

Fondamentaux des Mathématiques II

Stéphane Attal

26 mars 2018

Table des matières

I	ALGÈBRE	5
1	Fractions rationnelles	7
1.1	Définitions	7
1.2	Décomposition en éléments simples	10
1.3	Méthodes pratiques	13
1.3.1	La partie entière	13
1.3.2	Multiplicité 1, degré 1	13
1.3.3	Multiplicité quelconque, degré 1	14
1.3.4	Cas général	17
2	Espaces vectoriels	19
2.1	Structure d'espace vectoriel	20
2.2	Opérations sur les espaces vectoriels	22
2.3	Bases, dimension	24
2.4	Dimension des sous-espaces	26
3	Applications linéaires	29
3.1	Définitions	29
3.2	Noyau et image	30
3.3	Construction d'applications linéaires	31
3.4	Rang	32
4	Calcul matriciel	35
4.1	Représentation matricielle des applications linéaires	35
4.2	Opérations sur les matrices	37
II	ANALYSE	39
5	Développements limités	41
5.1	Comparaison de fonctions	41

5.2	Equivalence de fonctions	42
5.3	Les formules de Taylor	44
5.3.1	Un rappel concernant la dérivation	44
5.3.2	La formule de Taylor-Young	45
5.3.3	Développements limités	46
5.3.4	Développement des fonctions usuelles	48
5.3.5	Taylor-Lagrange, reste intégral	50
5.4	Propriétés des développements limités	52
5.4.1	Opérations sur les développements limités	52
5.4.2	Comportement local près des points critiques	54
6	Intégration	57
6.1	Fonctions en escalier	58
6.1.1	Subdivisions	58
6.1.2	Fonctions en escalier	58
6.1.3	Intégrales de fonctions en escalier	59
6.2	Fonctions Riemann-intégrables	61
6.2.1	Construction de l'intégrale de Riemann	61
6.2.2	Opérations sur les fonctions intégrables	64
6.2.3	Intégrales et inégalités	65
6.2.4	Intégrales et produits	67
6.3	Familles de fonctions intégrables	68
6.3.1	Manipulation de fonctions intégrables	68
6.3.2	Monotonie	69
6.3.3	Continuité	69
6.3.4	Convention et relation de Chasles	71
6.4	Primitives et intégrales	72
6.4.1	Le théorème fondamental	72
6.4.2	Intégration par parties	75
6.4.3	Changement de variables	75
6.5	Quelques résultats	78
6.6	Sommes de Riemann, de Darboux, surfaces etc.	80
6.6.1	Sommes de Darboux	80
6.6.2	Sommes de Riemann	81
6.6.3	Estimation d'erreurs	83
6.6.4	Lien avec les surfaces	85
6.7	Intégrales de suites de fonctions	85
6.7.1	Ce qui ne marche pas	85
6.7.2	Limite uniforme	87
7	Equations différentielles	91

Première partie

ALGEBRE

Chapitre 1

Fractions rationnelles

1.1 Définitions

Les *fractions rationnelles* font suite aux polynômes que vous avez étudiés au Semestre 1. Là encore l'esprit est algébrique plutôt qu'analytique. Il ne s'agit pas d'étudier les propriétés exactes des fonctions rationnelles (graphes, limites, dérivées etc.), mais plutôt les propriétés algébriques de ces fonctions : manipulation des expressions, simplifications, etc. Avec un objectif majeur : la *décomposition en éléments simples*. C'est une représentation très pratique et très simple des fractions rationnelles ; la principale application est le calcul des primitives. En effet, à la fin de ce chapitre nous serons capable de calculer la primitive de n'importe quelle fonction de la forme

$$F(x) = \frac{P(x)}{Q(x)}$$

où P et Q sont des polynômes. Par exemple

$$F(x) = \frac{3x^4 - 7x + 5}{x^3 - x^2 + x - 12}.$$

Aujourd'hui avec vos outils usuels de calcul de primitives, vous ne pouvez en aucune manière calculer une primitive (ou une intégrale) avec une telle fonction.

Nous pouvons attaquer les définitions.

Définition 1 Une fraction rationnelle sur un corps \mathbb{K} (qui sera \mathbb{Q} , \mathbb{R} ou \mathbb{C}) est une expression formelle de la forme

$$F(X) = \frac{P(X)}{Q(X)}$$

où P et Q sont des éléments de $\mathbb{K}[X]$.

On comprend cette expression comme un quotient de polynômes avec les propriétés algébriques qui en découlent, en particulier on identifie deux fractions rationnelles

$$\frac{P_1(X)}{Q_1(X)} = \frac{P_2(X)}{Q_2(X)}$$

si et seulement si

$$P_1(X)Q_2(X) = P_2(X)Q_1(X)$$

au sens de l'égalité dans $\mathbb{K}[X]$.

L'ensemble des fractions rationnelles sur le corps \mathbb{K} est noté $\mathbb{K}(X)$.

Proposition 1.1.1 *Toute fraction rationnelle $f(X)$ peut s'écrire*

$$F(X) = \frac{P(X)}{Q(X)}$$

avec $P \wedge Q = 1$ et Q unitaire. L'écriture de F sous cette forme est unique.

Démonstration Si $F(X) = P(X)/Q(X)$ et que $D = P \wedge Q$, on alors $P = P_1 D$ et $Q = Q_1 D$ donc $f = (P_1 D)/(Q_1 D)$. Mais avec la définition de l'égalité des fractions rationnelles, cette fraction est aussi égale à P_1/Q_1 .

On peut aussi factoriser le coefficient de plus haut degré de Q et simplifier le quotient, cela rend Q unitaire. Nous avons prouvé l'existence de cette représentation de f .

Pour l'unicité, par l'absurde supposons que $F = P_1/Q_1 = P_2/Q_2$ sont deux représentations de F avec ces propriétés. On a donc $P_1 Q_2 = P_2 Q_1$, donc Q_2 divise $P_2 Q_1$, mais comme $Q_2 \wedge P_2 = 1$ on a Q_2 divise Q_1 , par le Théorème de Gauss pour les polynômes. De la même façon on montre que Q_1 divise Q_2 , donc les polynômes Q_1 et Q_2 sont égaux à un facteur scalaire près. Mais comme ils sont tous deux unitaires, ils sont donc égaux. On obtient ensuite facilement que P_1 et P_2 sont aussi égaux. \square

Les polynômes usuels ($P(X) \in \mathbb{K}[X]$) sont vu comme des cas particulier de fractions rationnelles, en les identifiant à la fraction rationnelle $P(X)/\mathbb{1}$.

Définition 2 *Les opérations de somme et de produit pour les fractions rationnelles sont définies comme suit : si $F(X) = P_1(X)/Q_1(X)$ et $G(X) = P_2(X)/Q_2(X)$, on pose*

$$F(X) + G(X) = \frac{P_1(X)Q_2(X) + P_2(X)Q_1(X)}{Q_1(X)Q_2(X)}$$

et

$$F(X)G(X) = \frac{P_1(X)P_2(X)}{Q_1(X)Q_2(X)}.$$

Ce sont encore des éléments de $\mathbb{K}(X)$.

Muni de ces lois $(\mathbb{K}(X), +, \times)$ est un corps commutatif (remarque culturelle, hors programme).

Définition 3 On appelle degré d'une fraction rationnelle $F = P/Q$, la quantité

$$\deg F = \deg P - \deg Q,$$

où $\deg P$ et $\deg Q$ sont les degrés usuels de P et Q en tant que polynômes.

Notez que $\deg F$ appartient à \mathbb{Z} et pas seulement à \mathbb{N} .

Proposition 1.1.2 Le degré de F ne dépend pas de la représentation de F sous la forme P/Q .

On a les mêmes propriétés que pour les polynômes, à savoir

$$\deg(F + G) \leq \max\{\deg f, \deg g\}, \quad \deg FG = \deg f + \deg g.$$

Démonstration Si $F = P_1/Q_1 = P_2/Q_2$, alors $P_1Q_2 = P_2Q_1$, donc $\deg(P_1Q_2) = \deg(P_2Q_1)$, c'est à dire $\deg P_1 + \deg Q_2 = \deg P_2 + \deg Q_1$. Ce qui donne finalement $\deg P_1 - \deg Q_1 = \deg P_2 - \deg Q_2$. Ce qui prouve que $\deg F$ ne dépend pas de la représentation $f = P/Q$.

La formule de degré du produit est laissée en exercice. Faisons le cas de la somme. On a

$$\begin{aligned} \deg F + G &= \deg P_1Q_2 + P_2Q_1 - \deg Q_1Q_2 \\ &\leq \max\{\deg P_1Q_2, \deg P_2Q_1\} - \deg Q_1 - \deg Q_2 \\ &= \max\{\deg P_1 + \deg Q_2, \deg P_2 + \deg Q_1\} - \deg Q_1 - \deg Q_2. \end{aligned}$$

On distingue les deux cas :

– si $\deg P_1 + \deg Q_2 \leq \deg P_2 + \deg Q_1$ (ce qui revient à $\deg F \leq \deg G$), alors le dernier membre ci-dessus vaut $\deg P_2 + \deg Q_1 - \deg Q_1 - \deg Q_2 = \deg G$,

– si $\deg P_1 + \deg Q_2 \geq \deg P_2 + \deg Q_1$ (ce qui revient à $\deg f \geq \deg g$), alors le dernier membre ci-dessus vaut $\deg P_1 + \deg Q_2 - \deg Q_1 - \deg Q_2 = \deg F$.

