \exp\!\left(-\frac{2\,\pi}{T}\, \mathrm{i}\, n\, x\right) \mathrm{d}x.$$

VII. . Convergence en moyenne quadratique

Pour toute fonction $f \in D$, on a $\|f_n\|_2 \leqslant \|f\|_2$, soit

$$\|f_n\|_2^2 = \sum_{k=-n}^{n} |c_k|^2 = \frac{|a_0|^2}{4} + \sum_{k=1}^{n} \frac{|a_k|^2 + |b_k|^2}{2} \leqslant \|f\|_2^2.$$

Ceci est vrai plus généralement pour toute fonction f continue par morceaux et $2\,\pi$-périodique, car on peut toujours trouver une fonction $g \in D$ telle que $\|g\|_2 = \|f\|_2$.

Le terme de gauche admet une limite finie car il croît avec n et est majoré par le terme de droite. Pour toute fonction continue par morceaux et $2\,\pi$-périodique, on a donc

$$\lim_{n\to\infty} \|f_n\|_2^2 = \sum_{k=-\infty}^{\infty} |c_k|^2 = \frac{|a_0|^2}{4} + \sum_{k=1}^{\infty} \frac{|a_k|^2 + |b_k|^2}{2} \leqslant \|f\|_2^2 \qquad \text{(inégalité de Bessel).}$$

Pour une fonction f continue par morceaux et de période $2\,\pi$, les coefficients a_n, b_n et c_n sont bien définis, donc f_n aussi ; elle est même continue (c'est un polynôme trigonométrique). Le fait que $\lim_{n\to\infty} \|f_n\|_2^2$ existe permet de montrer que la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ est de Cauchy, mais ne permet pas de conclure que $\lim_{n\to\infty} f_n$ existe dans l'espace des fonctions continues, ou même seulement continues par morceaux, car ces espaces ne sont pas complets. En revanche, ceux-ci sont des sous-ensembles de l'espace des fonctions $2\,\pi$-périodiques de carré intégrable, c.-à-d. l'ensemble $L^2([0, 2\,\pi])$ des fonctions $g\colon \mathbb{R} \to \mathbb{C}$ telles que $\int_0^{2\,\pi} |g|^2$ [*2] existe. Or l'ensemble $L^2([0, 2\,\pi])$, muni du produit scalaire défini précedemment, est un *espace de Hilbert* (cf. annexe § VII.3) pour la norme $\|\cdot\|_2$ [*3] (enfin presque, car $\|\cdot\|_2$ est seulement une semi-norme). La suite de fonctions $(f_n)_{n\in\mathbb{N}}$ étant de Cauchy, elle converge donc vers une fonction $\widetilde{f}$ de $L^2([0, 2\,\pi])$ définie presque partout sur $[0, 2\,\pi]$.

Les familles $(e_n)_{n\in\mathbb{Z}}$ et $(t_n)_{n\in\mathbb{Z}}$ constituent des *bases hilbertiennes* orthonormées et dénombrables de $L^2([0, 2\,\pi])$. L'application $\|\cdot\|_2$ étant seulement une semi-norme, on peut juste affirmer que $\widetilde{f}$ ($:= \lim_{n\to\infty} f_n$) est égale presque partout à f, ce qu'on note $\widetilde{f}\, f$.

On peut alors énoncer le théorème suivant :

Théorème 73 (théorème de Parseval-Bessel)
Pour tout fonction f continue par morceaux et $2\,\pi$-périodique (ou, plus généralement, pour toute fonction de $L^2([0, 2\,\pi])$), les séries suivantes convergent (presque partout pour la première) et l'on a

$$\forall\, x \in \mathbb{R}, \quad f(x) \widetilde{f}(x) = \sum_{n=-\infty}^{\infty} c_n\, \mathrm{e}^{\mathrm{i}\, n\, x} = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(n\, x) + b_n \sin(n\, x)),$$

et

$$\|f\|_2^2 = \|\widetilde{f}\,\|_2^2 = \sum_{n=-\infty}^{\infty} |c_n|^2 = \frac{|a_0|^2}{4} + \frac{\sum_{n=1}^{\infty} (|a_n|^2 + |b_n|^2)}{2}.$$

⌟

2. Il s'agit d'une intégrale de Lebesgue, pas de Riemann, donc il n'est pas nécessaire que g soit définie sur tout l'intervalle : il suffit qu'elle le soit presque partout.
3. Il ne sera pas nécessairement complet pour d'autres normes s'il est de dimension infinie, car celles-ci peuvent ne pas être équivalentes.

Comme $f = f_n + (f - f_n)$ et que $f_n \in F_n$ et $f - f_n \perp F_n$, on a $\|f\|_2^2 = \|f_n\|_2^2 + \|f - f_n\|_2^2$. D'après ce qui précède, $\|f_n\|_2$ converge vers $\|\widetilde{f}\|_2 = \|f\|_2$, donc $\|f - \widetilde{f}\|_2 = 0$: on dit que $\widetilde{f}$ *converge en moyenne quadratique* vers f.

Exemple 62
Considérons la fonction valant $f(x) = 0$ sur $]-\pi, 0[$, $f(x) = 1$ sur $]0, \pi[$ et prolongée $2\,\pi$-périodiquement. On peut (mais ça n'est pas nécessaire, comme on vient de le faire remarquer) convenir que $f(0) = f(\pi) = 1/2$ de manière à ce que $f \in D$. Les coefficients de Fourier de f valent $c_0 = 1/2$ et $c_{n\neq 0} = (1 - (-1)^n)/(2\,\mathrm{i}\,\pi\,n)$. Donc, si $n = \pm 2\,k$ (avec $k \in \mathbb{N}^*$), $c_n = 0$; si $n = \pm(2\,k + 1)$ (avec $k \in \mathbb{N}$), $c_n = 1/(\mathrm{i}\,\pi\,n)$, soit $|c_{\pm(2\,k+1)}|^2 = 1/(\pi^2{\cdot}(2\,k + 1)^2)$.

Par ailleurs, $\|f\|^2 = (2\,\pi)^{-1} \int_0^\pi 1^2\,\mathrm{d}x = 1/2$. D'après le théorème de Parseval-Bessel, on a donc

$$\frac{1}{2} = \left(\frac{1}{2}\right)^2 + 2\sum_{k=0}^{\infty} \frac{1}{\pi^2{\cdot}(2\,k+1)^2},$$

soit

$$\sum_{k=0}^{\infty} \frac{1}{(2\,k+1)^2} = \frac{\pi^2}{8}.$$

⌟

VII. . Convergence ponctuelle

Nous avons vu ci-dessus que les coefficients du développement en série de Fourier sont définis pour toute fonction continue par morceaux et périodique. Bien que $\|f - \widetilde{f}\|_2$ soit égale à 0, ces conditions permettent seulement de dire que, « presque partout », $\widetilde{f}$ converge et ce vers f : la convergence en moyenne quadratique ne permet pas de conclure à la convergence ponctuelle en tout point.

VII. .1. Théorèmes de convergence

Les a_n et b_n étant déterminés à partir de la fonction f, la question est maintenant de savoir si le développement $\widetilde{f}$ converge, et, si oui, s'il converge vers f en tout point.

Nous admettrons le théorème suivant :

Théorème 74 (théorème de Dirichlet)
Soit $f : \mathbb{R} \to \mathbb{C}$ une application $2\,\pi$-périodique.

Si f est C^1 par morceaux, son développement en série de Fourier $\widetilde{f}$ est défini en tout point et

$$\forall\, x \in \mathbb{R}, \quad \widetilde{f}(x) = \frac{f(x^-) + f(x^+)}{2}.$$

Dans tout sous-intervalle où la fonction f est continue, la convergence de $\widetilde{f}$ vers f est uniforme. ⌟

Remarques

- Plutôt que l'expression de f sur $\mathbb{R}$, on donne souvent sa restriction sur un intervalle I de longueur $2\,\pi$, $f|_I$, avec $I = [a, a + 2\,\pi[$ ou $I =]a, a + 2\,\pi]$ et $a \in \mathbb{R}$; f est ensuite définie en prolongeant $f|_I$ par $2\,\pi$-périodicité, c.-à-d. que, pour tout $x \in \mathbb{R}$, on pose $f(x) := f|_I(x_0)$, où x_0 est l'unique élément de I tel que $x = x_0 \pmod{2\,\pi}$.

  Par exemple, si $I = [-\pi, \pi[$, $f(\pi^+) := f|_I([-\pi]^+) = f([-\pi]^+)$. Si $f|_I$ est C^1_{m}, on a alors

  $$\widetilde{f}(\pm\pi) = \frac{f(\pi^-) + f([-\pi]^+)}{2}.$$

- Attention, si f est définie par prolongement $2\,\pi$-périodique de sa restriction $f|_I$ à un intervalle I de longueur $2\,\pi$, le fait que $f|_I$ soit continue (en plus d'être C^1_{m}) ne suffit pas pour montrer que la convergence de $\widetilde{f}$ vers f est uniforme sur I : il faut que f (pas seulement $f|_I$) soit continue sur I.

  Le problème se pose aux bornes de $I = [a, a + 2\,\pi[$ ou $]a, a + 2\,\pi]$. Or f peut être discontinue en $a \pmod{2\,\pi}$ même si $f|_I$ est continue ou prolongeable par continuité en ce point.

  Exemple. — Soit f la fonction définie par $f(x) = x$ sur $[-\pi, \pi[$ et prolongée $2\,\pi$-périodiquement. $f|_{[-\pi,\,\pi[}$ est continue et est même prolongeable en une fonction continue sur $[-\pi, \pi]$. Pourtant, f (en tant que fonction définie sur $\mathbb{R}$) n'est pas continue en $\pm\pi$ puisque $f([\pm\pi]^+) = -\pi \neq f([\pm\pi]^-) = \pi$. ⌟

Aux conditions du théorème de Dirichlet, on préfère parfois l'une de celles-ci, plus générales :

- f n'a sur l'intervalle $[-\pi, \pi[$ qu'un nombre fini de discontinuités finies et a ailleurs une dérivée bornée ;

ou bien encore

- f est bornée et n'a qu'un nombre fini de minima et de maxima dans cet intervalle.

Sous ces conditions (« conditions de Dirichlet »), la conclusion précédente que la série converge en tout point vers $\frac{1}{2}\big(f(x^-)+f(x^+)\big)$ est encore vraie.

Il existe des conditions plus faibles assurant la convergence (conditions suffisantes). La version suivante, due à Jordan, semble la meilleure.

Théorème 75 (théorème de Jordan)
Si une fonction est à variation bornée sur l'intervalle période, la série de Fourier $\widetilde{f}$ converge en tout point vers la demi-somme $\frac{1}{2}\big(f(x^-)+f(x^+)\big)$. ⌟

Une fonction à variation bornée est définie comme suit. Pour une fonction définie sur un intervalle $[a, b]$, on considère une subdivision arbitraire de $[a, b]$, c.-à-d. un choix de points $x_0 = a < x_1 < \cdots < x_{n-1} < x_n = b$ et on calcule la *variation* correspondante $\sum_{i=1}^{n} |f(x_i) - f(x_{i+1})|$. La fonction est à variation bornée si et seulement si cette variation est bornée par un B indépendant de la subdivision. Une fonction monotone bornée est à variation bornée (on démontre d'ailleurs que toute fonction à variation bornée est la différence de deux fonctions croissantes bornées) ; une fonction bornée n'ayant qu'un nombre fini de maxima ou minima sur $[a, b]$ est aussi à variation bornée ; enfin, une fonction continue et à dérivée bornée est encore à variation bornée. (Cette grande variété de cas impliquant la propriété d'être à variation bornée explique la multiplicité de conditions suffisantes de convergence des séries de Fourier rencontrées dans la littérature. . .)

Plus la fonction est régulière, plus les coefficients de Fourier tendent rapidement vers 0.

Théorème 76
Soit $f : \mathbb{R} \to \mathbb{C}$ une application $2\,\pi$-périodique.
Si f est C^0_{m}, $(a_n, b_n, c_{\pm n}) = o_{n\to\infty}(1)$.
Si f est C^{k-1} et C^k_{m} (*4),

$$c_n\big(f^{(k)}\big) = \frac{1}{(\mathrm{i}\,n)^k}\,c_n(f) \quad \text{et} \quad (a_n, b_n, c_{\pm n}) = \underset{n\to\infty}{o}\left(\frac{1}{n^k}\right).$$
⌟

Preuve
Supposons d'abord f continue par morceaux. Les coefficients de Fourier sont alors bien définis. Comme $\sum_{n=-\infty}^{\infty} |c_n|^2$ converge (cf. inégalité de Bessel), on en déduit que $\lim_{n\to\infty} c_{\pm n} = 0$.

Supposons f désormais C^{k-1} et C^k_{m}. Calculons $c_n(f')$ par intégration par parties :

$$c_n(f') = \frac{1}{2\,\pi}\int_{x=0}^{2\,\pi} \mathrm{e}^{-\mathrm{i}\,n\,x}\, f'(x)\,\mathrm{d}x = \frac{1}{2\,\pi}\left(\Big[\mathrm{e}^{-\mathrm{i}\,n\,x}\, f(x)\Big]_{x=0}^{2\,\pi} - \int_{x=0}^{2\,\pi} (-\mathrm{i}\,n)\,\mathrm{e}^{-\mathrm{i}\,n\,x}\, f(x)\,\mathrm{d}x\right) = \mathrm{i}\,n\,c_n(f).$$

Par récurrence, on en déduit l'expression de $c_n(f^{(k)})$. $f^{(k)}$ étant C^0_{m}, on a $c_{\pm n}(f^{(k)}) = o(1)$, soit $c_{\pm n}(f) = o(1/n^k)$. ⌟

Remarque
Si f est définie par prolongement $2\,\pi$-périodique de $f|_I$, où I est un intervalle de longueur $2\,\pi$, il ne suffit pas que $f|_I$ soit C^{k-1} (en plus d'être C^k_{m}) pour que $c_n = o(1/n^k)$: il faut que f soit C^{k-1} sur I, ce qui est différent comme on l'a souligné dans la remarque sur le théorème de Dirichlet. ⌟

La convergence est donc moins bonne pour les dérivées successives. En fait, la dérivation terme à terme du développement de Fourier de f ne donne pas le développement de Fourier de f' si f est discontinue. Un exemple est donné par la fonction $x \mapsto f(x)$ périodique et valant x sur sa période $[-T/2, T/2[$, dont le développement de Fourier non trivial sera donné au § VII.E.2. Sa dérivée terme à terme est également non triviale et ne s'identifie pas au développement de Fourier de $x \mapsto f'(x) = 1$ qui se réduit au terme constant ! À l'inverse, la convergence est meilleure pour l'intégrale de f.

Exemple 63
Soit la fonction $f(x) = x^2$ sur l'intervalle $[-\pi, \pi[$, prolongée par périodicité de période $2\,\pi$ au-delà. On calcule aisément $c_0 = (2\,\pi)^{-1}\int_{-\pi}^{\pi} x^2\,\mathrm{d}x = \pi^2/3$ et, après deux intégrations par parties, $c_n = 2 \times (-1)^n/n^2$ pour $n \neq 0$. On a donc, puisque f est C^1_{m} et C^0,

$$x^2 = \frac{\pi^2}{3} + 2\sum_{n\neq 0}\frac{(-1)^n}{n^2}\,\mathrm{e}^{\mathrm{i}\,n\,x} = \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty}(-1)^n\,\frac{\cos(n\,x)}{n^2}.$$

Les deux séries convergent absolument et uniformément. On voit sur la figure VII.1 que le pic en $\pm\pi$ est de mieux en mieux représenté quand la série est tronquée à un ordre plus élevé. Ceci reflète bien le fait que les hautes fréquences (grands « harmoniques », c.-à-d. modes à n grand) donnent les détails fins de la fonction.