Donc au final $\deg F + G \leq \max\{\deg F, \deg G\}$. \square

Définition 4 Soit $F = P/Q \in \mathbb{K}(X)$ mis sous sa forme réduite. On appelle zéros de F les racines de P . On appelle pôles de F les racines de Q .

Les pôles sont les points de \mathbb{K} où la fonction f n'est pas définie.

La multiplicité d'un pôle x_0 de f est la multiplicité de x_0 comme racine de Q .

Définition 5 On définit la dérivée formelle de la fraction rationnelle $F = P/Q$ par

$$F' = \frac{P'Q - PQ'}{Q^2}.$$

C'est aussi une fraction rationnelle.

1.2 Décomposition en éléments simples

La première étape consiste à se ramener à des fractions rationnelles P/Q où $\deg P < \deg Q$.

Proposition 1.2.1 Soit $F \in \mathbb{K}(X)$, alors il existe une unique décomposition

$$F = E + G$$

où E est un polynôme ($E \in \mathbb{K}[X]$) et G est une fraction rationnelle telle que $\deg G < 0$.

Démonstration On prend une représentation $f = P/Q$. Si $\deg P < \deg Q$ il n'y a rien à faire. Si $\deg P \geq \deg Q$ alors on effectue une division euclidienne de P par Q , ce qui donne une décomposition

$$P = EQ + R$$

avec $E, R \in \mathbb{K}[X]$ et $\deg R < \deg Q$. Cela donne

$$F = E + \frac{R}{Q},$$

d'où la représentation annoncée.

Montrons l'unicité de cette représentation. Si $F = E_1 + G_1 = E_2 + G_2$ sont deux décompositions de cette forme, alors $E_1 - E_2 = G_2 - G_1$ est de degré strictement négatif, ce qui est impossible pour un polynôme, sauf pour le polynôme nul. On a donc $E_1 = E_2$ et donc $G_1 = G_2$. \square

On traite maintenant chaque pôle.

Proposition 1.2.2 Soit $F = P/Q$ une fraction rationnelle, avec un pôle a de multiplicité k . Soit $\tilde{Q} \in \mathbb{K}[X]$ tel que $Q(X) = (X - a)^k \tilde{Q}(X)$. Alors, il existe un unique couple $R, S \in \mathbb{K}[X]$ tel que $\deg S < k$ et

$$F = \frac{R}{\tilde{Q}} + \frac{S}{(X - a)^k}.$$

Démonstration Regardons tout de suite l'unicité. Si

$$F = \frac{R_1}{\tilde{Q}} + \frac{S_1}{(X-a)^k} = \frac{R_2}{\tilde{Q}} + \frac{S_2}{(X-a)^k}$$

on a alors

$$(R_1 - R_2)(X-a)^k = (S_1 - S_2)\tilde{Q}.$$

La racine a est de multiplicité au moins k dans le membre de gauche, alors qu'elle ne peut être que de multiplicité strictement inférieure dans le membre de droite. Donc on a forcément $R_1 - R_2 = 0$ et on obtient facilement l'unicité.

Montrons l'existence maintenant. Comme $(X-a)^k$ et \tilde{Q} sont premiers entre eux, il existe $U, V \in \mathbb{K}[X]$ tels que $P = U(X-a)^k + V\tilde{Q}$. On effectue ensuite la division euclidienne de V par $(X-a)^k$, ce qui donne $V = A(X-a)^k + B$, avec $\deg B < k$. On a ainsi

$$P = (U + A\tilde{Q})(X-a)^k + B\tilde{Q},$$

ce qui donne

$$F = \frac{U + A\tilde{Q}}{\tilde{Q}} + \frac{B}{(X-a)^k}.$$

C'est la décomposition annoncée. \square

Définition 6 La partie $S/(X-a)^k$ dans le théorème ci-dessus est appelée partie polaire de F associée au pôle a .

On peut maintenant énoncer le premier grand théorème de décomposition en éléments simples.

Théorème 1.2.3 (Décomposition en éléments simples dans $\mathbb{C}(X)$) Pour toute fraction rationnelle $F = P/Q \in \mathbb{C}(X)$ avec

$$Q(X) = (X-a_1)^{k_1} \dots (X-a_n)^{k_n}$$

il existe une unique décomposition

$$F = E + \sum_{i=1}^n \left(\frac{\lambda_1^i}{X-a_i} + \frac{\lambda_2^i}{(X-a_i)^2} + \dots + \frac{\lambda_{k_i}^i}{(X-a_i)^{k_i}} \right),$$

avec $E \in \mathbb{C}[X]$ et tous les $\lambda_j^i \in \mathbb{C}$.

Démonstration On applique la proposition 1.2.1, puis on répète la proposition 1.2.2 pour obtenir la représentation

$$F = E + \sum_{i=1}^n \frac{A_i}{(X - a_i)^{k_i}}, \quad (1.1)$$

avec $E, A_1, \dots, A_n \in \mathbb{K}[X]$ et $\deg A_i < k_i$. Il suffit maintenant de voir comment s'effectue le passage de $A_i/(X - a_i)^{k_i}$ à

$$\frac{\lambda_1^i}{X - a_i} + \frac{\lambda_2^i}{(X - a_i)^2} + \dots + \frac{\lambda_{k_i}^i}{(X - a_i)^{k_i}}.$$

Le polynôme A_i appartient à $\mathbb{C}_{k_i-1}[X]$, la famille $\mathbb{1}, (X - a_i), \dots, (X - a_i)^{k_i-1}$ est une base de $\mathbb{C}_{k_i-1}[X]$ (car échelonnée), donc il existe $\lambda_1^i, \dots, \lambda_{k_i}^i \in \mathbb{C}$ tels que

$$A_i = \sum_{j=1}^{k_i-1} \lambda_j^i (X - a_i)^{k_j}. \quad (1.2)$$

Ce qui donne la décomposition voulue.

Montrons l'unicité. Si on a deux telles décompositions alors les parties entières sont égales, en utilisant l'argument déjà vu de degré strictement négatif pour un polynôme.

Ensuite, dans la proposition 1.2.2, la partie polaire de a est unique. Donc la décomposition (1.1) est unique.

Ensuite la décomposition (1.2) est aussi unique car c'est une décomposition dans une base. On a montré l'unicité. \square

Dans $\mathbb{R}(X)$ la décomposition en éléments simples prend une forme différente.

Théorème 1.2.4 (Décomposition en éléments simples dans $\mathbb{R}(X)$) *Pour toute fraction rationnelle $F = P/Q \in \mathbb{R}(X)$ avec*

$$Q(X) = (X - a_1)^{k_1} \dots (X - a_n)^{k_n} (X^2 + \beta_1 X + \gamma_1)^{l_1} \dots (X^2 + \beta_m X + \gamma_m)^{l_m},$$

la décomposition de Q en facteurs irréductibles de $\mathbb{R}[X]$, il existe une unique décomposition

$$F = E + \sum_{i=1}^n \left(\frac{\lambda_1^i}{X - a_i} + \frac{\lambda_2^i}{(X - a_i)^2} + \dots + \frac{\lambda_{k_i}^i}{(X - a_i)^{k_i}} \right) + \sum_{i=1}^m \left(\frac{\mu_1^i X + \eta_1^i}{X^2 + \beta_i X + \gamma_i} + \dots + \frac{\mu_m^i X + \eta_m^i}{(X^2 + \beta_i X + \gamma_i)^{l_m}} \right)$$

avec $E \in \mathbb{R}[X]$ et tous les $\lambda_j^i, \mu_j^i, \eta_j^i \in \mathbb{R}$.

Démonstration On décompose Q en facteurs irréductibles dans $\mathbb{C}[X]$ et on considère la décomposition en éléments simples associée, pour F , sur $\mathbb{C}[X]$. Comme F est réelle, en particulier les racines de Q sont complexes conjuguées. Il est alors facile de voir, en considérant \overline{F} que les coefficients λ_j^i de la décomposition de F en éléments simples, sont conjugués pour les racines conjuguées, i.e. de la forme

$$\frac{\lambda}{(X - z)^k} + \frac{\overline{\lambda}}{(X - \overline{z})^k}.$$

Quand on regroupe ces deux termes on obtient quelque chose de la forme $P/(X^2 + \beta X + \gamma)^k$ où les polynômes sont réels maintenant et où $\deg P = k$.

En faisant des divisions euclidiennes successives de P par $X^2 + \beta X + \gamma$ on obtient une représentation de P sous la forme

$$P = \sum_{i=0}^k D_i (X^2 + \beta X + \gamma)^i$$

où les D_i sont de degré 1 au plus. Cela donne la représentation annoncée. \square

1.3 Méthodes pratiques

Nous allons passer en revue, à travers des exemples, des méthodes effectives pour calculer ces décompositions en éléments simples.

1.3.1 La partie entière

On rappelle ici brièvement que la partie entière de F est obtenue en effectuant la division euclidienne de P par Q . On suppose donc dans la suite que $\deg P < \deg Q$.

1.3.2 Multiplicité 1, degré 1

Regardons le cas où les zéros de Q sont tous de multiplicité 1 et qu'il n'y a pas de partie irréductible de degré 2. Typiquement

$$F = \frac{3X^2 - 7X + 5}{(X + 1)X(X - 2)}.$$

On sait que F va se décomposer sous la forme

$$F = \frac{a}{X + 1} + \frac{b}{X} + \frac{c}{(X - 2)}.$$

Si on multiplie à gauche et à droite par $(X + 1)$ on obtient

$$\frac{3X^2 - 7X + 5}{X(X - 2)} = a + \frac{b(X + 1)}{X} + \frac{c(X + 1)}{(X - 2)}.$$

En prenant $X = -1$, on trouve

$$\frac{15}{3} = a,$$

i.e. $a = 5$.

De la même façon, en multipliant par X et en faisant $X = 0$ on trouve $b = -5/2$. Enfin, en multipliant par $(X - 2)$ et en faisant $X = 2$, on obtient $c = 1/2$. D'où

$$\frac{3X^2 - 7X + 5}{(X + 1)X(X - 2)} = \frac{5}{X + 1} + \frac{-5/2}{X} + \frac{1/2}{(X - 2)}.$$

1.3.3 Multiplicité quelconque, degré 1

Ici on regarde toujours le cas où les parties irréductibles sont de degré 1, mais avec multiplicité. La meilleure technique ici est celle des développements limités.

Par exemple, si

$$F(X) = \frac{4X^2 - 3X + 1}{X(X - 1)^2}$$

on sait qu'on aura une décomposition en

$$F(X) = \frac{a}{X} + \frac{b}{(X - 1)} + \frac{c}{(X - 1)^2}.$$

On va déterminer b et c . Notez que dans cette méthode cela se fait indépendamment de la connaissance ou non de a (que l'on obtient facilement d'ailleurs par la méthode ci-dessus : multiplication par X , puis $X = 0$).

On multiplie tout par $(X - 1)^2$ et on pose $X = 1 + h$:

$$\frac{4(1 + h)^2 - 3(1 + h) + 1}{1 + h} = \frac{a}{1 + h}h^2 + bh + c = c + bh + o(h).$$

On fait un développement limité à l'ordre 1 du membre de gauche (en $h = 0$) :

$$(4 + 8h - 3 - 3h + 1 + o(h))(1 - h + o(h)) = 2 + 3h + o(h).$$

Ce qui donne directement $b = 3$ et $c = 2$.

On obtient $a = 1$ par la méthode du cas précédent, ce qui donne finalement

$$\frac{4X^2 - 3X + 1}{X(X-1)^2} = \frac{1}{X} + \frac{3}{(X-1)} + \frac{2}{(X-1)^2}.$$

Un autre exemple

$$F(X) = \frac{1}{X^3(X+1)^2} = \frac{a}{X} + \frac{b}{X^2} + \frac{c}{X^3} + \frac{d}{X+1} + \frac{e}{(X+1)^2}.$$

On multiplie par X^3 et on pose $X = h$:

$$\frac{1}{(h+1)^2} = ah^2 + bh + c + \frac{d}{h+1}h^3 + \frac{e}{(h+1)^2}h^3 = ah^2 + bh + c + o(h^2).$$

On effectue un D.L. à l'ordre 2 du membre de gauche :

$$\frac{1}{(h+1)^2} = 1 - (2h + h^2) + (2h + h^2)^2 + o(h^2) = 1 - 2h + 3h^2 + o(h^3).$$

Ce qui donne $a = 3, b = -2, c = 1$.

On multiplie par $(X+1)^2$ et on pose $X = -1 + h$:

$$\frac{1}{(-1+h)^3} = \frac{a}{(-1+h)}h^2 + \frac{b}{(-1+h)^2}h^2 + \frac{c}{(-1+h)^3}h^2 + \frac{d}{h} + e = e + dh + o(h).$$

On fait un D.L. à l'ordre 1 :

$$\frac{1}{(-1+h)^3} = -1 - 3h + o(h).$$

Ce qui donne $e = -1, d = -3$. Finalement

$$\frac{1}{X^3(X+1)^2} = \frac{3}{X} + \frac{-2}{X^2} + \frac{1}{X^3} + \frac{-3}{X+1} + \frac{-1}{(X+1)^2}.$$

Il y a une méthode sans D.L., mais plus longue ; la voici.

$$\frac{4X^2 - 3X + 1}{X(X-1)^2} = \frac{a}{X} + \frac{b}{(X-1)} + \frac{c}{(X-1)^2}.$$

On multiplie par $(X-1)^2$ et on fait $X = 1$:

$$\frac{4X^2 - 3X + 1}{X} = \frac{a}{X}(X-1)^2 + b(X-1) + c,$$

et finalement $c = 2$.

On retranche $2/(X-1)^2$ des deux côtés :

$$\frac{4X^2 - 3X + 1}{X(X-1)^2} - \frac{2}{(X-1)^2} = \frac{a}{X} + \frac{b}{(X-1)},$$

i.e.

$$\frac{4X^2 - 5X + 1}{X(X-1)^2} = \frac{a}{X} + \frac{b}{(X-1)}.$$

Le membre de gauche doit se simplifier par $X-1$... c'est le cas

$$\frac{4X-1}{X(X-1)} = \frac{a}{X} + \frac{b}{(X-1)}$$

et là on sait faire, par substitution ! ($a=1$, $b=3$)

De la même façon, on fait le deuxième exemple. On part de la forme

$$F(X) = \frac{1}{X^3(X+1)^2} = \frac{a}{X} + \frac{b}{X^2} + \frac{c}{X^3} + \frac{d}{X+1} + \frac{e}{(X+1)^2}.$$

On multiplie par X^3

$$\frac{1}{(X+1)^2} = aX^2 + bX + c + \frac{d}{X+1}X^3 + \frac{e}{(X+1)^2}X^3$$

et on fait $X=0$

$$1 = c.$$

On soustrait $1/X^3$

$$\begin{aligned} \frac{1}{X^3(X+1)^2} - \frac{1}{X^3} &= \frac{1 - (X+1)^2}{X^3(X+1)^2} = \frac{-X-2}{X^2(X+1)^2} \\ \frac{-X-2}{X^2(X+1)^2} &= \frac{a}{X} + \frac{b}{X^2} + \frac{d}{X+1} + \frac{e}{(X+1)^2} \end{aligned}$$

et on recommence : multiplication par X^2 , $X=0$, donnent

$$b = -2.$$

Ainsi de suite ...

$$\frac{-X-2}{X^2(X+1)^2} - \frac{-2}{X^2} = \frac{2X+3}{X(X+1)^2}$$

d'où

$$a = 3$$

puis

$$\frac{2X+3}{X(X+1)^2} - \frac{3}{X} = \frac{-3X-4}{(X+1)^2}$$

etc.

1.3.4 Cas général

Maintenant on incorpore les termes irréductibles de degré 2.
Commençons par

$$F(X) = \frac{1}{(X^2 + 1)(X - 1)} = \frac{aX + b}{X^2 + 1} + \frac{c}{X - 1},$$

avec a, b, c réels. On utilise la méthode de multiplication, substitution avec une des racines complexes de $X^2 + 1$:

$$\frac{1}{(X - 1)} = \frac{aX + b}{X^2 + 1} + \frac{c}{X - 1},$$

et en faisant $X = i$

$$-\frac{1}{2}(i + 1) = ai + b.$$

Comme a, b sont réels, on a directement $a = b = -1/2$. Ensuite on obtient $c = 1/2$ par la même méthode.

Dans un cas de multiplicité

$$F(X) = \frac{X}{(X^2 + 1)^2(X - 1)^2} = \frac{aX + b}{X^2 + 1} + \frac{cX + d}{(X^2 + 1)^2} + \frac{e}{X - 1},$$

On fait la méthode par substitutions successives (faire des D.L. dans \mathbb{C} est un peu délicat et hors programme) :

$$F(X) = \frac{X}{(X - 1)^2} = (aX + b)(X^2 + 1) + cX + d + \frac{e}{X - 1}(X^2 + 1)^2,$$

puis $X = i$

$$-\frac{1}{2} = ci + d,$$

d'où $c = 0, d = -1/2$. On soustrait

$$\frac{X}{(X^2 + 1)^2(X - 1)^2} - \frac{-1/2}{(X^2 + 1)^2} = \frac{1/2}{(X^2 + 1)(X - 1)^2}.$$

On trouve facilement $a = 0, b = 1/4$, et ainsi de suite ...

Chapitre 2

Espaces vectoriels

On aborde un chapitre absolument essentiel dans votre formation ; des notions vraiment nouvelles, pas toujours faciles et qui pourtant vont vous suivre durant absolument toutes vos études et pour certains durant toute leur carrière.