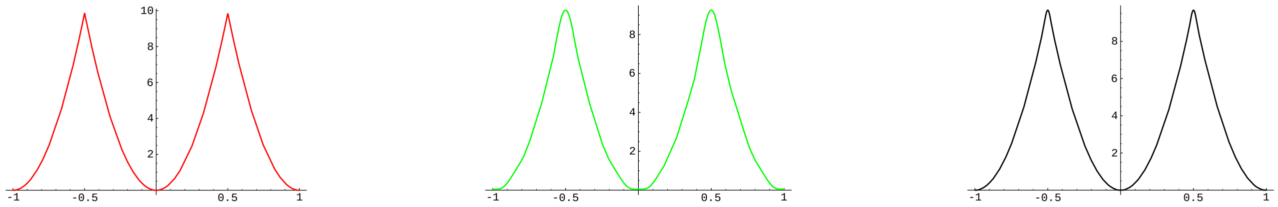


Figure VII.1 – À gauche, la fonction périodique valant x^2 sur sa période $[-\pi, \pi[$; au milieu (resp. à droite), son développement en série de Fourier jusqu'au 6e (resp. 20e) terme. La variable $x/(2\,\pi)$ est portée en abscisse.

En calculant f et $\widetilde{f}$ en $x = 0$ et $x = \pi$, on obtient respectivement que

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^2} = \frac{\pi^2}{12} \quad \text{et} \quad \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

⌟

Les exemples suivants seront étudiés en TD.

Exemple 64
Soient u et x deux variables réelles. En prenant les parties réelle et imaginaire de la série géométrique $\sum_{n=0}^{\infty} u^n\, e^{i\,n\,x} = 1/(1 - u\, e^{i\,x})$, absolument convergente pour $|u| < 1$ et pour tout x (donc uniformément convergente), on obtient deux séries de Fourier uniformément convergentes :

$$\begin{aligned} \frac{1 - u\cos x}{1 - 2\,u\cos x + u^2} &= \sum_{0}^{\infty} u^n \cos(n\,x) \\ \text{et} \quad \frac{u\sin x}{1 - 2\,u\cos x + u^2} &= \sum_{1}^{\infty} u^n \sin(n\,x). \end{aligned} \tag{12}$$

De même, en prenant les parties réelle et imaginaire de $\ln(1 - u\, e^{i\,x})$, on trouve

$$\begin{aligned} \ln(1 - 2\,u\cos x + u^2) &= -2\sum_{n=1}^{\infty} \frac{u^n}{n} \cos(n\,x) \\ \text{et} \quad \arctan\frac{u\sin x}{1 - u\cos x} &= \sum_{n=1}^{\infty} \frac{u^n}{n} \sin(n\,x). \end{aligned} \tag{13}$$

Montrer qu'on passe de (13) à (12) en dérivant par rapport à u. ⌟

Exemple 65
Soit la fonction $f(x) = \cos(a\,x)$, pour a non entier et $-\pi \leqslant x \leqslant \pi$. Sa série de Fourier est

$$\frac{\sin(a\,\pi)}{\pi} \cdot \left(\frac{1}{a} + \sum_{n \geqslant 1} (-1)^n\, \frac{2\,a\cos(n\,x)}{a^2 - n^2} \right).$$

Montrer que la série converge vers $f(x)$ pour $-\pi \leqslant x \leqslant \pi$. ⌟

On verra d'autres exemples de séries de Fourier dans la suite.

VII. .2. Phénomène de Gibbs

Que se passe-t-il au voisinage d'un point de discontinuité de la fonction f ?

Considérons par exemple la fonction $f : x \mapsto x$ sur l'intervalle $[-1/2, 1/2[$ et prolongée par périodicité de période 1 (noter que cette fonction est, à un changement d'échelle près, la dérivée de la fonction de l'exemple 63). On montre aisément (cf. TD) que son développement de Fourier s'écrit

$$f(x) = \frac{1}{\pi} \sum_{n=1}^{\infty} (-1)^{n-1}\, \frac{\sin(2\,n\,\pi\,x)}{n}$$

et, selon le théorème, ce développement converge vers f dans $]-1/2, 1/2[$, tandis qu'il vaut $f(-1/2) + f(1/2) = 0$ en $\pm 1/2$.

Représentons le développement tronqué à la N-ième harmonique, $\pi^{-1} \sum_{n=1}^{N} (-1)^{n-1} \sin(2\,n\,\pi\,x)/n$. On voit (figure VII.2) que quelques harmoniques reproduisent déjà très bien la discontinuité de la fonction en π. Conformément au théorème, quand on augmente le nombre de termes, la série de Fourier approxime de mieux en mieux la fonction en dehors de ses discontinuités, mais on voit qu'elle exagère la valeur de la discontinuité. Plus précisément, la position du maximum (ou minimum) tend vers le point de discontinuité x_0, mais la valeur à ce maximum ou minimum ne tend pas vers la valeur de $f(x_0^{\pm})$. C'est le *phénomène de Gibbs*. On peut calculer l'excédent et on trouve qu'il tend vers environ 18 % de la discontinuité (cf. TD) : la série de Fourier ne converge donc pas uniformément !

4. $f^{(k)}$ n'est pas définie aux points de discontinuité de $f^{(k-1)}$, mais ceci n'interdit pas de calculer ses coefficients de Fourier. Elle est en effet continue sur des intervalles $]x_i, x_{i+1}[$ et prolongeable par continuité sur $[x_i, x_{i+1}]$. L'intégrale $\int_{x_i}^{x_{i+1}} e^{-i\,n\,x} f(x)\, dx$ est donc faussement impropre.

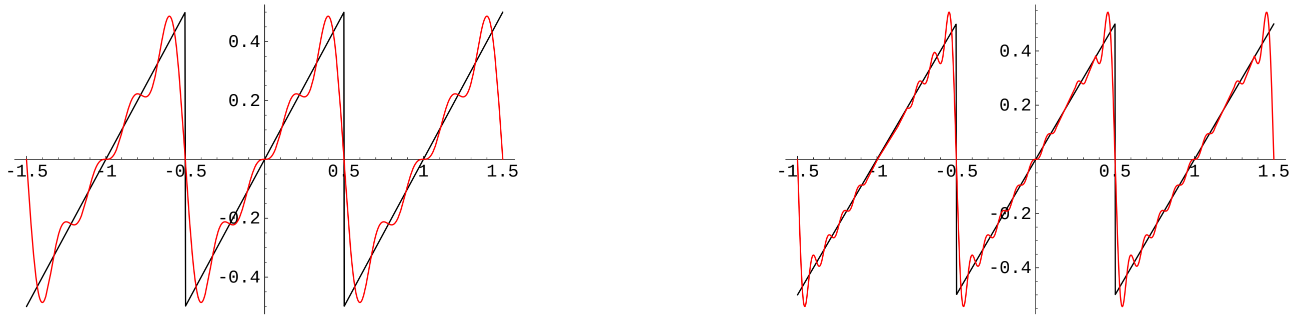


Figure VII.2 – Fonction $f : x \mapsto x$ sur $[-1/2, 1/2]$ et ses développements jusqu'au 4^e ordre (à gauche) et 10^e ordre (à droite).

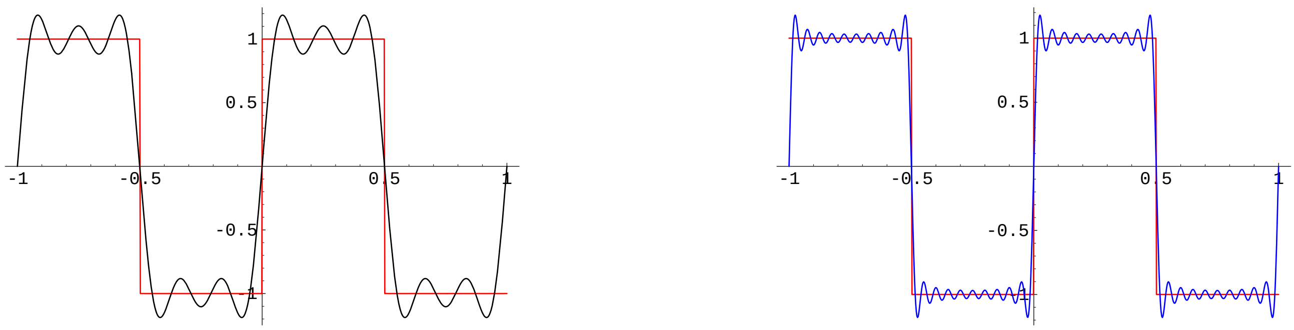


Figure VII.3 – La fonction périodique valant $\operatorname{sgn} x$ (signe de x) sur sa période $[-\pi, \pi]$ (« fonction créneau »), superposée à sa série de Fourier avec 6 termes (à gauche) et 20 termes (à droite). La variable $x/(2\pi)$ est portée en abscisse.

Autre exemple : la fonction créneau, périodique de période 2π, valant $f(x) = -1$ si $x \in]-\pi, 0[$, et $f(x) = 1$ si $x \in]0, \pi[$. Sa série de Fourier n'a que des termes en sinus, et un calcul simple conduit à

$$f(x) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \sin((2n+1)x).$$

La convergence est lente, en $1/n$, reflétant la présence d'une discontinuité, et on observe à nouveau le phénomène de Gibbs (voir figure VII.3).

VII. . Analyse d'un système périodique avec une bande passante finie

En pratique, le développement (7) (ou (8)) est rarement utilisé ou connu dans son intégralité. Ou bien, parce qu'on ne peut calculer en pratique qu'un nombre fini de termes, il est tronqué à un nombre de modes fini, la somme sur n courant de 0 à N (ou de $-N$ à N dans (8)). Ou bien, comme c'est souvent le cas en physique, la fonction représente un signal qui a pu être déformé lors de sa réception ou de son traitement, affectant ainsi les coefficients des différentes harmoniques, et tronquant éventuellement les basses ou les hautes fréquences (on parle de *filtre passe-haut* (resp. *passe-bas*) pour un dispositif qui « coupe » les basses (resp. les hautes) fréquences). On dira qu'on analyse la fonction f avec une *bande passante* finie si les hautes fréquences sont coupées ou fortement atténuées. Quel en est l'effet sur la reproduction de la fonction ?

On a vu au § VII.E.2 l'effet d'une coupure des harmoniques élevés sur les fonctions $x \mapsto x^2$, $x \mapsto x$ et $x \mapsto \operatorname{sgn} x$ prolongées de manière périodique : les détails fins de la fonction sont mal reproduits. En général, si les coefficients de la série décroissent lentement (par exemple, décroissance en $o(1)$ des coefficients de Fourier d'une fonction continue par morceaux mais non continue), la somme sera sensible à une coupure des hautes fréquences. Pour ces fonctions, dont les premiers modes de Fourier sont d'ordre 1, une coupure des basses fréquences n'est pas non plus anodine. . .

À titre d'exemple, considérons la fonction paire

$$f : x \mapsto e^{-\Gamma |x|}$$

définie sur l'intervalle $[-\pi, \pi[$, ou, si l'on préfère, pour tout x réel, en répétant l'intervalle $[-\pi, \pi[$ périodiquement. Pour Γ petit, la fonction est étalée autour de 0, et son pic se rétrécit quand Γ croît.

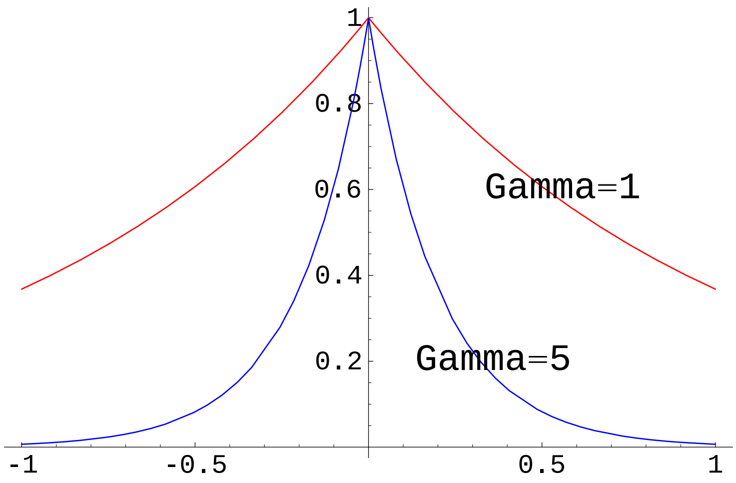


Figure VII.4 – La fonction $x \mapsto \mathrm{e}^{-\Gamma\,|x|}$ pour $\Gamma = 1$ et $\Gamma = 5$.

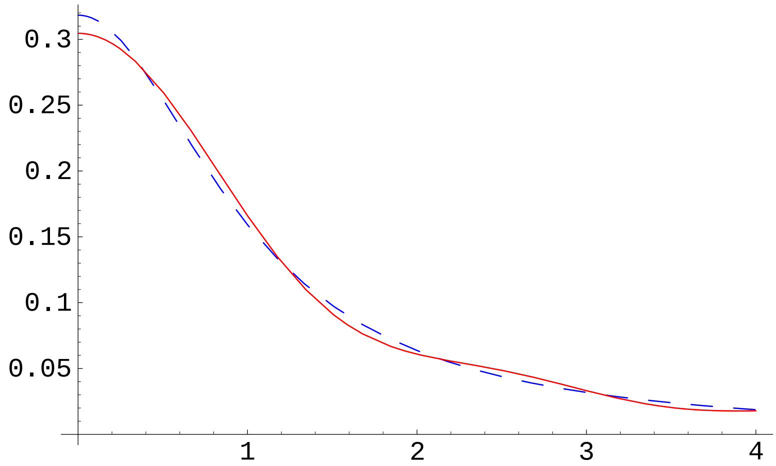


Figure VII.5 – Coefficients de Fourier a_n de la fonction f pour $\Gamma = 1$. Les a_n sont donnés en fonction de $n \in \mathbb{R}$ (au lieu de $\mathbb{N}$) par la formule (14) (trait plein) et par l'approximation lorentzienne (ligne brisée).

La fonction ainsi définie est continue et périodique ; selon les théorèmes précédents, son développement de Fourier doit converger. On calcule aisément les coefficients de Fourier :

$$a_n = \frac{2\,\Gamma}{\pi}\,\frac{1-(-1)^n\,\mathrm{e}^{-\Gamma\,\pi}}{\Gamma^2+n^2} \quad \text{et, en particulier,} \quad a_0 = \frac{2}{\pi\,\Gamma}\cdot\left(1-\mathrm{e}^{-\Gamma\,\pi}\right), \tag{14}$$

tandis que $b_n = 0$ (fonction paire). Pour $\Gamma\,\pi \gg 1$, on a $a_n \approx 2\,\Gamma/(\pi\cdot(\Gamma^2+n^2))$ qui est une fonction « lorentzienne » de la variable n (figure VII.5). Son graphe est une courbe en cloche plus étalée qu'une gaussienne [(*5)], dont on peut mesurer l'étalement en calculant la valeur de n où l'on atteint la demi-hauteur du maximum, soit $2\,\Gamma/(n^2+\Gamma^2) \approx 1/\Gamma$, c.-à-d. $|n| \approx \Gamma$.

La conclusion de ce calcul, qu'on pourrait répéter avec d'autres fonctions, est qu'à une fonction f « piquée » (resp. étalée) correspond une distribution large (resp. étroite) de ses coefficients de Fourier. En particulier, dans notre exemple, à grand Γ (fonction f très piquée), on doit utiliser beaucoup d'harmoniques ($|n| \sim \Gamma$) pour reproduire correctement la fonction. Inversement, une bande passante étroite (une limite basse sur les n) interdit de « voir » des détails fins de la fonction f. La bande passante doit être d'autant plus large que les détails de la fonction sont plus fins. Il s'agit là d'un aspect fondamental de l'analyse de Fourier, qu'on rencontre notamment en optique : avoir une bonne résolution spatiale, c.-à-d. pouvoir discerner des détails fins, impose d'utiliser des petites longueurs d'onde (c.-à-d. des grandes fréquences). Ce même principe apparaît sous des habits variés dans des situations multiples, telles les inégalités d'Heisenberg en mécanique quantique.