Avant de les attaquer je veux expliquer et motiver ce chapitre. En mathématiques, une des grandes activités, à un niveau élevé comme celui que vous atteignez, on ne regarde plus seulement des cas explicites (vecteurs dans $\mathbb{R}^2, \mathbb{R}^3$, fonctions réelles, convergence dans \mathbb{C} etc, mais on isole des structures générales dans lesquelles des résultats identiques se reproduisent, on établit des théorèmes généraux qui peuvent ensuite être utilisés dans plein de contextes différents.

Un exemple. Vous avez tous appris au lycée les vecteurs dans \mathbb{R}^2 et peut être \mathbb{R}^3 , vous avez appris leur structure de combinaison linéaires, peut-être même abordé certaines applications linéaires (projection, symétrie, etc), vous avez parlé de produit scalaire, de norme. Par exemple, vous savez que

$$|\vec{v} \cdot \vec{w}| \leq \|\vec{v}\| \|\vec{w}\| .$$

Pouvez-vous imaginer que cela a un rapport avec une inégalité (que vous ne connaissez pas, mais qui est vraie et très utile) du genre suivant : pour toute fonction f, g continue sur $[a, b]$ on a

$$\left(\int_a^b f(x)g(x) dx \right)^2 \leq \left(\int_a^b f(x)^2 dx \right) \left(\int_a^b g(x)^2 dx \right) ?$$

Apparemment pas ... et pourtant c'est la même chose, le même théorème, qui a trait aux espaces vectoriels munis d'un produit scalaire et de la norme associée. C'est le théorème de Schwarz.

Quel est le lien entre les vecteurs du plan et les systèmes d'équations linéaires à 5 inconnues ? Quelles sont les maths utilisées par l'algorithme PageRank de Google pour classer les sites web par importance ? Quelles sont les maths utilisées en mécanique quantique, en théorie quantique de l'information, pour effectuer des téléportations quantiques ou bien des ordinateurs quantiques ? La même chose : les espaces vectoriels, les applications linéaires, les matrices, les notions de vecteurs, de bases, de dimensions etc.

2.1 Structure d'espace vectoriel

Un espace vectoriel se définit d'abord sur un *corps* fixé. Cette notion n'est pas en détail au programme pour vous, il faut juste savoir en gros ce que c'est.

Définition 7 *Un corps est un ensemble \mathbb{K} muni de deux lois internes : une addition et une multiplication commutatives, associatives, la multiplication est distributive sur l'addition, il y a des éléments neutres 0 et 1 pour ces opérations et tout élément sauf 0 est inversible pour la multiplication.*

Dans la suite on ne travaillera que sur les corps \mathbb{R} ou \mathbb{C} , éventuellement \mathbb{Q} .

Définition 8 *Un espace vectoriel sur le corps \mathbb{K} est un ensemble E muni d'une loi d'addition interne +*

$$\begin{aligned} + & : E \times E \rightarrow E \\ & (x, y) \mapsto x + y. \end{aligned}$$

et d'une loi de multiplication externe \cdot

$$\begin{aligned} \cdot & : \mathbb{K} \times E \rightarrow E \\ & (\lambda, x) \mapsto \lambda \cdot x \end{aligned}$$

par le corps \mathbb{K} et qui vérifie les axiomes suivants.

- l'addition est commutative ($x + y = y + x$)
- l'addition admet un élément neutre 0_E ($0_E + x = x$)
- tout élément de E admet un inverse (appelé opposé) pour l'addition ($x + (-x) = 0_E$)
- pour tous $\alpha, \beta \in \mathbb{K}$, $x, y \in E$ on a

$$\begin{aligned} (\alpha + \beta) \cdot x &= \alpha \cdot x + \beta \cdot x \\ \alpha \cdot (\beta \cdot x) &= (\alpha\beta) \cdot x \\ \alpha \cdot (x + y) &= \alpha \cdot x + \alpha \cdot y \\ 1_{\mathbb{K}} \cdot x &= x \end{aligned}$$

On dit que E est un \mathbb{K} -espace vectoriel, les éléments de E sont appelés des vecteurs, le vecteur 0_E est le vecteur nul, les éléments de \mathbb{K} sont appelés des scalaires.

Tout de suite des exemples.

– Le plan $E = \mathbb{R}^2$ avec les lois habituelles d'addition des vecteurs et de multiplication par un scalaire :

$$\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x + x' \\ y + y' \end{pmatrix} ; \quad \lambda \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} ; \quad 0_E = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

– De la même façon, tous les ensembles du type \mathbb{R}^n (espace vectoriel sur \mathbb{R}), ou \mathbb{C}^n (espace vectoriel sur \mathbb{C}), avec les lois d'addition coordonnée par coordonnée et de multiplication scalaire aussi par coordonnée.

– Une extension de la structure sur \mathbb{R}^n est celle de l'espace vectoriel $E = \mathbb{R}^{\mathbb{N}}$, i.e. l'ensemble des suites (u_n) dans \mathbb{R} . Les lois d'espace vectoriel sont alors

$$(u_n) + (v_n) = (u_n + v_n) ; \quad \lambda \cdot (u_n) = (\lambda u_n) ; \quad 0_E = (0).$$

Je dis que c'est une extension du cas \mathbb{R}^n car on voit bien qu'ici la suite (u_n) est traitée comme un vecteur

$$u = \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_n \\ \vdots \end{pmatrix}$$

avec une infinité de coordonnées.

– Les espaces de fonctions admettent bien souvent des structures d'espace vectoriel définie comme suit. Par exemple prenons $E = \mathbb{R}[X]$. Les "vecteurs" de E sont donc des polynômes. Les lois sont

$$P + Q = R \text{ avec } R(X) = P(X) + Q(X) ; \quad \lambda \cdot P = S \text{ avec } S(X) = \lambda P(X) ;$$

et 0_E est le polynôme nul.

Notez bien qu'ici, si on écrit dans E une relation du type $\lambda \cdot P + \mu \cdot Q = 0_E$ (par exemple, juste pour situer), alors cela signifie que le polynôme $\lambda \cdot P + \mu \cdot Q$ est le polynôme nul, i.e. que $\lambda P(X) + \mu Q(X) = 0$ pour **tout** X .

– On a les mêmes structures d'addition et de multiplication scalaire sur des espaces comme $C([0, 1])$, $\mathbb{C}_4[X]$, $C^\infty(\mathbb{R})$, etc.

$$[f + g](x) = f(x) + g(x) ; \quad [\lambda \cdot f](x) = \lambda f(x) ; \quad [0_E](x) = 0, \quad \forall x.$$

Il y a beaucoup d'axiomes à vérifier pour montrer qu'une structure $(E, +, \cdot)$ est un espace vectoriel. Mais il y a une situation courante pour laquelle il y a peu de chose à vérifier : c'est quand votre ensemble F est un sous-ensemble d'un espace vectoriel E déjà identifié. Dans ce cas F hérite de la structure d'espace vectoriel de E , avec assez peu de conditions.

Définition 9 Soit E un espace vectoriel. Soit F un sous-ensemble de E . On dit que F est un sous-espace vectoriel (dans la suite s.e.v.) de E si

- $0_E \in F$,
- pour tous $\lambda, \mu \in \mathbb{K}$, tous $x, y \in F$ on a $\lambda \cdot x + \mu \cdot y \in F$.

Théorème 2.1.1 Tout s.e.v. est un espace vectoriel.

La démonstration est facile et laissée au lecteur.

Typiquement, les sous-espaces vectoriels de $E = \mathbb{R}^3$ sont

- $F = \{0\}$
- Toutes les droites passant par 0
- Tous les plans passant par 0
- $F = E$ tout entier.

Et c'est tout ! (il est facile de voir que ce sont des s.e.v. de E , par contre vous ne savez pas encore montrer que ce sont les seuls).

Dans la suite on arrête de noter le \cdot du produit par un scalaire. On note λx au lieu de $\lambda \cdot x$.

2.2 Opérations sur les espaces vectoriels

Proposition 2.2.1 Soit E un espace vectoriel et F, G deux s.e.v. de E . Alors $F \cap G$ est un s.e.v. de E .

Démonstration Il suffit donc de vérifier que $F \cap G$ vérifie les deux propriétés qui caractérisent un s.e.v.

Trivialement $0_E \in F \cap G$ puisqu'il appartient aux deux espaces F, G .

Maintenant, soit $\lambda, \mu \in \mathbb{K}$ et soit $x, y \in F \cap G$. Comme x, y appartiennent à F on a $\lambda x + \mu y \in F$. De même $\lambda x + \mu y$ appartient à G . Donc finalement $\lambda x + \mu y$ appartient à $F \cap G$. \square

Il est important de noter que la même propriété n'est plus vraie pour l'union de s.e.v. En effet, il suffit de considérer n'importe quelles 2 droites passant par 0, dans \mathbb{R}^2 ; leur union n'est pas un s.e.v., elle n'est pas stable par somme de vecteurs.

Définition 10 Pour tout s.e.v. F et G de E on définit

$$F + G = \{f + g; f \in F, g \in G\}.$$

De la même façon on définit la somme $\sum_{i=1}^n F_i$ d'un nombre quelconque de s.e.v. de E .