5. Rappelons que la fonction gaussienne, $x \mapsto \mathrm{e}^{-\alpha\,x^2}$, a pour graphe une courbe en cloche. Voir aussi l'exercice de TD sur la distribution de Maxwell.

Annexe VII.1. Espaces préhilbertiens

Définition 39
Soit E un $\mathbb{K}$-espace vectoriel, avec $\mathbb{K} = \mathbb{C}$. Une application

$$\begin{aligned}\langle\cdot,\cdot\rangle\colon\quad E^2 &\longrightarrow \mathbb{K}\\ (x, y) &\longmapsto \langle x, y\rangle\end{aligned}$$

est un *produit hermitien* si les conditions suivantes sont toutes réunies :

1. $\forall\,(x, y) \in E^2$, $\langle y, x\rangle = \overline{\langle x, y\rangle}$, où $\overline{k}$ désigne le conjugué d'un scalaire k ;
2. $\forall\,(x, y, z, \lambda) \in E^3 \times \mathbb{K}$, $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$ et $\langle x, \lambda\, y\rangle = \lambda\,\langle x, y\rangle$ (linéarité à droite) ;
3. $\forall\, x \in E$, $\langle x, x\rangle \geqslant 0$;
4. $\forall\, x \in E$, $\langle x, x\rangle = 0 \Longrightarrow x = 0$.

Si $\mathbb{K} = \mathbb{R}$, on parle de *produit scalaire*.

On note $\|\cdot\|_2$ l'application

$$\begin{aligned}\|\cdot\|_2\colon E &\longrightarrow \mathbb{R}^+\\ x &\longmapsto \|x\|_2 = \sqrt{\langle x, x\rangle}.\end{aligned}$$

On appelle *espace préhilbertien* complexe (resp. réel), et l'on note $(E, \langle\cdot,\cdot\rangle)$, un espace vectoriel sur le corps $\mathbb{C}$ (resp. $\mathbb{R}$) muni d'un produit hermitien (resp. scalaire). ⌟

Remarques
Le produit hermitien est parfois aussi appelé produit scalaire.

On déduit immédiatement des conditions n^os^ 1 et 2 que $\forall\,(\lambda, x, y) \in \mathbb{K} \times E^2$, $\langle\lambda\, x, y\rangle = \overline{\lambda}\,\langle x, y\rangle$, et que $\forall\,(x, y, z) \in E^3$, $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$.

$\langle x, x\rangle$ est un réel (positif), même si $\mathbb{K} = \mathbb{C}$.

Si $\mathbb{K} = \mathbb{R}$, la condition n° 1 se réduit à $\langle y, x\rangle = \langle x, y\rangle$. ⌟

Théorème 77 (inégalité de (Cauchy-)Schwarz)
Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien.

$\forall\,(x, y) \in E^2$,

$$|\langle x, y\rangle| \leqslant \|x\|_2\, \|y\|_2.$$ ⌟

Preuve
Pour tout $\mu \in \mathbb{C}$,

$$0 \leqslant \|x + \mu\, y\|_2^2 = \langle x + \mu\, y, x + \mu\, y\rangle = \langle x, x\rangle + \overline{\mu}\,\langle y, x\rangle + \mu\,\langle x, y\rangle + \overline{\mu}\,\mu\,\langle y, y\rangle = \|x\|_2^2 + \overline{\mu}\,\overline{\langle x, y\rangle} + \mu\,\langle x, y\rangle + |\mu|^2\,\|y\|_2^2. \quad (15)$$

Soit $\phi = \arg\langle x, y\rangle$. On a $\langle x, y\rangle = |\langle x, y\rangle|\,\mathrm{e}^{\mathrm{i}\,\phi}$ et $\overline{\langle x, y\rangle} = |\langle x, y\rangle|\,\mathrm{e}^{-\mathrm{i}\,\phi}$. L'équation (15) est vraie pour tout $\mu \in \mathbb{C}$, en particulier pour tout μ de la forme $\mu = \lambda\,\mathrm{e}^{-\mathrm{i}\,\phi}$, où λ est un réel quelconque. Pour tout $\lambda \in \mathbb{R}$, on a donc

$$0 \leqslant \|x\|_2^2 + (\overline{\lambda\,\mathrm{e}^{-\mathrm{i}\,\phi}})\cdot(\overline{|\langle x, y\rangle|\,\mathrm{e}^{\mathrm{i}\,\phi}}) + (\lambda\,\mathrm{e}^{-\mathrm{i}\,\phi})\cdot(|\langle x, y\rangle|\,\mathrm{e}^{\mathrm{i}\,\phi}) + \lambda^2\,\|y\|_2^2 = \|x\|_2^2 + 2\,\lambda\,|\langle x, y\rangle| + \lambda^2\,\|y\|_2^2.$$

Posons $a = \|y\|_2^2$, $b = 2\,|\langle x, y\rangle|$ et $c = \|x\|_2^2$. a, b et c sont des réels et a est positif. Pour que le trinôme $P(\lambda) := a\,\lambda^2 + b\,\lambda + c$ soit positif ou nul pour tout λ réel, il faut que ses racines λ_1 et λ_2 soient complexes ou que P ait une racine double, c.-à-d. que le discriminant $\Delta = b^2 - 4\,a\,c$ soit négatif ou nul.

En effet, si $\Delta > 0$, $P(\lambda)$ admet deux racines réelles distinctes, λ_1 et λ_2 ($> \lambda_1$). Comme $a \geqslant 0$ et que $P(\lambda) = a\cdot(\lambda - \lambda_1)\cdot(\lambda - \lambda_2)$, $P(\lambda) < 0$ pour tout $\lambda \in\,]\lambda_1, \lambda_2[$.

On a donc $\Delta \leqslant 0$, soit

$$(2\,|\langle x, y\rangle|)^2 \leqslant 4\,\|y\|_2^2\,\|x\|_2^2,$$

d'où l'inégalité de Cauchy-Schwarz. ⌟

Théorème-définition 1 (norme hermitienne, norme euclidienne)
Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien.

L'application $\|\cdot\|_2$ est une norme sur E. On l'appelle *norme hermitienne* si $\mathbb{K} = \mathbb{C}$ et *norme euclidienne* si $\mathbb{K} = \mathbb{R}$. ⌟

Preuve
D'après la définition du produit scalaire, $\|x\|_2 \geqslant 0$ pour tout $x \in E$, et $\|x\|_2 = 0 \Longrightarrow x = 0$.

Par ailleurs,

$$\|\lambda\, x\|_2 = \sqrt{\langle\lambda\, x, \lambda\, x\rangle} = \sqrt{\overline{\lambda}\,\lambda\,\langle x, x\rangle} = \sqrt{|\lambda|^2\,\langle x, x\rangle} = |\lambda|\,\sqrt{\langle x, x\rangle} = |\lambda|\,\|x\|_2.$$

Il reste à démontrer que $\|x + y\|_2 \leqslant \|x\|_2 + \|y\|_2$. Comme

$$\|x + y\|_2^2 = \langle x + y, x + y\rangle = \langle x, x\rangle + \langle y, x\rangle + \langle x, y\rangle + \langle y, y\rangle \leqslant \|x\|_2^2 + |\langle y, x\rangle| + |\langle x, y\rangle| + \|y\|_2^2$$

et que $|\langle y, x\rangle| = |\langle x, y\rangle| \leqslant \|x\|_2 \, \|y\|_2$, on obtient

$$\|x + y\|_2^2 \leqslant \|x\|_2^2 + 2\, \|x\|_2\, \|y\|_2 + \|y\|_2^2 = (\|x\|_2 + \|y\|_2)^2,$$

soit

$$\|x + y\|_2 \leqslant \|x\|_2 + \|y\|_2.$$

Si une application possède toutes les propriétés d'une norme sauf ($\|x\| = 0 \Longrightarrow x = 0$), on dit que c'est une *semi-norme*.

À toute norme sur un espace vectoriel E correspond une *distance*, $d(x, y) := \|x - y\|$ pour tous $(x, y) \in E^2$, (l'inverse n'est pas vrai), c.-à-d. une application de E^2 dans $\mathbb{R}^+$ possédant les propriétés suivantes :

- $d(x, y) = 0 \Longleftrightarrow x = y$;
- $\forall\, (x, y) \in E^2, d(x, y) = d(y, x)$;
- $\forall\, (x, y, z) \in E^3, d(x, z) \leqslant d(x, y) + d(y, z)$.

Définition 40 (distance hermitienne, distance euclidienne)
L'application

$$\begin{aligned} d_2 :\quad E^2 &\longrightarrow \mathbb{R}^+ \\ (x, y) &\longmapsto \|x - y\|_2 \end{aligned}$$

porte le nom de *distance hermitienne* (resp. *euclidienne*) si E est préhilbertien complexe (resp. réel).

Définition 41
Soit $(E, \langle\cdot, \cdot\rangle)$ un espace préhilbertien réel.

On appelle angle (non orienté) entre x et y l'unique $\theta \in [0, \pi]$ défini par

$$\theta = \arccos \frac{\langle x, y\rangle}{\|x\|_2\, \|y\|_2}.$$

Définition 42
Un vecteur u est dit normé ou unitaire si $\|u\|_2 = 1$.

Deux vecteurs u et v sont dits orthogonaux, ce que l'on note $u \perp v$, si $\langle u, v\rangle = 0$.

Deux sous-ensembles F et G de E sont dits orthogonaux, ce que l'on note $F \perp G$, si $\forall\, (u, v) \in F \times G, u \perp v$.

Soit I un ensemble dénombrable. Une famille de vecteurs $(u_i)_{i\in I}$ est dite orthogonale si $\forall\, (i, j) \in I^2, i \neq j \Longrightarrow u_i \perp u_j$. Elle est dite orthonormée si tous ses éléments sont en plus unitaires.

Remarque
Un ensemble dénombrable est un ensemble qu'on peut mettre en bijection avec $\mathbb{N}$ ou une partie de $\mathbb{N}$. Exemples d'ensembles dénombrables : $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, tout sous-ensemble de ces ensembles, tout ensemble fini. En revanche, $\mathbb{R}$ n'est pas dénombrable, de même que $\mathbb{C}$, $\mathbb{R}^n$ ou tout intervalle (réel) non réduit à un point.

Théorème 78 (Pythagore)
Soient x et y deux vecteurs. Si $x \perp y$, $\|x + y\|_2^2 = \|x\|_2^2 + \|y\|_2^2$.

Théorème 79
Toute famille orthogonale et dénombrable de vecteurs non nuls est libre.

Preuve
Soit $(u_i)_{i\in I}$ une famille de vecteurs orthogonaux non nuls.

La famille $(u_i)_{i\in I}$ est libre si, pour toute sous-famille *finie* J de I et pour toute famille $(\lambda_i)_{i\in J}$ de scalaires,

$$\sum_{i\in J} \lambda_i\, u_i = 0 \Longrightarrow \forall\, i \in J, \lambda_i = 0.$$

Supposons que $\sum_{i\in J} \lambda_i\, u_i = 0$. Pour tout $j \in J$, on a

$$\langle u_j, \sum_{i\in J} \lambda_i\, u_i\rangle = 0 = \sum_{i\in J} \lambda_i\, \langle u_j, u_i\rangle = \lambda_j\, \langle u_j, u_j\rangle = \lambda_j\, \|u_j\|_2^2.$$

$u_j \neq 0$, donc $\|u_j\|_2 \neq 0$ et $\lambda_j = 0$ pour tout $j \in J$: la famille est bien libre.

Annexe VII.2. Espaces hermitiens et euclidiens

Définition 43
On appelle espace hermitien (resp. euclidien) un espace vectoriel de dimension finie muni d'un produit hermitien (resp. scalaire). ⌟

Théorème 80
Soit E un espace hermitien ou euclidien de dimension n.

Toute famille orthonormée de n vecteurs constitue une base orthonormée de E. ⌟

Preuve
Dans un espace vectoriel E de dimension n, toute famille libre de n vecteurs constitue une famille génératrice de E, donc une base de E. Une famille orthonormée de n vecteurs étant libre, elle constitue une base orthonormée de E. ⌟

Partant d'une base quelconque de E, on peut toujours construire une base orthonormée par la procédure dite de Gram-Schmidt.

Théorème 81
Soient $(E, \langle\cdot,\cdot\rangle)$ un espace hermitien ou euclidien de dimension n et $B = (u_i)_{i\in[\![1,n]\!]}$ une base orthonormée de E.

Pour tout $x \in E$, on a

$$x = \sum_{i=1}^{n} \langle u_i, x\rangle\, u_i. \tag{16}$$

$x_i := \langle u_i, x\rangle$ est la i-ième coordonnée de x dans la base B.

Pour tous $(x, y) \in E^2$, on a

$$\langle x, y\rangle = \sum_{i=1}^{n} \overline{x_i}\, y_i, \tag{17}$$

où $y_i = \langle u_i, y\rangle$.

Enfin,

$$\|x\|_2^2 = \sum_{i=1}^{n} |x_i|^2. \tag{18}$$

⌟

Théorème-définition 2
Soient $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien (éventuellement de dimension infinie), P un sous-espace vectoriel de E de dimension $n \in \mathbb{N}$, et x un élément de E.

Il existe un et un seul $x' \in P$ tel que $x - x'$ soit orthogonal à tout vecteur de P. Les coordonnées x'_i de x' dans une base orthonormée (u_i) de P valent $x'_i = \langle u_i, x\rangle$. L'application

$$\begin{aligned} p\colon E &\longrightarrow P \\ x &\longmapsto x' = \textstyle\sum_{i=1}^{n} \langle u_i, x\rangle\, u_i \end{aligned}$$

est un projecteur, c.-à-d. une application linéaire telle que $p \circ p = p$. Le vecteur $p(x)$ est appelé projection orthogonale de x sur P et l'on a

$$x = \underbrace{p(x)}_{\in P} + \underbrace{(x - p(x))}_{\perp P}.$$

L'élément de P le plus proche de E au sens de la norme $\|\cdot\|_2$ est la projection orthogonale de x sur P. On a $\|p(x)\|_2^2 \leqslant \|x\|_2^2$. ⌟

Preuve
Soit $(u_i)_{i\in[\![1,n]\!]}$ une base orthonormée de P et $x' = \sum_{i=1}^{n} x'_i\, u_i$ un vecteur de P. Le vecteur x' est la projection orthogonale de x sur P si $x - x'$ est perpendiculaire à tout vecteur de P, ou, ce qui revient au même, si $\forall\, j \in [\![1, n]\!]$, $\langle u_j, x - x'\rangle = 0$.

$$\langle u_j, x - x'\rangle = \langle u_j, x - \sum_{i=1}^{n} x'_i\, u_i\rangle = \langle u_j, x\rangle - \sum_{i=1}^{n} x'_i\, \langle u_j, u_i\rangle = \langle u_j, x\rangle - x'_j.$$

La projection orthogonale de x sur P est donc le vecteur $x' = \sum_{i=1}^{n} \langle u_i, x\rangle\, u_i$; les coordonnées de x' dans la base $(u_i)_{i\in[\![1,n]\!]}$ sont donc les scalaires $\langle u_i, x\rangle$.

Montrons que x' est le vecteur de P le plus proche de x. Soit y un autre vecteur de P. On a

$$\begin{aligned}\|y-x\|_2^2 &= \langle (y-x')+(x'-x), (y-x')+(x'-x)\rangle \\ &= \langle y-x', y-x'\rangle + \langle x'-x, y-x'\rangle + \langle y-x', x'-x\rangle + \langle x'-x, x'-x\rangle \\ &= \|y-x'\|_2^2 + \|x'-x\|_2^2 \geqslant \|x'-x\|_2^2,\end{aligned}$$

puisque $x'-y \in P$ et $x-x' \perp P$. ⌟

On retrouve ces résultats en géométrie dans l'espace. Notons $\mathcal{E}$ l'espace affine usuel à 3 dimensions, E l'espace vectoriel associé, $\mathcal{P}$ un plan ou une droite affine de $\mathcal{E}$, P l'espace vectoriel associé à $\mathcal{P}$, p la projection orthogonale sur P et A un point quelconque de $\mathcal{P}$: pour tout point M de $\mathcal{E}$, le point M' défini par $\overrightarrow{AM'} = p(\overrightarrow{AM})$ s'appelle la projection orthogonale de M sur $\mathcal{P}$ (M' ne dépend pas du point A de $\mathcal{P}$ choisi) ; c'est le point de $\mathcal{P}$ le plus proche de M.

Annexe VII.3. Espaces de Hilbert

Les relations (16), (17) et (18) ont été établies dans un espace hermitien ou euclidien, donc de dimension finie ; de même, les propriétés de la projection orthogonale n'ont été démontrées que pour une projection sur un sous-espace vectoriel de dimension finie.

Les espaces de Hilbert permettent, sous certaines conditions, d'étendre ces résultats au cas de la dimension infinie.

Définition 44
On appelle *espace de Hilbert* un espace préhilbertien complet pour la norme $\|\cdot\|_2$. ⌟

Remarque. — Si l'espace est de dimension finie et est complet pour une norme, il est complet pour toute autre norme.

Si $(u_i)_{i\in\mathbb{N}}$ est une base (algébrique) de E, on souhaiterait pouvoir écrire que, pour tout $x \in E$, il existe une et une seule famille de scalaires $(\lambda_i)_{i\in\mathbb{N}}$ telle que $x = \sum_{i\in\mathbb{N}} \lambda_i\, u_i$. Cette écriture est problématique, car même si $(u_i)_{i\in\mathbb{N}}$ est une famille libre, pour qu'elle soit une base, il faut qu'elle soit génératrice, c.-à-d. que pour tout $x \in E$, il existe une sous-famille *finie* J telle que $x = \sum_{i\in J} \lambda_i\, u_i$. Par ailleurs, la somme étant infinie, $\sum_{i\in\mathbb{N}} \lambda_i\, u_i$ dépend a priori de l'ordre des u_i, lequel est arbitraire.

Pour pouvoir écrire que $x = \sum_{i\in\mathbb{N}} \lambda_i\, u_i$, il faut introduire le concept de base hilbertienne.

Définition 45
Soit E un espace de Hilbert de dimension infinie et $(u_i)_{i\in\mathbb{N}}$ une famille libre et dénombrable de vecteurs de E
La famille $(u_i)_{i\in\mathbb{N}}$ est une *base hilbertienne* de E si et seulement si

$$\forall\, x \in E, \forall\, \epsilon > 0, \text{ il existe un ensemble fini } J \text{ et une famille } (\lambda_i)_{i\in J} \text{ de scalaires tels que } \|x - \sum_{i\in J} \lambda_i\, u_i\|_2 \leqslant \epsilon.$$

⌟

Théorème 82
Soit E un espace de Hilbert de dimension infinie et $(u_i)_{i\in\mathbb{N}}$ est une base hilbertienne orthonormée de E.
Pour tous $(x, y) \in E^2$, on a alors

$$x = \sum_{i=0}^{\infty} \langle u_i, x\rangle\, u_i,$$

$$\|x\|_2^2 = \sum_{i=0}^{\infty} |\langle u_i, x\rangle|^2$$

et

$$\langle x, y\rangle = \sum_{i=0}^{\infty} \overline{\langle u_i, x\rangle}\, \langle u_i, y\rangle.$$

⌟

Chapitre VIII

É ´

VIII. . Généralités sur les équations différentielles

VIII. .1. Définitions

On appelle *équation différentielle d'ordre n* une équation reliant une fonction inconnue $y(x)$ et les n premières de ses dérivées, avec éventuellement une dépendance explicite dans la variable x, donc de la forme générale

$$F(y(x), y'(x), \ldots, y^{(n)};\ x) = 0. \tag{1}$$

Une *solution* de cette équation est une fonction $y(x)$ telle que (1) soit satisfaite pour tout x (dans un ensemble de définition donné). Le problème consiste à trouver *la* ou *les* solutions d'une telle équation, éventuellement complétée par des informations supplémentaires, comme on va voir.

Un cas particulier de la forme (1) est celui qui s'écrit

$$y^{(n)} = f(y, y', \ldots, y^{(n-1)};\ x), \tag{2}$$

c.-à-d. où $y^{(n)}$ a été explicitement exprimée en fonction de y et de ses $n-1$ premières dérivées. La forme (2), bien que moins générale, facilite souvent la discussion de l'existence des solutions et leur construction.

La variable x peut représenter le temps : l'équation régit alors la variation temporelle d'une variable dynamique y (position, température, pression, courant ou charge électrique, nombre de particules, etc.). Mais x peut aussi représenter une variable spatiale ; c'est par exemple le cas des ondes stationnaires dans une corde vibrante : si y désigne le déplacement transverse d'une corde tendue entre deux points d'abscisses $x = 0$ et $x = L$, et vibrant dans le mode de longueur d'onde λ, $y(x)$ obéit à l'équation $y'' + (2\,\pi/\lambda)^2\, y = 0$ pour $0 \leqslant x \leqslant L$.

VIII. .2. Systèmes d'équations différentielles

Plutôt qu'une équation différentielle sur une fonction inconnue, on peut aussi avoir à traiter des équations différentielles couplées, formant un *système d'équations différentielles*. C'est par exemple le cas pour

$$y_1' = f_1(y_1, y_2, \ldots, y_n;\ x), \tag{3}$$
$$y_2' = f_2(y_1, y_2, \ldots, y_n;\ x), \tag{4}$$
$$\vdots \tag{5}$$
$$y_n' = f_n(y_1, y_2, \ldots, y_n;\ x), \tag{6}$$

qui est un système de n équations du premier ordre.

Il est bon d'observer que toute équation différentielle du n-ième ordre de la forme (2) peut se mettre sous la forme d'un système de n équations du premier ordre. Il suffit de définir $y_1 \coloneqq y$, $y_2 \coloneqq y'$, …, $y_n \coloneqq y^{(n-1)}$. On vérifie aisément (par récurrence par exemple) que $y_1' = y_2$, $y_2' = y_3$, …, $y_{n-1}' = y_n$, $y_n' = f(y_1, y_2, \ldots, y_n;\ x)$, qui est bien un système de la forme (3).

Un exemple familier qui conduit à un système d'équations différentielles couplées est fourni par le circuit électrique LC constitué d'une bobine d'inductance L et d'un condensateur de capacité C (figure VIII.1a). Rappelons brièvement l'analyse de ce système. La tension aux bornes du condensateur portant une charge Q est $V_C = Q/C$ tandis que celle aux bornes de la bobine est $V_L = L\,\mathrm{d}I/\mathrm{d}t$, où $I = \mathrm{d}Q/\mathrm{d}t$ est l'intensité du courant dans le circuit. En écrivant que $V_C + V_L = 0$, on trouve le système différentiel

$$L\,\frac{\mathrm{d}I}{\mathrm{d}t} + \frac{Q}{C} = 0, \tag{7}$$

$$I - \frac{\mathrm{d}Q}{\mathrm{d}t} = 0. \tag{8}$$

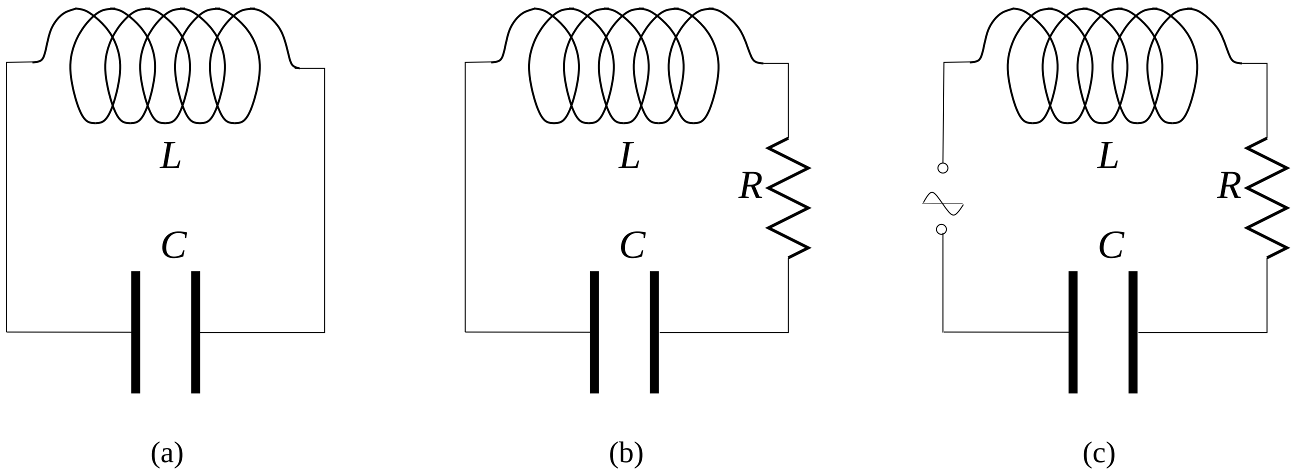


Figure VIII.1 – Circuits *LC* et *RLC*, isolés ou couplés à une tension extérieure.

En éliminant I entre ces deux équations, on obtient pour Q une équation du second ordre :

$$\frac{d^2Q}{dt^2} + \frac{Q}{L\,C} = 0. \tag{9}$$

Quelle équation décrit la situation de la figure VIII.1b, où la bobine a aussi une résistance interne ? Et celle de la figure VIII.1c, où le circuit est couplé à une tension extérieure V_e ?

VIII. . Conditions initiales. Théorème de Cauchy. Conditions aux limites

VIII. .1. Existence et unicité des solutions

Considérons un problème de mécanique élémentaire, celui du mouvement d'un point matériel soumis à la seule gravité. Sa coordonnée $z(t)$ verticale le long d'un axe orienté vers le haut satisfait l'équation

$$\ddot{z} = -g, \tag{10}$$

comme conséquence du principe fondamental de la dynamique. Cette équation différentielle se résout par deux *quadratures* (c.-à-d. deux intégrations) successives ; on a d'abord

$$\dot{z}(t) - \dot{z}(0) = -g \int_{t'=0}^{t} dt' \Longrightarrow \dot{z}(t) = -g\,t + v_0$$

où $v_0 := \dot{z}(0)$ est la vitesse à l'instant $t = 0$. Puis une nouvelle quadrature fournit

$$z(t) - z(0) = \int_{t'=0}^{t} (-g\,t' + v_0)\,dt' \Longrightarrow z(t) = -\frac{1}{2}\,g\,t^2 + v_0\,t + z_0, \tag{11}$$

où $z_0 := z(0)$ est la position à l'instant $t = 0$. On trouve donc que la solution est complètement déterminée par la donnée de deux constantes, les *conditions initiales*.

Le principe selon lequel la solution d'une équation différentielle *existe et est unique* si on se donne des conditions initiales appropriées est en fait très général.

On admettra le « théorème » [(*1)] suivant.

Théorème 83 (« théorème » de Cauchy)
Si les valeurs de y et de ses $n-1$ premières dérivées en un point x_0 sont données (notons-les $y(x_0) = a_0$, $y'(x_0) = a_1, \ldots, y^{(n-1)}(x_0) = a_{n-1}$) et si la fonction $f(y_0, y_1, y_2, \ldots, y_{n-1};\ x)$ est « suffisamment régulière » par rapport aux variables $y_0, y_1, y_2, \ldots, y_{n-1}, x$ dans un « intervalle » ouvert contenant le point $(a_0, a_1, \ldots, a_{n-1};\ x_0)$, alors l'équation

$$y^{(n)} = f(y, y', \ldots, y^{(n-1)};\ x), \tag{12}$$

admet une et une seule solution. ⌟

1. Entre guillemets en raison de la formulation excessivement vague que nous en donnons. Des notions de topologie et sur les fonctions de plusieurs variables seraient nécessaires pour en donner un énoncé rigoureux.

Les valeurs des $y(x_0)$, $y'(x_0)$, …, $y^{(n-1)}(x_0)$ sont appelées *conditions de Cauchy* (ou conditions initiales). Nous ne préciserons pas ici les conditions exactes de ce théorème. Dans tous les exemples rencontrés dans ce cours, ces conditions seront supposées satisfaites, et, moyennant la donnée des conditions initiales, la solution existera et sera unique.

Cette propriété n'a rien d'évident et on peut construire des exemples d'équations dont la solution n'est pas unique, à conditions initiales fixées. Considérons ainsi l'équation

$$y' = 3\, y^{2/3},$$

avec la condition initiale $y(0) = 0$. On vérifie immédiatement qu'elle admet *deux* solutions, $y(x) = 0$ et $y(x) = x^3$. C'est la singularité de la fonction $y^{2/3}$ (sa non-dérivabilité par rapport à y en $y = 0$) qui est responsable de cette non-unicité.

Noter que ce théorème d'existence et d'unicité ne nous donne pas d'information sur la solution, ni sur la méthode pour la déterminer. De fait, la résolution effective d'une équation différentielle est en général un problème difficile, qui n'admet que rarement une solution explicite en termes de fonctions simples. Il faut bien souvent recourir à des approximations numériques ou analytiques (développement en approximations successives). Dans la suite, nous allons nous borner à considérer des cas suffisamment simples pour être complètement solubles et en même temps suffisamment intéressants pour la physique.

VIII. .2. Conditions initiales et conditions aux bords

On a supposé données une ou des conditions initiales sur la fonction cherchée $y(x)$ pour une valeur x_0 donnée. En physique, c'est typiquement le type de données dont on dispose dans un problème de dynamique : par exemple, la vitesse et le position du mobile (deux vecteurs à 3 dimensions s'il s'agit d'un point matériel dans l'espace ordinaire) sont connues à l'« instant initial » ; ou encore, la charge portée par le condensateur et l'intensité circulant dans un circuit *RLC* sont connues au temps 0, etc. Mais il existe d'autres types de problèmes où la fonction cherchée est contrainte pour deux valeurs différentes de la variable x. *Le théorème de Cauchy n'est alors pas applicable*.