Proposition 2.2.2 $F + G$ est un s.e.v. de E .

Démonstration Comme $0_E = 0_E + 0_E$, on a $0_E \in F + G$.

Soient $x, y \in F + G$ et $\lambda, \mu \in \mathbb{K}$. On sait en particulier que $x = f + g$ et $y = f' + g'$ avec $f, f' \in F$ et $g, g' \in G$. On a donc

$$\lambda x + \mu y = (\lambda f + \mu f') + (\lambda g + \mu g').$$

Ce qui montre que $\lambda x + \mu y$ est bien un élément de $F + G$.

On a prouvé que $F + G$ est un s.e.v. de E . □

Définition 11 Soit A une partie quelconque d'un espace vectoriel E . On appelle espace vectoriel engendré par A l'intersection de tous les s.e.v. de E qui contiennent A . On le note $\text{Vect}(A)$.

Ce s.e.v. existe car il y a toujours au moins l'espace E qui contient A .

Ce s.e.v. est le plus petit s.e.v. de E qui contient A .

Proposition 2.2.3

$$\text{Vect}(F \cup G) = F + G.$$

Démonstration L'espace $\text{Vect}(F \cup G)$ est le plus petit s.e.v. de E qui contient $F \cup G$; d'autre part $F + G$ est un s.e.v. qui contient F et G . Donc $\text{Vect}(F \cup G)$ est inclus dans $F + G$.

L'espace $\text{Vect}(F \cup G)$ est un espace vectoriel qui contient F et qui contient G . Donc il contient tous les éléments de E de la forme $f + g$, $f \in F$, $g \in G$. Donc il contient $F + G$.

On a prouvé l'égalité des deux ensembles. □

Définition 12 Soient F_1, \dots, F_n des s.e.v. de E , on dit que la somme $F = \sum_{i=1}^n F_i$ est directe si pour tout $x \in F$ il existe une **unique** décomposition $x = x_1 + \dots + x_n$, où $x_i \in F_i$, $i = 1, \dots, n$.

Dans ce cas on note cette somme $F = \bigoplus_{i=1}^n F_i$, ou encore $F_1 \oplus \dots \oplus F_n$.

Proposition 2.2.4 La somme $F = \sum_{i=1}^n F_i$ est directe si et seulement si la seule décomposition possible de 0_E sous la forme

$$0_E = x_1 + \dots + x_n, \text{ avec } x_i \in F_i$$

est donnée par $x_i = 0_E$ pour tout i .

Démonstration Si la somme est directe, la décomposition de 0_E est unique.

Inversement, si la décomposition de x dans $\sum_i F_i$ n'est pas unique, alors en faisant la différence des deux décompositions, on obtient une décomposition non triviale de 0_E . \square

Définition 13 Deux sous-espaces F et G de E sont dits supplémentaires si $E = F \oplus G$.

Autrement dit, si et seulement si pour tout $x \in E$ il existe un unique couple $(f, g) \in F \times G$ tel que $x = f + g$.

2.3 Bases, dimension

Définition 14 Soit E un \mathbb{K} -espace vectoriel, soient v_1, \dots, v_n des vecteurs de E , soient $\lambda_1, \dots, \lambda_n$ des scalaires. On dit que le vecteur $v = \sum_{i=1}^n \lambda_i v_i$ est une combinaison linéaire des vecteurs v_i .

Pour une famille infinie $(v_i)_{i \in I} \subset E$, les combinaisons linéaires des v_i sont les combinaisons linéaires **finies** d'éléments de cette famille.

Définition 15 Une famille $\{v_1, \dots, v_n\} \subset E$ est dite génératrice dans E si tout vecteur $v \in E$ est combinaison linéaire des v_i .

Autrement dit, si $\text{Vect}(v_1, \dots, v_n) = E$.

Pour une famille infinie $(v_i)_{i \in I}$, la définition est la même : $\text{Vect}(\{v_i; i \in I\}) = E$, i.e. les combinaisons linéaires finies des v_i engendrent tous les vecteurs de E .

Par exemple :

- Deux vecteurs non colinéaires sont générateurs dans \mathbb{R}^2 .
- Un exemple de 4 vecteurs générateurs dans \mathbb{R}^3
- Les vecteurs X^k , $k \in \mathbb{N}$ forment une famille génératrice de $\mathbb{R}[X]$.
- Les vecteurs $X^2 + X + 1$ et $3X^2 - X - 2$ ne sont pas générateurs de $\mathbb{R}_2[X]$. Ca engendre seulement les polynômes $\alpha X^2 + \beta X + \gamma$ tels que $\alpha - 5\beta + 4\gamma = 0$. Donc, par exemple pas X^2

Définition 16 Une famille $\{v_1, \dots, v_n\}$ de E est dite libre si la seule combinaison linéaire des v_i qui vaut 0 est la combinaison nulle. Autrement dit, si $\sum_{i=1}^n \lambda_i v_i = 0$ entraîne $\lambda_i = 0$ pour tout i .

Dans le cas d'une famille infinie, la définition est la même, la seule combinaison linéaire finie des v_i qui vaut 0 est la combinaison nulle.

Par exemple :

- Deux vecteurs non colinéaires sont libres dans \mathbb{R}^2 .
- Un exemple de 2 vecteurs libres dans \mathbb{R}^3
- Les vecteurs X^k , $k \in \mathbb{N}$ forment une famille libres de $\mathbb{R}[X]$.
- Les vecteurs $X^2 + X + 1$, $3X^2 - X - 2$, $X + 1$, $X^2 - 7X + 4$ ne sont pas libres dans $\mathbb{R}_2[X]$.

Définition 17 Une famille de vecteurs $\{v_1, \dots, v_n\}$ (ou bien $\{v_i; i \in I\}$) est une base de E si elle est à la fois libre et génératrice.

Théorème 2.3.1 Si $\mathcal{V} = \{v_1, \dots, v_n\}$ (ou $\{v_i; i \in I\}$) est une base de E , alors pour tout $x \in E$ il existe une unique décomposition

$$x = \sum_{i=1}^n \lambda_i v_i.$$

Définition 18 Ces uniques λ_i associés à x sont appelés coordonnées de x dans la base \mathcal{V} .

Définition 19 Si E possède une base de cardinal fini, on dit que E est de dimension finie.

Théorème 2.3.2 Si E est de dimension finie, toutes les bases de E ont même cardinal.

Admis (un peu long et laborieux, mais pas compliqué au fond).

Définition 20 Si E est de dimension finie, ce cardinal unique des bases de E est appelé dimension de E .

Si E n'est pas de dimension finie, on dit qu'il est de dimension infinie.

Théorème 2.3.3 (Théorème de la base incomplète) Soit E un \mathbb{K} -espace vectoriel de dimension fini.

- 1) De toute famille génératrice de E on peut extraire une base de E . En particulier toute famille génératrice a un cardinal $\geq \dim E$.
- 2) Toute famille libre dans E peut être complétée en une base de E . En particulier toute famille libre a un cardinal $\leq \dim E$.

Démonstration

1) Soit $\mathcal{W} = \{w_1, \dots, w_m\}$ une famille génératrice de E . Si elle est libre, c'est que c'est une base et c'est fini. Sinon, c'est qu'un des w_i s'écrit comme combinaison linéaire des autres. Disons que c'est w_m . On le retire à la liste \mathcal{W} qui reste génératrice ! Et ainsi de suite jusqu'à qu'on trouve une sous-famille libre, qui sera génératrice par construction. Le procédé s'arrête forcément, c'est clair.

2) Soit $\mathcal{V} = \{v_1, \dots, v_p\}$ une famille libre dans E . On se donne une famille $\mathcal{W} = \{w_1, \dots, w_m\}$ génératrice de E et on applique l'algorithme suivant. Si $w_1 \in \text{Vect}(\mathcal{V})$ on ne fait rien, si $w_1 \notin \text{Vect}(\mathcal{V})$, on remplace \mathcal{V} par $\mathcal{V} \cup \{w_1\}$. Cette nouvelle famille est libre aussi (car dire qu'elle est liée revient à dire que $w_1 \in \text{Vect}(\mathcal{V})$).

Ainsi de suite on répète l'opération jusqu'à obtenir une famille libre et génératrice (ce qui arrivera forcément car si on épuise \mathcal{W} on a bien une famille génératrice. \square)

Théorème 2.3.4 *Soit E un espace vectoriel de dimension finie. Soit $n = \dim E$. Il y a équivalence entre les assertions suivantes.*

- i) $\{e_1, \dots, e_n\}$ est une base de E .
- ii) $\{e_1, \dots, e_n\}$ est une famille libre de E .
- iii) $\{e_1, \dots, e_n\}$ est une famille génératrice de E .

Démonstration i) implique ii) est trivial. Voyons ii) implique i). Si $\{e_1, \dots, e_n\}$ est une famille libre de E , alors on peut compléter cette famille en une base de E : $\{e_1, \dots, e_n, e_{n+1}, \dots, e_m\}$. Mais toutes les bases de E doivent avoir le même cardinal, donc $m = n$ et donc $\{e_1, \dots, e_n\}$ était déjà une base de E .