Exemple 66
Reprenons l'étude du point matériel soumis à la gravité et obéissant à (10). Quelle est sa trajectoire passant par l'origine $z_0 = 0$ au temps $t_0 = 0$ et par la position z_1 au temps t_1 ? Réponse : la loi du mouvement (11) dépend des deux constantes v_0 et z_0 qui sont déterminées par les deux conditions ci-dessus : $z_0 = 0$ et v_0 solution de $-\frac{1}{2}\, g\, t_1^2 + v_0\, t_1 = z_1$. La résolution de cette dernière équation nous indique avec quelle vitesse initiale v_0 le corps doit être mis en mouvement à l'instant $t = 0$ pour atteindre le point z_1 à l'instant requis. La connaissance de z_0 et v_0 détermine complètement la solution. La donnée des *conditions aux limites* $z(0) = z_0$, $z(t_1) = z_1$ a donc suffi ici à déterminer de façon unique la solution de l'équation différentielle. ⌟

Exemple 67
La donnée de conditions aux limites ne conduit pas toujours à un problème bien posé, admettant une solution et une seule. Examinons par exemple le cas de la corde vibrante obéissant à $y'' + k^2\, y = 0$. La corde étant supposée attachée à ses deux extrémités d'abscisses $x = 0$ et $x = L$, son déplacement transverse s'annule en ces deux points. On cherche donc des solutions de l'équation différentielle satisfaisant $y(0) = y(L) = 0$. Puisque la solution la plus générale de l'équation s'écrit

$$y(x) = A \cos(k\, x) + B \sin(k\, x),$$

la condition $y(0) = 0$ impose que $A = 0$, d'où l'on déduit, grâce à la condition $y(L) = 0$, que $B \sin(k\, L) = 0$. Si on écarte la solution triviale $B = 0$ où la corde reste au repos, la deuxième condition impose donc que $k\, L = n\, \pi$, ce qui contraint un coefficient de l'équation différentielle, mais ne détermine pas le coefficient arbitraire restant, B. Physiquement, le « mode » dans lequel la corde peut vibrer (la longueur d'onde de ses oscillations) n'est pas arbitraire, mais est (partiellement) contraint par la géométrie (la longueur de la corde). Dans ce cas, la donnée de deux conditions aux bords ne suffit pas à déterminer uniquement la solution : une condition supplémentaire, par exemple, l'amplitude maximale de la corde est nécessaire pour fixer le coefficient B. ⌟

Exemple 68
Dans une autre classe de problèmes physiques, les conditions aux limites sont remplacées par des *conditions asymptotiques*. C'est le cas de l'équation

$$y'' = k^2\, y \qquad (k > 0),$$

dont la solution la plus générale est

$$y(x) = A\, e^{k\, x} + B\, e^{-k\, x}.$$

Supposons qu'on s'intéresse au comportement à grand $x > 0$ de $y(x)$ et que telle ou telle contrainte physique (ou tout simplement le bon sens !) impose à y de rester bornée (c.-à-d. de ne pas tendre vers l'infini) quand x croît. (Nous rencontrerons cette situation dans l'étude de la propagation de la chaleur, cf. TD 5.) Cette condition asymptotique implique que $A = 0$; la solution est donc de la forme $y(x) = B\,\mathrm{e}^{-k\,x}$, ce qu'on peut encore écrire $y(x) = y_0\,\mathrm{e}^{-k\,x}$, où y_0 est la valeur de y en $x = 0$. Dans ce cas, la condition asymptotique plus la condition initiale ou au bord $y(0) = y_0$ déterminent complètement et uniquement la solution. ⌟

VIII. . Équations différentielles linéaires

VIII. .1. Propriétés générales

Nous allons nous intéresser tout spécialement à des équations différentielles du n-ième ordre du type suivant

$$\sum_{k=0}^{n} C_k(x)\, y^{(k)}(x) = \phi(x), \tag{13}$$

ce que nous réécrivons encore sous la forme

$$\mathcal{L}(y[x]) := \left(\sum_{k=0}^{n} C_k[x]\,\frac{\mathrm{d}^k}{\mathrm{d}x^k}\right) y(x) = \phi(x), \tag{14}$$

où $\phi(x)$ est une fonction donnée, les $C_k(x)$ sont des fonctions données de x, et on cherche à déterminer la fonction inconnue $y(x)$, sujette à n conditions initiales, en $x = 0$ par exemple. La fonction ϕ représente la « source » (une action extérieure) qui « excite » le système représenté par y, et on cherche à calculer la « réponse » y à ϕ.

Le membre de gauche de l'équation ci-dessus a la propriété fondamentale d'être *linéaire* en y : y et ses dérivées n'y apparaissent qu'à la puissance 1. C'est ce que nous avons souligné par la notation $\mathcal{L}$ pour l'*opérateur différentiel* apparaissant au membre de gauche. La propriété de linéarité exprime que, pour toute paire de fonctions $(y_1[x], y_2[x])$ et pour tous coefficients constants λ_1 et λ_2, éventuellement complexes,

$$\forall\, x, \quad \mathcal{L}(\lambda_1\, y_1[x] + \lambda_2\, y_2[x]) = \lambda_1\, \mathcal{L}(y_1[x]) + \lambda_2\, \mathcal{L}(y_2[x]). \tag{15}$$

Cela a pour conséquence que, si les fonctions $y_1(x)$ et $y_2(x)$ sont solutions de (14) respectivement pour deux « sources » $\phi_1(x)$ et $\phi_2(x)$, c.-à-d. $\mathcal{L}(y_1) = \phi_1$, $\mathcal{L}(y_2) = \phi_2$, alors, quelles que soient les *constantes* λ_1 et λ_2, $\lambda_1\, y_1(x) + \lambda_2\, y_2(x)$ est solution de l'équation pour la source $\lambda_1\, \phi_1 + \lambda_2\, \phi_2$, autrement dit $\mathcal{L}(\lambda_1\, y_1[x] + \lambda_2\, y_2[x]) = \lambda_1\, \phi_1(x) + \lambda_2\, \phi_2(x)$. Nous utiliserons ce principe dans la résolution d'équations par décomposition en modes de Fourier (cf. §VIII.c.5.*b*.i).

Avant d'aller plus loin, il convient de souligner que les *systèmes linéaires*, c.-à-d. régis par des équations différentielles linéaires jouent un rôle fondamental en physique et dans toutes les sciences exactes. Nous leur consacrerons le chapitre IX de ce cours.

VIII. .2. Équations linéaires homogènes, équations linéaires inhomogènes

On appelle équation différentielle linéaire *homogène* une équation du type (14) dans laquelle $\phi(x) = 0$:

$$\mathcal{L}(y[x]) = \left(\sum_{k=0}^{n} C_k[x]\,\frac{\mathrm{d}^k}{\mathrm{d}x^k}\right) y(x) = 0. \tag{16}$$

Inversement, on appelle équation différentielle linéaire *inhomogène*, ou *avec second membre*, une équation du type (14) dans laquelle $\phi(x) \neq 0$.

VIII. .2.*a*. Indépendance linéaire de *n* fonctions

La propriété de linéarité implique qu'étant données deux solutions $y_1(x)$ et $y_2(x)$ de l'équation homogène, une combinaison linéaire $y(x) = \lambda_1\, y_1(x) + \lambda_2\, y_2(x)$ à coefficients constants est aussi solution. Bien sûr, choisir y_2 proportionnelle à y_1 est possible mais n'apporte aucune information nouvelle. On dira que les deux fonctions $y_1(x)$ et $y_2(x)$ sont *linéairement dépendantes* si on peut trouver deux constantes λ_1 et λ_2 non toutes deux nulles telles que, *pour tout* x, on ait $\lambda_1\, y_1(x) + \lambda_2\, y_2(x) = 0$. Si cette condition est réalisée, avec par exemple $\lambda_1 \neq 0$, on a

$$\forall\, x \quad \left(\lambda_1\, y_1[x] + \lambda_2\, y_2[x] = 0 \Rightarrow y_1[x] = -\frac{\lambda_2}{\lambda_1}\, y_2[x]\right) \tag{17}$$

et y_1 et y_2 sont proportionnelles. Inversement, on dit que deux fonctions $y_1(x)$ et $y_2(x)$ sont *linéairement indépendantes* si la condition $\lambda_1\, y_1(x) + \lambda_2\, y_2(x) = 0$ est impossible à réaliser avec deux constantes λ_1 et λ_2 non nulles.

La définition suivante généralise ces notions.

Définition 46 (indépendance linéaire)
n fonctions $y_1(x), \ldots, y_n(x)$ sont *linéairement indépendantes* (ou encore, « constituent une *famille libre* ») si on ne peut pas trouver n constantes $\lambda_1, \ldots, \lambda_n$ non toutes nulles telles que $\forall\, x,\ \sum_{i=1}^{n} \lambda_i\, y_i(x) = 0$ ou, de manière équivalente, si

$$(\forall\, x,\ \sum_{i=1}^{n} \lambda_i\, y_i(x) = 0) \Longrightarrow \lambda_1 = \lambda_2 = \ldots = \lambda_n = 0. \tag{18}$$

Dans le cas contraire, les fonctions sont *linéairement dépendantes* (ou encore, « constituent une *famille liée* »). ⌟

(Si les fonctions sont linéairement dépendantes on peut toujours exprimer l'une de ces fonctions comme combinaison linéaire des $n-1$ autres.) Par exemple, on démontre que les fonctions $\cos(\omega_i\, x)$ et $\sin(\omega_i\, x)$, pour tout ensemble de n nombres positifs ω_i distincts, sont linéairement indépendantes ; de même pour les fonctions $\exp(k_i\, x)$, pour tout ensemble de n nombres k_i réels ou complexes distincts, sont linéairement indépendantes.

Exercice
En appliquant le critère ci-dessus, démontrer que les fonctions $\mathrm{e}^{k\,x}$ et $x\,\mathrm{e}^{k\,x}$ sont linéairement indépendantes.
Les fonctions $\mathrm{e}^{\mathrm{i}\,k\,x}$, $\cos(k\,x)$ et $\sin(k\,x)$ sont-elles indépendantes ? (Ici, on admet dans (18) des coefficients λ_i complexes.) ⌟

Exercice
Montrer que deux fonctions dérivables y_1 et y_2 sont linéairement indépendantes si et seulement si leur *wronskien* $W(y_1, y_2) := y_1'\, y_2 - y_2'\, y_1$ n'est pas identiquement nul (c.-à-d. nul pour tout x).
Plus généralement, on définit le wronskien de n fonctions $y_i(x)$ ($i \in [\![1, n]\!]$) supposées $(n-1)$ fois dérivables comme le *déterminant* (cf. cours d'algèbre)

$$W(x) = \det\begin{pmatrix} y_1(x) & y_2(x) & \ldots & y_n(x) \\ y_1'(x) & y_2'(x) & \ldots & y_n'(x) \\ y_1''(x) & y_2''(x) & \ldots & y_n''(x) \\ \vdots & & & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \ldots & y_n^{(n-1)}(x) \end{pmatrix}. \tag{19}$$

⌟

Si les fonctions y_i sont linéairement dépendantes, il existe n constantes non toutes nulles λ_i telles que $\sum_{i=1}^{n} \lambda_i\, y_i(x) = 0$, et cette relation est aussi vérifiée par les dérivées successives des y_i, donc $\forall\, p \in [\![0, n-1]\!]$, $\sum_{i=1}^{n} \lambda_i\, y_i^{(p)}(x) = 0$. Les n colonnes du déterminant sont donc linéairement dépendantes et le déterminant $W(x)$ est nul pour tout x. Inversement, il suffit que le wronskien $W(x)$ soit non nul en un point x_0 pour qu'on puisse conclure que les fonctions y_i sont linéairement *indépendantes*. L'étude du wronskien offre donc un moyen pratique d'étudier l'indépendance linéaire de n fonctions.

VIII. .2.*b*. Théorèmes

Ces définitions étant acquises, on peut démontrer deux théorèmes fondamentaux sur les équations différentielles linéaires.

Théorème 84 (équations homogènes)
Si les fonctions C_k ($k = 0, \ldots, n$) sont continues dans un même intervalle I, avec $C_n(x)$ ne s'annulant pas dans cet intervalle, il existe n solutions linéairement indépendantes de l'équation différentielle homogène (16) dans I, et la solution la plus générale est une combinaison linéaire à coefficients constants arbitraires de ces n fonctions. En outre, il existe une unique solution de (16) satisfaisant à des conditions initiales $y(x_0) = a_0$, $y'(x_0) = a_1, \ldots,$ $y^{(n-1)}(x_0) = a_{n-1}$, avec $x_0 \in I$. ⌟

Les solutions de l'équation homogène constituent donc un espace vectoriel de dimension n dont les y_i sont une base.

Théorème 85 (équations inhomogènes)
Avec les mêmes hypothèses sur les fonctions C_k qu'au théorème 84, et ϕ étant supposée en outre continue dans I, la solution générale dans I de l'équation inhomogène (14) est la somme d'une solution particulière de (14) et de la solution générale de l'équation homogène (16). Il existe une unique solution de (14) satisfaisant à des conditions initiales $y(x_0) = a_0, \ldots, y^{(n-1)}(x_0) = a_{n-1}$, avec $x_0 \in I$. ⌟

L'existence de n solutions indépendantes de (16) selon le théorème 84 n'est pas aisée à démontrer et nous l'admettrons en général ; nous en donnerons une ébauche de preuve pour des coefficients constants au VIII.c.4.*a*.
Le théorème 85, lui, se démontre aisément : supposons que $\tilde{y}(x)$ soit une solution particulière de (14). Alors pour toute autre solution $y(x)$ de (14), en soustrayant membre à membre les deux équations différentielles satisfaites par $y(x)$ et par $\tilde{y}(x)$, on trouve que la différence $y(x) - \tilde{y}(x)$ est une solution de l'équation *homogène* : en invoquant alors le théorème 84, on montre que la solution générale y de (14) est la somme d'une solution

particulière $\tilde{y}(x)$ et d'une solution quelconque (« générale ») de (16). La réciproque est évidente : on obtient bien une solution de (14) en ajoutant à la solution particulière $\tilde{y}(x)$ une solution quelconque de l'équation homogène. Noter que, là encore, ces théorèmes énoncent l'existence et l'unicité de solutions sans nous dire comment les construire...

VIII. .3. Équations différentielles linéaires du premier ordre à coefficients variables

VIII. .3.*a*. Équation homogène

Pour toute équation linéaire homogène du premier ordre, écrite comme

$$y'(x) + q(x)\, y(x) = 0, \tag{20}$$

la solution générale s'obtient facilement par une quadrature, c.-à-d. à l'aide d'un calcul de primitive de $q(x)$ (*2) :

$$y(x) = A \exp\left(-\int^x q[x']\, \mathrm{d}x'\right). \tag{21}$$

VIII. .3.*b*. Équation inhomogène

Dans le cas d'une équation du premier ordre, il existe une méthode générale pour construire une solution particulière de l'équation inhomogène, la *méthode de variation de la constante*. Considérons une équation de la forme

$$y'(x) + q(x)\, y(x) = \phi(x). \tag{22}$$

L'équation *homogène* admet selon la discussion précédente la solution générale $y(x) = A \exp\left(-\int^x q[x']\, \mathrm{d}x'\right)$, où A est une *constante*. Cherchons une solution de l'équation *inhomogène* (22) de la forme

$$y(x) = A(x) \exp\left(-\int^x q[x']\, \mathrm{d}x'\right),$$

où $A(x)$ est désormais une *fonction*. La dérivée de y est

$$\begin{aligned} y'(x) &= A'(x) \exp\left(-\int^x q[x']\, \mathrm{d}x'\right) + A(x) \frac{\mathrm{d}}{\mathrm{d}x} \exp\left(-\int^x q[x']\, \mathrm{d}x'\right) \\ &= A'(x) \exp\left(-\int^x q[x']\, \mathrm{d}x'\right) + A(x) \frac{\mathrm{d}\left(-\int^x q[x']\, \mathrm{d}x'\right)}{\mathrm{d}x} \exp\left(-\int^x q[x']\, \mathrm{d}x'\right) \\ &= A'(x) \exp\left(-\int^x q[x']\, \mathrm{d}x'\right) - A(x)\, q(x) \exp\left(-\int^x q[x']\, \mathrm{d}x'\right). \end{aligned}$$

En reportant dans (22), on trouve

$$(A'[x] - A[x]\, q[x] + A[x]\, q[x]) \exp\left(-\int^x q[x']\, \mathrm{d}x'\right) = \phi(x),$$

soit

$$A'(x) = \exp\left(\int^x q[x']\, \mathrm{d}x'\right) \phi(x),$$

et le problème de trouver une solution particulière est donc réduit au calcul d'une primitive de la fonction $\exp\left(\int^x q[x']\, \mathrm{d}x'\right) \phi(x)$.

VIII. .4. Équations linéaires homogènes à coefficients constants

VIII. .4.*a*. Résolution de l'équation homogène

Considérons maintenant un cas particulièrement important d'équations différentielles linéaires homogènes, celui où les coefficients C_k de l'équation sont des constantes :

$$\sum_{k=0}^{n} C_k\, y^{(k)}(x) = 0. \tag{23}$$

2. La définition à une constante additive près de la primitive de $q(x)$ introduit-elle un arbitraire ?

Cherchons des solutions de forme exponentielle, $y(x) = e^{\alpha x}$, où α peut être *a priori* réel ou complexe. Une telle fonction est effectivement solution à condition que

$$\left(\sum_{k=0}^{n} C_k \, \alpha^k\right) e^{\alpha x} = 0,$$

et donc que α satisfasse l'équation polynomiale, dite *équation caractéristique*,

$$\sum_{k=0}^{n} C_k \, \alpha^k = 0. \tag{24}$$

Or le théorème fondamental de l'algèbre énonce que tout polynôme $P(\alpha) = \sum_{k=0}^{n} C_k \, \alpha^k$ de degré n admet n *racines*, réelles ou complexes, qui sont les n solutions de l'équation $P(\alpha) = 0$. Si nous supposons d'abord que ces n racines α_i ($i = 1, \ldots, n$) sont distinctes, nous avons trouvé n solutions $e^{\alpha_i x}$, et ces n fonctions sont bien linéairement indépendantes, selon une remarque faite plus haut. Donc, dans ce cas, nous savons construire explicitement une base de n solutions, dont toute autre solution est une combinaison linéaire.

Plus généralement, on peut démontrer le théorème suivant.

Théorème 86
Soit $P(\alpha) = \sum_{k=0}^{n} C_k \, \alpha^k$. Notons $\alpha_i \in \mathbb{C}$ ses p racines et $m_i \in \mathbb{N}$ leur multiplicité. (On a donc $P(\alpha) = C_n \prod_{i=1}^{p} (\alpha - \alpha_i)^{m_i}$ et $\sum_{i=1}^{p} m_i = n$.)

Les fonctions $x^j \, e^{\alpha_i x}$, où $j \in [\![0, m_i - 1]\!]$, constituent une base de l'espace des solutions de (23). ⌟

À titre d'exemple, montrons dans le cas où la racine α_i est double que $x \, e^{\alpha_i x}$ est également une solution de l'équation différentielle. On a

$$\frac{d}{d\alpha}\left(\sum_{k=0}^{n} C_k \frac{d^k e^{\alpha x}}{dx^k}\right) = \sum_{k=0}^{n} C_k \frac{d^k}{dx^k}\left(\frac{d(e^{\alpha x})}{d\alpha}\right) = \sum_{k=0}^{n} C_k \frac{d^k (x \, e^{\alpha x})}{dx^k}.$$

Or, si α_i est double, $P(\alpha) = (\alpha - \alpha_i)^2 \, Q(\alpha)$, où Q est un polynôme.

$$\frac{dP}{d\alpha} = 2\,(\alpha - \alpha_i)\, Q(\alpha) + (\alpha - \alpha_i)^2 \frac{dQ}{d\alpha},$$

donc $(dP/d\alpha)_{\alpha=\alpha_i} = 0$. Comme

$$\frac{d}{d\alpha}\left(\sum_{k=0}^{n} C_k \frac{d^k e^{\alpha x}}{dx^k}\right) = \frac{d(P \, e^{\alpha x})}{d\alpha} = \left(\frac{dP}{d\alpha} + x \, P[\alpha]\right) e^{\alpha x}$$

s'annule en $\alpha = \alpha_i$, on en déduit que $x \, e^{\alpha_i x}$ est bien une solution.

Ce résultat se généralise sans difficulté au cas où la multiplicité de α_i est $m_i > 2$ en utilisant la formule de Leibniz,

$$\frac{d^p (f \, g)}{dx^p} = \sum_{j=0}^{p} \binom{p}{j} \frac{d^j f}{dx^j} \frac{d^{p-j} g}{dx^{p-j}}.$$

VIII. .4.*b*. Équation du second ordre à coefficients constants

Examinons plus en détail le cas de l'équation du second ordre, que nous retrouverons au chapitre IX dans différents contextes physiques. Réécrivons-la sous la forme

$$a \, y''(x) + b \, y'(x) + c \, y(x) = 0, \tag{25}$$

où a, b, c et y sont réels et $a \neq 0$. Son équation caractéristique est

$$a \, \alpha^2 + b \, \alpha + c = 0. \tag{26}$$

La solution de (26) dépend du signe de $\Delta := b^2 - 4\,a\,c$:

- Si $\Delta > 0$, les racines $\alpha_\pm$ sont réelles et distinctes :

  $$\alpha_\pm = \frac{-b \pm \sqrt{\Delta}}{2\,a}$$

  et les deux fonctions $e^{\alpha_\pm x}$ forment une base de solutions, ce qui veut dire que la solution générale de (25) est

  $$y(x) = A_+ \, e^{\alpha_+ x} + A_- \, e^{\alpha_- x}, \tag{27}$$

  avec deux constantes arbitraires $A_\pm$ fixées par les conditions initiales.

- Si $\Delta < 0$, les racines sont complexes conjuguées et distinctes :

$$\alpha_\pm = \frac{-b \pm \mathrm{i}\,\sqrt{-\Delta}}{2\,a}.$$

Comme base de solutions, on peut choisir soit les deux exponentielles oscillantes $\mathrm{e}^{\alpha_\pm\,x}$, soit

$$\cos\left(\frac{\sqrt{-\Delta}}{2\,a}\,x\right)\exp\left(\frac{-b\,x}{2\,a}\right) \quad \text{et} \quad \sin\left(\frac{\sqrt{-\Delta}}{2\,a}\,x\right)\exp\left(\frac{-b\,x}{2\,a}\right).$$

La solution générale de l'équation (25) peut donc s'écrire sous différentes formes équivalentes :

$$y(x) = B_+\,\mathrm{e}^{\alpha_+\,x} + B_-\,\mathrm{e}^{\alpha_-\,x}, \tag{28}$$

$$\text{ou encore} \quad y(x) = \left(A_1\cos\left[\frac{\sqrt{-\Delta}}{2\,a}\,x\right] + A_2\sin\left[\frac{\sqrt{-\Delta}}{2\,a}\,x\right]\right)\exp\left(\frac{-b\,x}{2\,a}\right), \tag{29}$$

$$\text{ou encore} \quad y(x) = A\cos\left(\frac{\sqrt{-\Delta}}{2\,a}\,x + \varphi\right)\exp\left(\frac{-b\,x}{2\,a}\right) \tag{30}$$

avec des paires de constantes (B_+, B_-), (A_1, A_2) ou (A, φ). Sous la première forme, B_+ et B_- doivent être des complexes conjugués pour que $y(x)$ soit réel : $B_+ = B_-^*$.

- Si $\Delta = 0$, la racine $\alpha_0 = -b/(2\,a)$ est double et on vérifie immédiatement que

$$\mathrm{e}^{\alpha_0\,x} \quad \text{et} \quad x\,\mathrm{e}^{\alpha_0\,x}$$

sont deux solutions, donc que la solution générale de l'équation (25) s'écrit

$$y(x) = (A_0\,x + B_0)\,\mathrm{e}^{\alpha_0\,x} \tag{31}$$

avec deux constantes A_0 et B_0.

Dans tous les cas, on a trouvé une solution générale dépendant de *deux* constantes réelles (puisque $\mathrm{Re}\,B_+ = \mathrm{Re}\,B_-$ et $\mathrm{Im}\,B_+ = -\,\mathrm{Im}\,B_-$ dans la première expression du cas $\Delta < 0$). Ces deux constantes sont déterminées de façon unique par *deux* conditions initiales, par exemple la donnée de $y(x_0)$ et $y'(x_0)$.

VIII. .5. Équations linéaires inhomogènes à coefficients constants

Pour une équation linéaire à coefficients constants, on sait donc de façon générale construire les n solutions indépendantes du théorème 84 et résoudre complètement l'équation homogène. En ce qui concerne l'équation inhomogène à coefficients constants, sa solution repose sur la connaissance d'*une* solution particulière, selon le théorème 85.

VIII. .5.*a*. Second membre constant

Un cas évident est celui où $\phi(x)$ est constante. Dans ce cas, $y(x) = \phi/C_0$ est une solution particulière.

VIII. .5.*b*. Second membre périodique

VIII. .5.*b*.i. Découplage des modes de Fourier

On considère l'équation différentielle du ℓ-ième ordre à coefficients C_k constants

$$\mathcal{L}(y[x]) := \sum_{k=0}^{\ell} C_k\,\frac{\mathrm{d}^k}{\mathrm{d}x^k}\,y(x) = \phi(x), \tag{32}$$

où $\phi(x)$ est une fonction périodique de x, de période T, supposée satisfaire les hypothèses du théorème de Dirichlet. On peut alors représenter ϕ par son développement de Fourier,

$$\phi(x) = \sum_{n=-\infty}^{\infty} \phi_n\,\mathrm{e}^{\mathrm{i}\,\omega\,n\,x}, \tag{33}$$

où $\omega = 2\,\pi/T$. Cherchons une solution particulière $y(x)$ T-périodique et supposons-la également développable en série de Fourier :

$$y(x) = \sum_{n=-\infty}^{\infty} y_n\,\mathrm{e}^{\mathrm{i}\,n\,\omega\,x}. \tag{34}$$

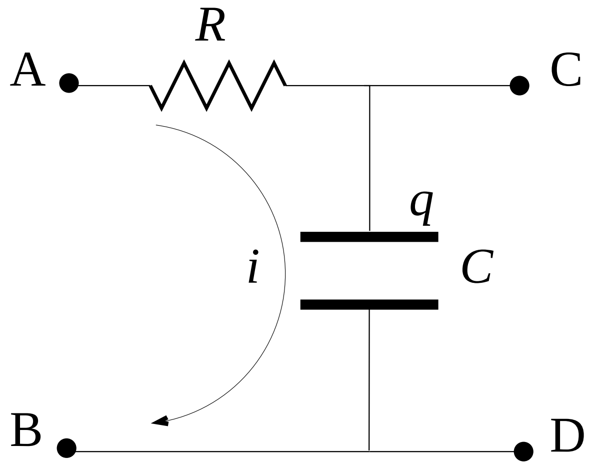


Figure VIII.2 – Filtre *RC*.

Si l'on suppose $y(x)$ suffisamment régulière (cf. VII.E.1), on peut dériver terme à terme le développement de $y(x)$, et comme dériver k fois revient à multiplier le n-ième « mode » (la n-ième composante de Fourier) par $(\mathrm{i}\,n\,\omega)^k$,

$$y^{(k)}(x) = \sum_{n=-\infty}^{\infty} y_n \cdot (\mathrm{i}\,n\,\omega)^k\, \mathrm{e}^{\mathrm{i}\,n\,\omega\,x}.$$

La propriété de linéarité implique que l'action de $\mathcal{L}$ sur les différents modes de $y(x)$ s'écrit

$$\mathcal{L}(y[x]) = \sum_n \sum_{k=0}^{\ell} C_k \cdot (\mathrm{i}\,n\,\omega)^k\, y_n\, \mathrm{e}^{\mathrm{i}\,n\,\omega\,x} \tag{35}$$

et qu'on obtient *une solution particulière* de (32) en sommant une solution pour chacun des modes,

$$\sum_{k=0}^{\ell} C_k \cdot (\mathrm{i}\,n\,\omega)^k\, y_n = \phi_n,$$

soit

$$y_n = \frac{\phi_n}{\sum_{k=0}^{\ell} C_k \cdot (\mathrm{i}\,n\,\omega)^k} \quad \text{et} \quad y(x) = \sum_{-\infty}^{\infty} \frac{\phi_n}{\sum_{k=0}^{\ell} C_k \cdot (\mathrm{i}\,n\,\omega)^k}\, \mathrm{e}^{\mathrm{i}\,n\,\omega\,x}. \tag{36}$$

Le n-ième mode de y ne « répond » qu'au n-ième mode de ϕ. On dit qu'il y a « découplage » des différents modes de Fourier. L'ensemble $\{Z_n\}$ des coefficients de proportionnalité entre le n-ième mode ϕ_n de la source et le n-ième mode de la réponse y_n,

$$Z_n = \frac{1}{\sum_{k=0}^{\ell} C_k \cdot (\mathrm{i}\,n\,\omega)^k}, \tag{37}$$

constitue ce qu'on appelle la *susceptibilité* du système.

VIII. .5.*b*.ii. Filtre *RC*

À titre d'exemple, considérons le circuit électrique représenté sur la figure VIII.2.

Soit V_e le signal d'entrée (tension entre A et B), V_s le signal de sortie (entre C et D). On a $V_A - V_C = R\,i = R\,\dot{q}$, $V_\mathrm{s} = V_C - V_D = V_C - V_B = q/C$ et $V_\mathrm{e} = V_A - V_B = R\,\dot{q} + q/C$. Si le signal d'entrée $V_\mathrm{e}(t)$ est donné, le signal de sortie $V_\mathrm{s}(t)$ est solution de l'équation différentielle

$$R\,C\,\dot{V}_\mathrm{s}(t) + V_\mathrm{s}(t) = V_\mathrm{e}(t). \tag{38}$$

Si $V_\mathrm{e}(t)$ est un signal périodique de période $T = 2\,\pi/\omega$, on introduit sa série de Fourier $V_\mathrm{e}(t) = \sum_{-\infty}^{\infty} V_n\, \mathrm{e}^{\mathrm{i}\,n\,\omega\,t}$ et on procède comme au § VIII.C.5.*b*.i. Pour chaque mode $V_n\, \mathrm{e}^{\mathrm{i}\,n\,\omega\,t}$, on cherche une solution particulière de (38) de la forme $Z_n\, V_n\, \mathrm{e}^{\mathrm{i}\,n\,\omega\,t}$, d'où $(\mathrm{i}\,R\,C\,n\,\omega + 1)\,Z_n = 1$, soit

$$Z_n = \frac{1}{1 + \mathrm{i}\,R\,C\,n\,\omega} = \frac{1}{1 + \mathrm{i}\,n\,\omega/\omega_0}, \tag{39}$$

avec $\omega_0 \coloneqq 1/(R\,C)$ (qui a bien la dimension d'une fréquence). La solution particulière de (38) qu'on vient de construire,

$$\widehat{V}_{\mathrm{s}}(t) = \sum_{-\infty}^{\infty} Z_n\, V_n\, \mathrm{e}^{\mathrm{i}\,n\,\omega\,t}, \tag{40}$$

est la transformée du signal d'entrée $V_{\mathrm{e}}(t)$ par le circuit ou « filtre » RC.

Remarques
Si $\omega \gg \omega_0$ et donc que le signal d'entrée est à haute fréquence (par rapport à la fréquence caractéristique ω_0 du circuit), on a $Z_n \approx \omega_0/(\mathrm{i}\,n\,\omega)$ si $|n| \geqslant 1$. Si $V_0 = 0$ (signal d'entrée nul en moyenne), le circuit RC a alors pour effet d'*intégrer* le signal d'entrée (à la multiplication par un facteur ω_0 près) :

$$\widehat{V}_{\mathrm{s}}(t) \approx \sum_{-\infty}^{\infty} \frac{\omega_0}{\mathrm{i}\,n\,\omega}\, V_n\, \mathrm{e}^{\mathrm{i}\,n\,\omega\,t} = \omega_0 \int^t V_{\mathrm{e}}(t')\,\mathrm{d}t', \tag{41}$$

(où $\int^t V_{\mathrm{e}}(t')\,\mathrm{d}t'$ est une certaine primitive de $V_{\mathrm{e}}(t)$). Il est donc légitime d'appeler ce filtre RC un *intégrateur*.