Pour l'équivalence entre i) et iii) c'est le même genre de raisonnement. \square

2.4 Dimension des sous-espaces

Théorème 2.4.1 *Si E est un espace vectoriel de dimension finie et si F est un sous-espace vectoriel de E , alors*

$$\dim F \leq \dim E.$$

Et on a $\dim F = \dim E$ si et seulement si $F = E$.

Démonstration Soit $\{f_1, \dots, f_p\}$ une base de F , c'est une famille libre dans F , donc libre dans E , donc $p \leq \dim E$.

Si $\dim F = \dim E$, ça veut dire qu'il existe $\{f_1, \dots, f_n\}$ base de F , en particulier libre dans E . Donc c'est une base de E et du coup $E = \text{Vect}(\{f_1, \dots, f_n\}) = F$. \square

Théorème 2.4.2 (Formule de Grassmann) *Si F, G sont des sous-espaces de E alors*

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G).$$

Démonstration On se donne une base $\{h_1, \dots, h_q\}$ de $F \cap G$. On complète cette famille en une base de $F : \{h_1, \dots, h_q, f_1, \dots, f_{n-q}\}$. On fait de même pour $G : \{h_1, \dots, h_q, g_1, \dots, g_{m-q}\}$.

Les vecteurs f_i (ainsi que leurs combinaisons linéaires) sont dans F mais pas dans G (car sinon ils seraient dans $F \cap G$ et ne seraient pas libres avec les h_j). De la même manière les vecteurs g_i (et leur combinaisons linéaires) sont dans G mais pas dans F .

Considérons la famille $\{h_1, \dots, h_q, f_1, \dots, f_{n-q}, g_1, \dots, g_{m-q}\}$. Elle est libre, car si par exemple f_i est combinaison linéaire des autres, on rassemble les f_i dans la combinaison linéaire et on trouve que cela forme un élément de G . Ce qui ne peut pas être le cas, on l'a vu. Cette famille est incluse dans $F + G$, elle est donc libre dans $F + G$.

Elle est aussi génératrice de $F + G$ car elle permet d'écrire tous les éléments de F ainsi que tous les éléments de G , donc tous les éléments de $F + G$.

C'est donc une base de $F + G$. Il suffit maintenant de vérifier les dimensions :

$$q + n - q + m - q = n + m - q.$$

□

Corollaire 2.4.3 *Si F et G sont en somme directe dans E , alors*

$$\dim(F \oplus G) = \dim F + \dim G.$$

Théorème 2.4.4 *En dimension finie, tout sous-espace F de E admet un supplémentaire. Et tous les supplémentaires de F ont même dimension.*

Démonstration On prend une base $\{f_1, \dots, f_p\}$ de F , on la complète en une base $\{f_1, \dots, f_p, e_1, \dots, e_{n-p}\}$ de E . L'espace $G = \text{Vect}(\{e_1, \dots, e_{n-p}\})$ est un supplémentaire de F car $F + G = E$ et $F \cap G = \{0\}$.

Pour la dimension c'est évident avec ce qui a été obtenu plus haut. □

Chapitre 3

Applications linéaires

3.1 Définitions

Soient E, F deux espaces vectoriels sur \mathbb{K} , on appelle *application linéaire de E dans F* toute application $u : E \mapsto F$ qui vérifie

$$u(\lambda x + \mu y) = \lambda u(x) + \mu u(y),$$

pour tout $x, y \in E$, $\lambda, \mu \in \mathbb{K}$.

On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F . C'est un espace vectoriel avec les opérations usuelles sur les fonctions.

Lorsque $F = \mathbb{K}$ on dit que u est une *forme linéaire*. L'espace $\mathcal{L}(E, \mathbb{K})$ est alors noté E^* et appelé *dual de E* .

Lorsque $F = E$ on dit que u est un *endomorphisme* de E . L'espace $\mathcal{L}(E, E)$ est alors noté $\mathcal{L}(E)$.

Si $u : E \mapsto F$ est linéaire et bijective on dit que c'est un *isomorphisme* et l'espace des isomorphismes est noté $GL(E, F)$.

Si $u : E \mapsto E$ est linéaire et bijective on dit que c'est un *automorphisme* et l'espace des automorphismes est noté $GL(E)$.

Proposition 3.1.1 *Si $u \in \mathcal{L}(E, F)$ alors $u(0_E) = 0_F$.*

Démonstration

$$u(0_E) = u(0x) = 0u(x) = 0_F.$$

□

Proposition 3.1.2

- 1) *L'image d'un s.e.v. par une application linéaire est un s.e.v.*
- 2) *L'image réciproque d'un s.e.v. par une application linéaire est un s.e.v.*

3) Si $u \in \mathcal{L}(E, F)$ et $A \subset E$ alors

$$u(\text{Vect}(A)) = \text{Vect}(u(A)).$$

Démonstration

1) Soit G un s.e.v de E , si $z_1, z_2 \in u(G)$ et si $\lambda_1, \lambda_2 \in \mathbb{K}$ alors $z_i = u(x_i)$ pour un $x_i \in E$ et

$$\lambda_1 z_1 + \lambda_2 z_2 = \lambda_1 u(x_1) + \lambda_2 u(x_2) = u(\lambda_1 x_1 + \lambda_2 x_2).$$

Comme G est un s.e.v. on a $\lambda_1 x_1 + \lambda_2 x_2 \in G$ et donc $\lambda_1 z_1 + \lambda_2 z_2 \in u(G)$. On a montré que $u(G)$ est un s.e.v.

2) Soit H un s.e.v de F , si $x_1, x_2 \in u^{-1}(H)$ et si $\lambda_1, \lambda_2 \in \mathbb{K}$ alors $u(x_i) \in H$ et

$$u(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 u(x_1) + \lambda_2 u(x_2)$$

appartient à H car c'est un s.e.v. Donc $\lambda_1 z_1 + \lambda_2 z_2 \in u^{-1}(H)$. On a montré que $u^{-1}(H)$ est un s.e.v.

3) On a $A \subset \text{Vect}(A)$ donc $u(A) \subset u(\text{Vect}(A))$ et finalement $\text{Vect}(u(A)) \subset u(\text{Vect}(A))$.

Inversement, si $x \in u(\text{Vect}(A))$, alors $x = u(\lambda_1 a_1 + \dots + \lambda_n a_n)$ pour des $\lambda_i \in \mathbb{K}$ et des $a_i \in A$. Donc $x = \lambda_1 u(a_1) + \dots + \lambda_n u(a_n)$ et $x \in \text{Vect}(u(A))$.
□

3.2 Noyau et image

On appelle *noyau* de u l'ensemble

$$\text{Ker } u = \{x \in E; u(x) = 0_F\}.$$

C'est un s.e.v. de E , d'après la proposition ci-dessus.

On appelle *image* de u l'ensemble

$$\text{Im } u = \{u(x); x \in E\}.$$

C'est un s.e.v. de F .

Proposition 3.2.1

- 1) u est injective si et seulement si $\text{Ker } u = \{0\}$.
- 2) u est surjective si et seulement si $\text{Im } u = F$.

Démonstration

2) est évidente, c'est la définition même de la surjectivité.

1) Si u est injective et si $x \in \text{Ker } u$ alors $u(x) = 0$. Mais on a aussi $u(0) = 0$, donc avec l'injectivité on a $x = 0$ et on a montré que $\text{Ker } u$ est réduit à $\{0\}$.

Inversement si $\text{Ker } u = \{0\}$ et si $u(x) = u(y)$ alors $u(x - y) = 0$ et $x - y \in \text{Ker } u$. D'où $x - y = 0$ et $x = y$. On a montré que u est injective. \square

Proposition 3.2.2

1) Si u est surjective et si $\{e_1, \dots, e_p\}$ est génératrice dans E , alors $\{u(e_1), \dots, u(e_p)\}$ est génératrice dans F .

2) Si u est injective et si $\{e_1, \dots, e_p\}$ est libre dans E , alors $\{u(e_1), \dots, u(e_p)\}$ est libre dans F .

Démonstration

1) Si $y \in F$ alors $y = u(x)$ pour un $x \in E$. Ce x s'écrit $x = \sum_i \lambda_i e_i$ donc $y = u(x) = \sum_i \lambda_i u(e_i)$.

2) Si $\sum_i \lambda_i u(e_i) = 0$ alors $u(\sum_i \lambda_i e_i) = 0$ et donc $\sum_i \lambda_i e_i = 0$. ce qui implique $\lambda_i = 0$ pour tout i . \square

Théorème 3.2.3 Soit $\{e_1, \dots, e_n\}$ une base de E et $u \in \mathcal{L}(E, F)$.

1) u est injective si et seulement si $\{u(e_1), \dots, u(e_n)\}$ est libre.

2) u est surjective si et seulement si $\{u(e_1), \dots, u(e_n)\}$ est génératrice.

3) u est bijective si et seulement si $\{u(e_1), \dots, u(e_n)\}$ est une base.

Démonstration Pour 1) et 2) on a déjà fait un sens. Avec 1) et 2) on a 3) trivialement.

1) Si $\{u(e_1), \dots, u(e_n)\}$ est libre et si $u(x) = 0$ alors $x = \sum_i \lambda_i e_i$ et $\sum_i \lambda_i u(e_i) = 0$. Donc les λ_i sont tous nuls et $x = 0$.

2) Si $\{u(e_1), \dots, u(e_n)\}$ est génératrice et si $y \in F$ alors $y = \sum_i \lambda_i u(e_i)$, donc $y = u(\sum_i \lambda_i e_i)$ et donc $y \in \text{Im } u$. \square

3.3 Construction d'applications linéaires

Théorème 3.3.1 Si $\{e_1, \dots, e_n\}$ est une base de E et si $\{v_1, \dots, v_n\} \subset F$, alors il existe une unique application linéaire u telle $u(e_i) = v_i$ pour tout i .

Démonstration Pour tout $x \in E$, on a $x = \sum_i \lambda_i e_i$ (et cette écriture est unique!), on pose alors $u(x) = \sum_i \lambda_i v_i$. On vérifie facilement que c'est linéaire. L'unicité est facile aussi. \square

Corollaire 3.3.2 *Si deux applications linéaires coïncident sur une base, alors elles sont égales.*

Deux espaces vectoriels de même dimension sont isomorphes.

3.4 Rang

On appelle *rang de u* la dimension de $\text{Im } u$.

$$\text{rg } u = \dim \text{Im } u .$$

Proposition 3.4.1

Si $u \in \mathcal{L}(E, F)$ alors $\text{rg } u \leq \dim E$, avec égalité si et seulement si u est injective.

Si $u \in \mathcal{L}(E, F)$ alors $\text{rg } u \leq \dim F$, avec égalité si et seulement si u est surjective.

Démonstration

Soit $\{e_1, \dots, e_n\}$ une base de E , alors $\text{Im } u = u(\text{Vect}(e_1, \dots, e_n)) = \text{Vect}(u(e_1), \dots, u(e_n))$. Donc la dim de $\text{Im } u$ est $\leq n$. On a égalité si et seulement si $\dim \text{Vect}(u(e_1), \dots, u(e_n)) = n$, c'est à dire u injective.

Le fait que $\text{rg } u \leq \dim F$ est trivial. Le cas d'égalité est évidemment le cas surjectif. \square

Proposition 3.4.2 *Si $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, G)$, alors $\text{rg } v \circ u \leq \min\{\text{rg } u, \text{rg } v\}$.*

Démonstration On a

$$\text{rg } v \circ u = \dim \text{Im } v \circ u = \dim v(u(E)) \leq \dim v(F) = \text{rg } v .$$

D'autre part, $v \circ u = v|_{u(E)}$ et en appliquant la proposition ci-dessus $\text{rg } v \circ u \leq \dim u(E)$. \square

Théorème 3.4.3 (Théorème du rang) *Soit $u \in \mathcal{L}(E, F)$ et soit S un supplémentaire de $\text{Ker } u$ dans E . Alors la restriction de u à S est un isomorphisme de S dans $\text{Im } u$.*

Démonstration Notons $v = u|_S$. Alors $v : S \mapsto \text{Im } u$ et

$$\text{Ker } v = \{x \in S ; u(x) = 0\} = \text{Ker } u \cap S = \{0\} .$$

Donc v est injective.

Soit $y \in \text{Im } u$, alors $y = u(x)$ et se décompose $x = a + b$ avec $a \in \text{Ker } u$ et $b \in S$. Donc $y = u(a) + u(b) = u(b) = v(b)$. Donc $y \in \text{Im } v$. On a montré que $\text{Im } v = \text{Im } u$. Ainsi v est surjective. Finalement v est une bijection de S dans $\text{Im } u$. \square

Théorème 3.4.4 (Théorème du rang) *Si $u \in \mathcal{L}(E, F)$ alors*

$$\dim E = \dim \text{Ker } u + \text{rg } u.$$

Théorème 3.4.5 (Théorème d'isomorphisme) *Si $\dim E = \dim F = n$ et si $u \in \mathcal{L}(E, F)$ alors on a équivalence entre*

i) *u est un isomorphisme*

ii) *u est injective*

iii) *u est surjective*

iv) *$\text{rg } u = n$*

v) *$\exists v : F \mapsto E ; v \circ u = \text{Id}_E$*

v) *$\exists w : F \mapsto E ; u \circ w = \text{Id}_F$*

Dans ce cas on a

$$w = v = u^{-1}.$$

Démonstration

i) implique ii) trivial

ii) implique iii) c'est le théorème du rang : $n = \text{rg } u + \dim \text{Ker } u$ donc $\text{Ker } u = \{0\}$ implique $\text{rg } u = n$.

iii) implique ii) c'est aussi le théorème du rang.

iii) implique i) parce que iii) implique ii) et que iii)+ii) impliquent i).

iii) et iv) sont évidemment équivalents.

i) implique clairement v) et vi).

v) implique ii) car si $v \circ u$ est injectif alors u est injectif.

vi) implique iii) car si $u \circ w$ est surjectif alors u est surjectif. □

Proposition 3.4.6 *Une application linéaire $u : E \mapsto F$ n'est jamais inversible si $\dim E \neq \dim F$.*

Démonstration Si $\dim E > \dim F$, par le théorème du rang $\dim \text{Ker } u = \dim E - \text{rg } u \geq \dim E - \dim F > 0$. Donc u n'est pas injective.

Si $\dim E < \dim F$, par le théorème du rang $\text{rg } u = \dim E - \dim \text{Ker } u \leq \dim E < \dim F$. Donc u n'est pas surjective. □

Chapitre 4

Calcul matriciel

4.1 Représentation matricielle des applications linéaires

On se donne deux espaces vectoriels E et F sur le même corps \mathbb{K} ; ils seront de dimension finie n et m respectivement, avec des bases fixées $\mathcal{E} = \{e_1, \dots, e_n\}$ et $\mathcal{F} = \{f_1, \dots, f_m\}$, respectivement.

Soit $u : E \mapsto F$ une application linéaire. Pour tout $i = 1, \dots, n$, on a $u(e_i)$ qui est un élément de F , donc il s'écrit dans la base \mathcal{F}

$$u(e_i) = \sum_{j=1}^m u_{ji} f_j,$$

pour des $u_{ji} \in \mathbb{K}$.

Tout vecteur $v \in E$ admet des coordonnées dans la base $\mathcal{E} : v = \sum_{i=1}^n v_i e_i$, que l'on note

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{\mathcal{E}}, \quad \text{ou } v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

si le choix de base est clair. On a alors

$$\begin{aligned} u(v) &= \sum_{i=1}^n v_i u(e_i) \\ &= \sum_{i=1}^n \sum_{j=1}^m v_i u_{ji} f_j \\ &= \sum_{j=1}^m \left(\sum_{i=1}^n u_{ji} v_i \right) f_j. \end{aligned}$$

Autrement dit,

$$u(v) = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}_{\mathcal{F}}$$

avec

$$y_i = \sum_{k=1}^n u_{ik} v_k.$$

On voit qu'on a ainsi une méthode pour calculer l'image de tout vecteur par u .

1) On construit un tableau de nombre $\{u_{ij}; i = 1, \dots, n; j = 1, \dots, m\}$, appelée *matrice*, de la forme suivante

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ u_{21} & u_{22} & u_{23} & \dots & u_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u_{m1} & u_{m2} & u_{m3} & \dots & u_{mn} \end{pmatrix}$$

Le tableau a n colonnes et m lignes.

2) Ses colonnes sont constituées des coordonnées des $u(e_i)$ dans la base \mathcal{F} :

$$u(e_1) = \begin{pmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{m1} \end{pmatrix}, \quad u(e_i) = \begin{pmatrix} u_{1i} \\ u_{2i} \\ \vdots \\ u_{mi} \end{pmatrix}$$

3) Cette matrice agit sur les vecteurs de longueur n et les transforme en vecteurs de longueur m , car u agit sur E de dimension n et les transforme en vecteurs de F , de dimension m .

Elle agit de la façon suivante :

$$u(v) = \begin{pmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ u_{21} & u_{22} & u_{23} & \dots & u_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u_{m1} & u_{m2} & u_{m3} & \dots & u_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

avec

$$y_i = \sum_{j=1}^n u_{ij} v_j$$

pour tout $i = 1, \dots, m$.

La matrice ainsi obtenue est appelée *matrice associée à u dans les bases \mathcal{E} et \mathcal{F}* . On la note $\text{Mat}_{\mathcal{E},\mathcal{F}}(u)$.

On note $\mathcal{M}_{m,n}(\mathbb{K})$ l'ensemble des matrices à m lignes et n colonnes. Dans le cas où $m = n$ on note ça simplement $\mathcal{M}_n(\mathbb{K})$. Notez que $\mathcal{M}_{m,n}(\mathbb{K})$ est l'espace des matrices qui vont d'un espace de dimension n vers un espace de dimension m .