Si par contre $\omega \ll \omega_0$, pour les petites valeurs de n (pour n tel que $n\,\omega \ll \omega_0$), on a $Z_n \approx 1$, et le signal de sortie $\widehat{V}_{\mathrm{s}}$ aux basses fréquences ne diffère guère de celui d'entrée V_{e}. On dit que le circuit RC est alors un *filtre passe-bas*. ⌟

VIII. .5.*b*.ii.*α*. Signal d'entrée en créneaux : solution particulière

Prenons un cas concret, celui d'un signal d'entrée donné par la « fonction créneau » $V_{\mathrm{e}}^{\mathrm{cr}}$, de période T, valant $-V$ sur $[0, T/2]$ et $+V$ sur $[T/2, T]$, étudiée à de petites modifications de notations près au § VII.E.2. Sa série de Fourier s'écrit

$$V_{\mathrm{e}}^{\mathrm{cr}}(t) = \sum_{n=0}^{\infty} \frac{-4\,V}{\pi\cdot(2\,n+1)} \sin([2\,n+1]\,\omega\,t). \tag{42}$$

(En fait, il n'y a convergence du membre de droite vers la fonction créneau, et donc égalité des deux membres de cette relation, qu'aux points autres que les points de discontinuité.)

Si $\omega \gg \omega_0$, le filtre transforme ce signal d'entrée de moyenne nulle en

$$\widehat{V}_{\mathrm{s}}^{\mathrm{cr}}(t) = \sum_{n=0}^{\infty} \frac{4\,\omega_0\,V}{\pi\,\omega\cdot(2\,n+1)^2} \cos([2\,n+1]\,\omega\,t), \tag{43}$$

qui doit donc être la série de Fourier d'une primitive de $\omega_0\,V_{\mathrm{e}}^{\mathrm{cr}}(t)$, c.-à-d. la série de Fourier d'une fonction continue et linéaire par morceaux. Cette fonction vaut $K - V\,\omega_0\,t$ sur $[0, T/2]$ et $K + V\,\omega_0\cdot(t-T)$ sur $[T/2, T]$, où K est une constante. C'est une « fonction triangle » (du type de celle étudiée au TD 4, à de petits changements de variable et de période près, et à une constante additive près). En $t = 0$, $\widehat{V}_{\mathrm{s}}^{\mathrm{cr}}(0) = K$, donc

$$K = \sum_{n=0}^{\infty} \frac{4\,\omega_0\,V}{\pi\,\omega\cdot(2\,n+1)^2} = \frac{\pi\,\omega_0}{2\,\omega}\,V = \frac{T\,\omega_0}{4}\,V.$$

VIII. .5.*b*.ii.*β*. Signal d'entrée en créneaux : solution générale

Selon la théorie générale des équations différentielles linéaires, à la solution particulière $\widehat{V}_{\mathrm{s}}(t)$ de (40), on doit ajouter la solution générale de l'équation sans second membre $\dot{V}_{\mathrm{s}}^{(0)} + \omega_0\,V_{\mathrm{s}}^{(0)} = 0$, soit $V_{\mathrm{s}}^{(0)}(t) = V_1\,\mathrm{e}^{-\omega_0\,t}$. La solution générale de (38) est donc $V_{\mathrm{s}}(t) = \widehat{V}_{\mathrm{s}}(t) + V_1\,\mathrm{e}^{-\omega_0\,t}$, avec V_1 une constante déterminée par les conditions initiales. Mais le deuxième terme s'atténue très rapidement, avec une échelle de temps caractéristique $t_1 = 1/\omega_0$: au bout de ce temps t_1, la fonction $V_{\mathrm{s}}^{(0)}(t)$ a décru d'un facteur e ; au bout de $2\,t_1$, elle a décru de e^2, etc. On conçoit donc qu'après un *régime transitoire*, ce terme est négligeable devant la solution périodique $\widehat{V}_{\mathrm{s}}^{\mathrm{cr}}(t)$, qui décrit le *régime permanent* du système.

Pour bien étudier le régime transitoire, examinons plus en détail le signal de sortie $V_{\mathrm{s}}(t)$ pour un signal d'entrée donné par la fonction créneau $V_{\mathrm{e}}^{\mathrm{cr}}(t)$. Autrement dit, étudions le circuit auquel on applique la tension $-V$ aux bornes A et B du circuit pendant le temps $T/2$, puis la tension opposée V pendant la deuxième demi-période, etc. Supposons par exemple qu'à l'instant $t = 0$, le potentiel $V_{\mathrm{s}} = V$.

On doit donc résoudre l'équation (38) avec un $V_{\mathrm{e}}(t) = \mp V$ selon qu'on est dans une demi-période $t \in [n\,T, (n+1/2)\,T]$ ou dans $t \in [(n+1/2)\,T, (n+1)\,T]$, soit

$$\omega_0^{-1}\,\dot{V}_{\mathrm{s}} + V_{\mathrm{s}} = \begin{cases} -V & \text{si } t \in [n\,T, (n+1/2)\,T], & (44\text{a}) \\ +V & \text{si } t \in [(n+1/2)\,T, (n+1)\,T]. & (44\text{b}) \end{cases}$$

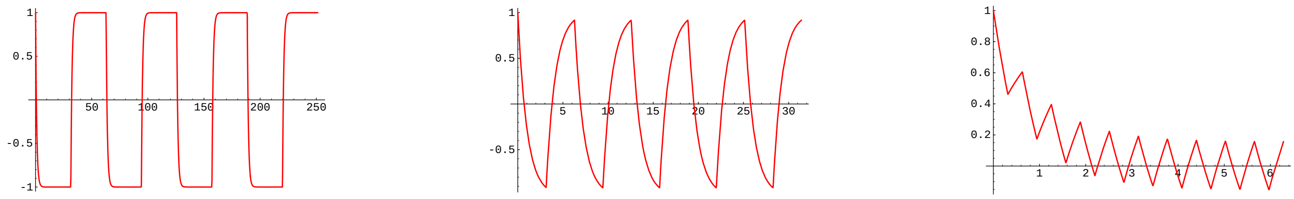


Figure VIII.3 – Courbe $V_s(t)$ pour $\omega/\omega_0 = 0{,}1$, 1, 10. On a pris $V = 1$, $\omega_0 = 1$.

Dans chacun de ces intervalles, la solution est comme toujours de la forme « solution particulière de l'équation (soit ici $\mp V$) + solution générale de l'équation sans second membre $\dot{V}_s + \omega_0 V_s = 0$ (donc une fonction proportionnelle à $e^{-\omega_0 t}$) » :

$$V_s(t) = \begin{cases} A_n e^{-\omega_0 t} - V & \text{si } t \in [n T, (n+1/2) T], & (45a) \\ B_n e^{-\omega_0 t} + V & \text{si } t \in [(n+1/2) T, (n+1) T]. & (45b) \end{cases}$$

Le coefficient A_0 est déterminé par la condition initiale $V_s(0) = V$, d'où $A_0 = 2V$, puis les autres coefficients B_n et A_n sont déterminés en imposant la continuité de la charge q, donc de $V_s(t)$ à la fin de chaque demi-période $t = (n+1/2) T$ ou $t = n T$. Par exemple à l'instant $t = T/2 = \pi/\omega$, la continuité entre les deux formes (45a) et (45b) ci-dessus donne

$$B_0 e^{-\omega_0 \pi/\omega} = 2V \cdot (1 - e^{-\omega_0 \pi/\omega}).$$

Notant

$$\varepsilon := e^{-\omega_0 \pi/\omega},$$

donc un nombre réel < 1, on vérifie par récurrence que

$$A_n = 2V \cdot (1 - \varepsilon^{-1} + \varepsilon^{-2} - \cdots + \varepsilon^{-2n}) = 2V \frac{1 + \varepsilon^{-2n-1}}{1 + \varepsilon^{-1}} \tag{46}$$

$$B_n = 2V \cdot (1 - \varepsilon^{-1} + \varepsilon^{-2} - \cdots - \varepsilon^{-2n-1}) = 2V \frac{1 - \varepsilon^{-2n-2}}{1 + \varepsilon^{-1}} \tag{47}$$

et donc

$$V_s([n+1/2] T) = -V + A_n \varepsilon^{2n+1} = -V + 2V \frac{\varepsilon + \varepsilon^{2n+2}}{1 + \varepsilon} = -V \frac{1 - \varepsilon}{1 + \varepsilon} + O(\varepsilon^{2n+2}) \tag{48}$$

$$V_s([n+1] T) = V + B_n \varepsilon^{2n+2} = -V + 2V \frac{-\varepsilon + \varepsilon^{2n+3}}{1 + \varepsilon} = V \frac{1 - \varepsilon}{1 + \varepsilon} + O(\varepsilon^{2n+3}). \tag{49}$$

Si on néglige les termes $O(\varepsilon^{2n+2})$, on voit que V_s prend aux demi-périodes la valeur $\pm V_\infty$, où

$$V_\infty = \frac{1 - \varepsilon}{1 + \varepsilon}.$$

L'interprétation est qu'après un temps $n T$, plus ou moins long selon la valeur du rapport ω_0/ω, le terme $e^{-\omega_0 \pi T} = \varepsilon^{2n}$ est devenu négligeable. C'est la fin du régime transitoire mentionné plus haut. Le système est alors dans un régime périodique où le condensateur se charge et se décharge, avec une tension à ses bornes oscillant entre les valeurs $\pm V_\infty$.

Examinons les deux cas limites où ω_0/ω est petit ou grand :

1. $\omega/\omega_0 \gg 1$, c.-à-d. que l'on excite le filtre à de hautes fréquences par rapport à ω_0. D'une part, le régime transitoire dure assez longtemps, car ε est proche de 1. De l'autre, on peut approximer la fonction $V_s(t)$ par une fonction linéaire par morceaux, et on retrouve notre fonction triangle et le circuit intégrateur.
2. $\omega/\omega_0 \ll 1$, le régime transitoire est beaucoup plus bref, car $\varepsilon \ll 1$. Dans la limite $\omega/\omega_0 \to 0$, on retrouve une fonction créneau pour $V_s(t)$: le filtre à basse fréquence est presque « transparent » ; c'est un filtre passe-bas. Noter que les discontinuités de la fonction créneau ont été adoucies : la fonction $V_s(t)$ est continue.

Exercice

(À résoudre par exemple à l'aide de Maple.)

Quel est l'effet d'une condition initiale $V(0)$ différente de celle choisie plus haut ? Par exemple, si $V(0) = V/2$, ou $V(0) = -V$, tracer la courbe de $V_s(t)$ pour $\omega/\omega_0 = 0{,}1$, 1, 10. Montrer qu'au bout d'un temps fini, qui dépend de ω/ω_0, les courbes reproduisent celles de la figure VIII.3. ⌟

VIII. . Systèmes d'équations différentielles linéaires

Comme on l'a déjà fait remarquer, les équations différentielles peuvent se présenter sous forme de systèmes d'équations différentielles couplées portant sur différentes fonctions $y_1, y_2, \ldots, y_p$. Le cas de systèmes d'équations couplées linéaires est particulièrement intéressant.

Considérons par exemple le système linéaire de deux équations du premier ordre

$$y_1' = a\, y_1 + b\, y_2, \tag{50}$$

$$y_2' = c\, y_1 + d\, y_2. \tag{51}$$

1. En dérivant la première équation par rapport à x et en y reportant l'expression de y_2' donnée par la seconde, (c.-à-d. en *éliminant* y_2' entre ces deux équations), on obtient la paire d'équations

$$y_1'' = a\, y_1' + b\cdot(c\, y_1 + d\, y_2), \tag{52}$$
$$y_1' = a\, y_1 + b\, y_2. \tag{53}$$

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

$$y_1'' = (a + d)\, y_1' + (b\, c - a\, d)\, y_1. \tag{54}$$

Cette dernière équation est une équation différentielle linéaire du second ordre à coefficients constants, qui tombe donc sous le coup de l'analyse menée plus haut : la solution générale en $y_1(x)$ (et de même en $y_2(x)$) est une combinaison linéaire de fonctions exponentielles $\mathrm{e}^{\alpha_1 x}$ et $\mathrm{e}^{\alpha_2 x}$, où α_1 et α_2 sont solutions de l'équation

$$\alpha^2 - (a + d)\, \alpha + (a\, d - b\, c) = 0. \tag{55}$$

En résumé, le système de deux équations linéaires du premier ordre à coefficients constants (50) est donc équivalent à une équation linéaire du second ordre.

2. Discutons maintenant une autre méthode : en effectuant des combinaisons linéaires des deux équations de (50) avec des coefficients λ et μ quelconques, on trouve que

$$(\lambda\, y_1 + \mu\, y_2)' = (a\, \lambda + c\, \mu)\, y_1 + (b\, \lambda + d\, \mu)\, y_2. \tag{56}$$

Supposons qu'on sache trouver λ et μ tels que

$$a\, \lambda + c\, \mu = \alpha\, \lambda \tag{57}$$
$$b\, \lambda + d\, \mu = \alpha\, \mu, \tag{58}$$

pour un certain nombre α. Alors, en reportant (57) dans (56), on voit que l'équation (56) est une équation différentielle linéaire du premier ordre particulièrement simple pour la fonction $y := \lambda\, y_1 + \mu\, y_2$,

$$y' = \alpha\, y, \tag{59}$$

ce qui s'intègre immédiatement en $y = C\, \mathrm{e}^{\alpha x}$. Or le système (57), qui est un système d'équations algébriques du premier degré à deux variables, se résout aisément. En éliminant par exemple μ entre les deux équations, on trouve que

$$c\, \mu = (\alpha - a)\, \lambda \quad \text{et} \quad ([d - \alpha]\cdot[a - \alpha] - b\, c)\, \lambda = 0. \tag{60}$$

Cette dernière équation n'admet de solution non nulle en λ que si α satisfait l'équation du second degré

$$(d - \alpha)\cdot(a - \alpha) - b\, c = 0. \tag{61}$$

Supposons les racines distinctes : il existe donc deux valeurs de α, donc aussi deux paires (λ, μ) (à un facteur global près) remplissant les conditions ci-dessus, c.-à-d. conduisant à une équation du type (59) pour la combinaison linéaire correspondante. Chacune de ces deux combinaisons linéaires est appelée un *mode propre* du système initial (50). Par les combinaisons algébriques que nous avons effectuées, nous avons donc réduit le système initial, équivalent à une équation du second ordre selon la discussion du point 1 ci-dessus, à la solution de deux équations du premier ordre découplées. On note que les deux approches ont en commun de faire jouer un rôle central à l'équation « caractéristique » (55)-(61).

3. Ceci se généralise à des systèmes impliquant plus de fonctions ou des équations d'ordres plus élevés. Réécrivons le système linéaire (50) sous une forme matricielle :

$$\frac{\mathrm{d}}{\mathrm{d}x} Y = A\, Y, \tag{62}$$

où

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Supposons maintenant que la matrice A puisse se diagonaliser par un changement de base $A = P\, D\, P^{-1}$, où $P^{-1} = \left(\begin{smallmatrix} \lambda_1 & \mu_1 \\ \lambda_2 & \mu_2 \end{smallmatrix}\right)$ est une matrice 2×2 indépendante de x, $D = \left(\begin{smallmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{smallmatrix}\right)$ une matrice diagonale, et α_1 et α_2 sont les *valeurs propres* de A. En posant $Z = P^{-1}\, Y = \left(\begin{smallmatrix} z_1 \\ z_2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} \lambda_1\, y_1 + \mu_1\, y_2 \\ \lambda_2\, y_1 + \mu_2\, y_2 \end{smallmatrix}\right)$, on a, en multipliant à gauche les deux membres de l'équation matricielle (62) par P^{-1},

$$\frac{\mathrm{d}}{\mathrm{d}x} Z = D\, Z,$$

soit

$$\begin{pmatrix} z_1' \\ z_2' \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \tag{63}$$

c.-à-d. deux équations découplées :

$$z_1' = \alpha_1 \, z_1 \qquad \text{et} \qquad z_2' = \alpha_2 \, z_2. \tag{64}$$

La démarche reproduit celle suivie au point 2. Les valeurs propres α_1 et α_2 sont les racines de l'équation (61) ; la matrice est diagonalisable, donc P existe et est inversible, car les deux racines sont distinctes ; et les combinaisons z_1 et z_2 sont les deux « modes propres » définis en 2.

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

Rappel

Soient A une matrice carrée de taille $n \in \mathbb{N}^*$, à coefficients quelconques dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et I_n la matrice identité. Un scalaire α est une *valeur propre* de A s'il existe un vecteur $X \in \mathbb{K}^n$ non nul, appelé *vecteur propre*, tel que $A\,X = \alpha\,X$, ou encore $(A - \alpha\,I_n)\,X = 0$.

Remarquons que si $X_{i,1}$ et $X_{i,2}$ sont deux vecteurs propres pour la même valeur propre α_i, toute combinaison linéaire $\lambda_1\,X_{i,1} + \lambda_2\,X_{i,2}$ avec $(\lambda_1, \lambda_2) \neq (0, 0)$ est encore un vecteur propre pour α_i. Les vecteurs propres associés à la valeur propre α_i constituent donc un sous-espace vectoriel, E_i, de $\mathbb{K}^n$.

Si $A - \alpha\,I_n$ est inversible, la seule solution de l'équation $(A - \alpha\,I_n)\,X = 0$ est le vecteur nul. Pour que A admette des vecteurs propres pour un certain α, il faut donc que $A - \alpha\,I_n$ ne soit pas inversible, c.-à-d. que $\det(A - \alpha\,I_n) = 0$. Or $\det(A - \alpha\,I_n)$ est un polynôme d'ordre n en la variable α. Il admet donc au plus n racines. Notons $\alpha_1, \ldots, \alpha_k$ ($k \leqslant n$) les racines et $m_1, \ldots, m_k$ leurs multiplicités ($m_1 + \cdots + m_k = n$).

Le sous-espace propre E_i est au moins de dimension 1 puisque le vecteur nul n'est pas un vecteur propre par définition. On peut en fait montrer que $\dim E_i \in [\![1, m_i]\!]$. Si, $\forall\, i \in [\![1, k]\!]$, $\dim E_i = m_i$ [*3], la matrice A est diagonalisable, c.-à-d. qu'on peut l'écrire sous la forme $A = P\,D\,P^{-1}$.

Si A est diagonalisable, choisissons, pour chaque valeur propre, m_i vecteurs propres linéairement indépendants $X_{i,j}$ ($j \in [\![1, m_i]\!]$) dans E_i. On peut alors écrire D et P de la manière suivante :

$$D = \begin{pmatrix} \alpha_1 & 0 & \ldots & & & \ldots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & \alpha_1 & & & & \\ & & & \ddots & & & \\ & & & & \alpha_k & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \ldots & & & \ldots & 0 & \alpha_k \end{pmatrix}$$

$$\begin{array}{cccccccc} & X_{1,1} & \ldots & X_{1,m_1} & \ldots\ldots & X_{k,1} & \ldots & X_{k,m_k} \\ & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ P = & \multicolumn{7}{c}{\begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}} \end{array}$$

┘

Exemple 69

Soit

$$A = \begin{pmatrix} 1 & 3 \\ 1 & -1 \end{pmatrix}.$$

$$A - \alpha\,I_2 = \begin{pmatrix} 1 - \alpha & 3 \\ 1 & -1 - \alpha \end{pmatrix},$$

3. C'est notamment le cas si toutes les racines sont distinctes, puisqu'on a alors $1 \leqslant \dim E_i \leqslant m_i = 1, \forall\, i$.

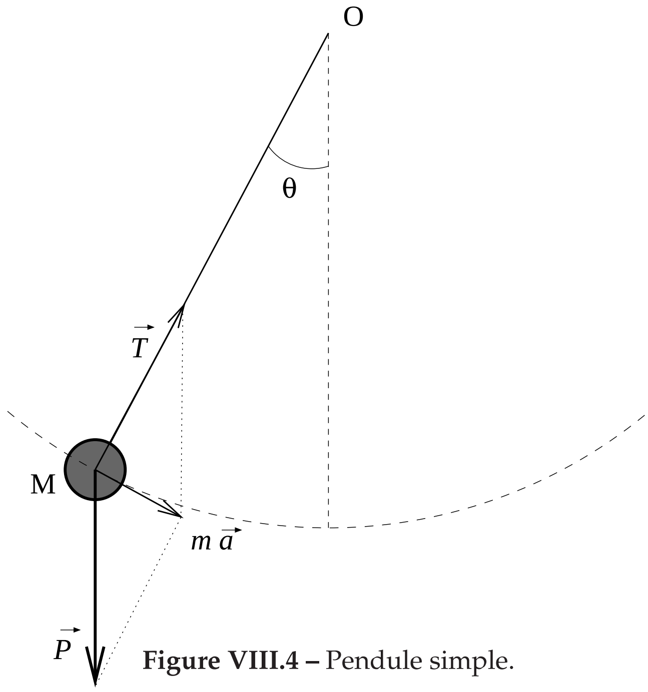


Figure VIII.4 – Pendule simple.

donc $\det(A - \alpha\, I_2) = \alpha^2 - 4$: A a deux valeurs propres distinctes, $\alpha_1 = 2$ et $\alpha_2 = -2$. $X_1 = (3, 1)$ et $X_2 = (-1, 1)$ sont des vecteurs propres respectivement associés à α_1 et α_2. On a donc

$$D = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \qquad \text{et} \qquad P = \begin{pmatrix} 3 & -1 \\ 1 & -1 \end{pmatrix}.$$

$$P^{-1} = \begin{pmatrix} 1/4 & 1/4 \\ -1/4 & 3/4 \end{pmatrix},$$

et l'on vérifie sans difficulté que $A = P\, D\, P^{-1}$. ⌟

VIII. . Autres méthodes de résolution d'équations différentielles

VIII. .1. Intégrales premières

Il existe des situations où l'on peut réduire l'ordre de l'équation différentielle en effectuant explicitement une intégration. Considérons par exemple un pendule simple constitué d'une masse ponctuelle m fixée à l'extrémité d'un fil de masse négligeable de longueur ℓ. Le pendule oscille dans un plan vertical. Il effectue son mouvement sous la seule action de son poids, $\vec{P} = -m\, g\, \widehat{z}$, où $\widehat{z}$ désigne un vecteur unitaire vertical ascendant, et de la tension $\vec{T}$ du fil. D'après la relation fondamentale de la dynamique, $m\,\vec{a} = \sum \vec{F} = \vec{T} - m\, g\, \widehat{z}$. En projetant la tension inconnue $\vec{T}$ sur une direction tangente à la trajectoire circulaire, on obtient

$$m\, \ell\, \ddot{\theta} = -m\, g \sin\theta. \tag{65}$$

Multiplions cette équation par $\dot{\theta}$ et reconnaissons dans $\ddot{\theta}\, \dot{\theta}$ la dérivée par rapport au temps de $\frac{1}{2}\,\dot{\theta}^2$ et dans $-\dot{\theta}\sin\theta$ celle de $\cos\theta$. L'équation (65) s'écrit donc

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\,\ell\, m\, \dot{\theta}^2 - g\, m \cos\theta\right) = 0, \tag{66}$$

qui s'intègre en

$$\frac{1}{2}\, m\, \ell\, \dot{\theta}^2 = m\, g\cdot(\cos\theta - \cos\theta_{\mathrm{m}}), \tag{67}$$

où θ_{m} est une constante d'intégration que l'on peut choisir dans $[0, \pi]$. Pour $\theta = \pm\theta_{\mathrm{m}}$, le membre de droite s'annule, donc celui de gauche aussi : la vitesse angulaire $\dot{\theta}$ s'annule alors, donc θ_{m} est l'amplitude d'oscillation du pendule.

Réécrivons l'équation précédente (multipliée par ℓ) sous la forme

$$\frac{1}{2}\, m\, \ell^2\, \dot{\theta}^2 - m\, g\, \ell \cos\theta = -m\, g\, \ell \cos\theta_{\mathrm{m}} := E, \tag{68}$$

avec E indépendante du temps. Sur le plan mathématique, nous avons réduit notre équation de départ du second ordre (65) à une équation du premier ordre, ce qui est un progrès substantiel. Sur le plan physique, ce que cette « constante du mouvement » (ou « intégrale première » de l'équation initiale) exprime est bien sûr la loi de conservation de l'énergie interne du système isolé {Terre, pendule} (ou, de manière équivalente, de l'énergie mécanique du pendule) : $\frac{1}{2}\, m\, \ell^2\, \dot{\theta}^2$ est l'énergie cinétique et $m\, g\, \ell \cos\theta$ l'énergie potentielle.

VIII. .2. Équations à variables séparées

On appelle *équation à variables séparées* une équation différentielle du premier ordre $F(y, y', x) = 0$ qui peut se mettre sous la forme $f(y)\,\mathrm{d}y = g(x)\,\mathrm{d}x$ et donc s'intégrer en

$$\int^{y} f(y')\,\mathrm{d}y' = \int^{x} g(x')\,\mathrm{d}x' + \text{constante}. \tag{69}$$

Exemple 70
Montrer que l'équation $\dot{x} = -\alpha^2 - x^2$ est à variables séparées. En chercher les solutions. ⌟

Exemple 71
On suppose qu'un corps en chute libre est freiné par une force proportionnelle au carré de sa vitesse (parachute). Montrer que son équation du mouvement s'écrit

$$\ddot{z} = k\, \dot{z}^2 - g$$

et que l'équation différentielle satisfaite par $\dot{z}$ ressemble beaucoup à celle de l'exemple précédent. ⌟

Exemple 72
Examinons l'équation du pendule simple que nous avons obtenue au paragraphe précédent,

$$\frac{1}{2}\, \ell\, \dot{\theta}^2 = g\cdot(\cos\theta - \cos\theta_{\mathrm{m}}). \tag{70}$$

Puisque $|\theta| \leqslant \theta_{\mathrm{m}}$, le signe du membre de droite est bien positif. Cette équation se réécrit

$$\frac{|\mathrm{d}\theta|}{\sqrt{\cos\theta - \cos\theta_{\mathrm{m}}}} = \sqrt{\frac{2\, g}{\ell}}\,\mathrm{d}t,$$

où toute la dépendance dans la variable θ a été mise au membre de gauche, toute celle en t (ici triviale) au membre de droite. Comme au paragraphe précédent, supposons le pendule lâché à l'instant $t = 0$ avec une vitesse nulle d'un angle $\theta = -\theta_{\mathrm{m}} \leqslant 0$. Dans la première demi-oscillation du pendule, θ croît, donc $\mathrm{d}\theta \geqslant 0$. On peut résoudre l'équation en intégrant chaque membre :

$$\int_{\theta'=-\theta_{\mathrm{m}}}^{\theta} \frac{\mathrm{d}\theta'}{\sqrt{\cos\theta' - \cos\theta_{\mathrm{m}}}} = \sqrt{\frac{2\, g}{\ell}}\, t. \tag{71}$$

Il reste à calculer effectivement l'intégrale du membre de gauche et à en extraire la fonction cherchée $\theta(t)$. Cette intégration mène à des fonctions non élémentaires, dites fonctions elliptiques, et nous ne la poursuivrons pas ici. ⌟

VIII. .3. Changement de variable

Dans certains cas, une équation linéaire à coefficients non constants peut se ramener à une équation à coefficients constants, qu'on sait résoudre assez aisément. C'est par exemple le cas des équations dites d'Euler, $\sum C_k\, x^k\, y^{(k)}(x) = 0$, étudiées pour $x > 0$, avec les C_k constants. Montrer que le changement de variable $x \mapsto t = \ln x$ permet de se ramener à une équation différentielle en t, linéaire et à coefficients constants.

VIII. .4. Changement de la fonction inconnue

Il existe de nombreux cas où un changement judicieux de la fonction inconnue, solution cherchée de l'équation différentielle, simplifie considérablement la discussion. Il s'agit donc, étant donnée une équation différentielle satisfaite par une fonction $y(x)$, d'écrire et d'étudier l'équation différentielle qui en résulte pour la nouvelle fonction $z(x) = F(y[x])$.

Étudions par exemple l'équation de Riccati,

$$y' + a(x)\, y^2 + b(x)\, y + c(x) = 0. \tag{72}$$

Supposons qu'on connaisse une solution $y_0(x)$ et posons

$$y(x) = y_0(x) + \frac{1}{z(x)}$$

avec une nouvelle fonction inconnue $z(x)$. En reportant cette expression dans (72) et en utilisant le fait que $y_0(x)$ satisfait elle-même (72), on obtient

$$-\frac{z'}{z^2} + a\cdot\left(2\,\frac{y_0}{z} + \frac{1}{z^2}\right) + \frac{b}{z} = 0$$

(toutes les dépendances en x sont sous-entendues), soit l'équation *linéaire*

$$-z' + a\cdot(2\, y_0\, z + 1) + b\, z = 0, \tag{73}$$

à laquelle on applique les théorèmes et méthodes du § VIII.c.

Un autre exemple, celui de l'équation de Bernoulli, sera étudié en TD.

VIII. .5. Résolution d'une équation différentielle par une série entière

Considérons, à titre d'exemple, le cas de l'équation d'Airy,

$$f''(x) = x\, f(x) \tag{74}$$

et cherchons-en une solution sous forme d'une série entière,

$$f(x) = \sum_{n=0}^{\infty} a_n\, x^n.$$

En dérivant terme à terme et en insérant dans (74), on trouve

$$2\, a_2 + \sum_{n=1}^{\infty} ([n+1]\cdot[n+2]\, a_{n+2} - a_{n-1})\, x^n = 0,$$

soit, en identifiant à zéro le coefficient de chaque terme x^n,

$$a_2 = 0 \quad \text{et} \quad a_{n+2} = \frac{a_{n-1}}{(n+1)\cdot(n+2)} \quad \text{pour } n \geqslant 1.$$

Cela fournit une relation de récurrence entre les a_n, qui, complétée par la donnée de $a_0 = f(0)$ et $a_1 = f'(0)$, deux conditions initiales indispensables à la résolution de l'équation du second ordre (74), permet de déterminer tous les a_n. Comme a_n se comporte comme $1/(n!)^{2/3}$, le rayon de convergence de la série (cf. § VI.c) est infini : la série converge pour tout x.

Les deux solutions de base, notées $\mathrm{Ai}\, x$ et $\mathrm{Bi}\, x$, sont distinguées par leur comportement à l'infini : $\mathrm{Ai}\, x \to 0$ et $\mathrm{Bi}\, x \to \infty$.

La fonction d'Airy intervient dans des problèmes de physique, telle la description quantitative de l'arc-en-ciel. . .

Inversement, connaître l'équation différentielle satisfaite par la somme d'une série entière permet parfois de calculer cette somme. Considérons par exemple la somme

$$\sum_{k=0}^{\infty}\left(\frac{1}{16}\right)^k \left(\frac{4}{8\,k+1} - \frac{2}{8\,k+4} - \frac{1}{8\,k+5} - \frac{1}{8\,k+6}\right). \tag{75}$$

On démontre aisément qu'elle est convergente (le vérifier !). Pour calculer sa somme, construisons la série entière

$$S(x) = \sum_{k=0}^{\infty}\left(\frac{x^8}{16}\right)^k \cdot \left(\frac{4\,x}{8\,k+1} - \frac{2\,x^4}{8\,k+4} - \frac{x^5}{8\,k+5} - \frac{x^6}{8\,k+6}\right). \tag{76}$$

Vérifier les points suivants :

1. la série converge pour $|x| < \sqrt{2}$;
2. la fonction $S(x)$ satisfait
$$S'(x) = \frac{4 - 2\,x^3 - x^4 - x^5}{1 - x^8/16}$$
et $S(0) = 0$;
3. la fonction
$$\varphi(x) := -4\arctan(1-x) + 2\ln(2-x^2) - 2\ln(x^2 - 2\,x + 2)$$
satisfait $\phi'(x) = S'(x)$ et $\varphi(0) = -\pi$.

En conclure que $S(x) = \varphi(x) + \pi$, donc que la somme (75) cherchée est $S(1) = \varphi(1) + \pi = \pi$. On pourra utiliser un logiciel de calcul formel comme Maple pour vérifier le point 3.