Proposition 4.1.1 *Soient $u, v \in \mathcal{L}(E, F)$, avec des bases \mathcal{E}, \mathcal{F} , de matrice associée $(u_{ij}), (v_{ij})$, alors pour tout $\lambda, \mu \in \mathbb{K}$, la matrice associée à $\lambda u + \mu v$ a pour coefficients*

$$\lambda u_{ij} + \mu v_{ij}.$$

Soient $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, G)$, de bases respectives $\mathcal{E}, \mathcal{F}, \mathcal{G}$, de dimensions respectives n, m, p et de matrices respectives (u_{ij}) et (v_{ij}) , alors la matrice de $w = v \circ u \in \mathcal{L}(E, G)$ a pour matrice dans les bases \mathcal{E}, \mathcal{G} :

$$w_{ij} = \sum_{k=1}^m v_{ik} u_{kj},$$

$$i = 1, \dots, p, j = 1, \dots, n.$$

4.2 Opérations sur les matrices

On l'a vu ci-dessus, les ensembles $\mathcal{M}_{n,m}(\mathbb{K})$ sont des \mathbb{K} -espaces vectoriels. On a une loi d'addition, une loi de multiplication scalaire, un zéro (qui est la matrice nulle).

Les éléments de $\mathcal{M}_{n,m}(\mathbb{K})$ agissent sur des vecteurs à m coordonnées (ou encore des matrices dans $\mathcal{M}_{m,1}(\mathbb{K})$) pour en faire des vecteurs à n coordonnées ($\in \mathcal{M}_{n,1}$), de la façon suivante

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1m} \\ u_{21} & u_{22} & u_{23} & \dots & u_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u_{n1} & u_{n2} & u_{n3} & \dots & u_{nm} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

avec

$$y_i = \sum_{j=1}^m u_{ij} v_j$$

pour tout $i = 1, \dots, n$.

On peut "multiplier" à droite des matrices $A \in \mathcal{M}_{n,m}(\mathbb{K})$ par des matrices $B \in \mathcal{M}_{m,p}(\mathbb{K})$ grâce au produit qui est donné par la composition des applications linéaires :

$$(AB)_{ij} = \sum_{k=1}^m A_{ik} B_{kj}.$$

Ce qui donne une matrice $AB \in \mathcal{M}_{n,p}(\mathbb{K})$.

Notez qu'en général, on ne peut pas considérer aussi le produit BA , si on n'a pas $n = m = p$.

Sur les espaces $\mathcal{M}_n(\mathbb{K})$ on a aussi un produit interne (qui fait que $\mathcal{M}_n(\mathbb{K})$ est ce qu'on appelle une *algèbre*). Il s'agit du produit qui est donné par la composition des applications linéaires :

$$(AB)_{ij} = \sum_{k=1}^n A_{ik} B_{kj}.$$

Notez qu'en général on a

$$AB \neq BA.$$

Une matrice particulière dans $\mathcal{M}_n(\mathbb{K})$

$$\text{Id}_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

C'est l'élément neutre de $\mathcal{M}_n(\mathbb{K})$ pour le produit matriciel.

Sur $\mathcal{M}_n(\mathbb{K})$ on a une notion d'inverse de matrice quand ça existe. On dit que la matrice B est *l'inverse* de $A \in \mathcal{M}_n(\mathbb{K})$ (quand ça existe) si

$$AB = BA = \text{Id}_n.$$

Dans ce cas on note $B = A^{-1}$. Notez que B est forcément unique et qu'il suffit de prouver $AB = \text{Id}_n$ ou bien $BA = \text{Id}_n$ pour avoir l'autre égalité.

Deuxième partie

ANALYSE

Chapitre 5

Développements limités

Nous attaquons ici un chapitre fondamental de l'analyse : celui où l'on va aborder les développements limités. Il s'agit d'un outil incontournable pour l'étude fine du comportement des fonctions, c'est un outil redoutablement efficace pour le calcul de limites difficiles.

5.1 Comparaison de fonctions

Dans ce qui suit a est soit un réel, soit $\pm\infty$.

Définition 21 *On dit qu'une fonction f est négligeable devant une fonction g au voisinage de a , s'il existe une fonction ε telle que $\lim_{x \rightarrow a} \varepsilon(x) = 0$ et*

$$f(x) = \varepsilon(x)g(x).$$

On écrit alors que $f = o(g)$ au voisinage de a , ou encore $f = o_a(g)$.

Par exemple $x^3 = o(x^2)$ au voisinage de 0, puisque $x^3 = x \times x^2$ et que $\varepsilon(x) = x$ tend vers 0 en 0. Ou encore $x^2 = o(x^5)$ au voisinage de $+\infty$, puisque $x^2 = x^5 \times 1/x^3$.

Je vous rappelle ici un théorème que vous avez vu au 1er semestre, le théorème des croissances comparées, exprimé ici en terme de négligeabilité.

Théorème 5.1.1 *Quelques soient $\alpha, \beta, \gamma > 0$ on a*

$$(\ln(x))^\alpha = o(x^\beta) \quad \text{et} \quad x^\beta = o(e^{\gamma x}),$$

au voisinage de $+\infty$ et on a

$$|\ln(x)|^\alpha = o(x^{-\beta})$$

au voisinage de 0.

Quelques propriétés et manipulations avec les o .

Proposition 5.1.2

- 1) Au voisinage de a , si $f = o(g)$ et si $g = o(h)$ alors $f = o(h)$.
- 2) Au voisinage de a , si $f_1 = o(g_1)$ et $f_2 = o(g_2)$ alors $f_1 f_2 = o(g_1 g_2)$.
- 3) Au voisinage de a , si $f_1 = o(g)$ et $f_2 = o(g)$ alors $f_1 + f_2 = o(g)$.
- 4) Au voisinage de a , si $f = o(g)$ alors $1/g = o(1/f)$.

Toutes les démonstrations des résultats ci-dessus sont évidentes et laissées au lecteur. Par contre, notez bien qu'il ne faut pas faire les erreurs suivantes.

– Si $f_1 = o(g_1)$ et $f_2 = o(g_2)$ alors il n'est pas vrai que $f_1 + f_2 = o(g_1 + g_2)$. En effet, au voisinage de 0, on a $x^2 = o(x)$ et $-x^3 = o(-x + x^2)$ mais par contre $x^2 - x^3$ n'est pas un $o(x^2)$.

– Avec les quotients de fonctions rien de général ne marche.

Notez qu'avec nos notations on écrit $f = o(1)$ au voisinage de a pour dire f tend vers 0 en a .

5.2 Equivalence de fonctions

Définition 22 On dit que f est équivalente à g au voisinage de a si

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 1.$$

On note cela $f \sim_a g$, ou bien $f \sim g$ au voisinage de a . Notez que c'est équivalent à $f = g + o(g)$ au voisinage de a .

Théorème 5.2.1

- 1) La relation \sim au voisinage de a est une relation d'équivalence entre fonctions : elle est symétrique, réflexive et transitive.
- 2) Si $f \sim_a g$ alors $\lim_{x \rightarrow a} f(x)$ existe si et seulement si $\lim_{x \rightarrow a} g(x)$ existe. Dans ce cas les limites sont égales.
- 3) On peut prendre le produit, le quotient (si bien défini) et les puissances des équivalents.
- 4) Si $\lim_{x \rightarrow b} \phi(x) = a$ et si $f \sim_a g$ alors $f \circ \phi \sim_b g \circ \phi$.

Démonstration

- 1) Le fait que $f \sim_a f$ est une évidence.

Si $f \sim_a g$ alors $f(x) = (1 + \varepsilon(x))g(x)$ où $\lim_{x \rightarrow a} \varepsilon(x) = 0$. Mais dans ce cas, en posant

$$\varepsilon'(x) = \frac{\varepsilon(x)}{1 + \varepsilon(x)}$$

on a $\lim_{x \rightarrow a} \varepsilon'(x) = 0$ et $g(x) = (1 - \varepsilon'(x))f(x)$. Ainsi $g \sim_a f$.

Enfin, si $f \sim_a g$ et $g \sim_a h$ alors $f = (1 + \varepsilon)g$ et $g = (1 + \varepsilon')h$ d'où $f = (1 + \varepsilon + \varepsilon' + \varepsilon\varepsilon')h$ ce qui donne le résultat facilement.

2) On a $g(x) = (1 + \varepsilon(x))f(x)$ donc $\lim_{x \rightarrow a} g(x) = l$.

4) On a $f(\phi(x)) = (1 + \varepsilon(\phi(x)))g(\phi(x))$, mais $\lim_{x \rightarrow b} \varepsilon(\phi(x)) = \lim_{x \rightarrow a} \varepsilon(x) = 0$.

3) Si $f_1 \sim_a g_1$ et $f_2 \sim_a g_2$ alors

$$\begin{aligned} f_1(x)f_2(x) &= (1 + \varepsilon_1(x))(1 + \varepsilon_2(x))g_1(x)g_2(x) \\ &= (1 + \varepsilon_1(x) + \varepsilon_2(x) + \varepsilon_1(x)\varepsilon_2(x))g_1(x)g_2(x). \\ \frac{f_1(x)}{f_2(x)} &= \frac{1 + \varepsilon_1(x)}{1 + \varepsilon_2(x)} \frac{g_1(x)}{g_2(x)} \\ &= (1 + \varepsilon_1(x))(1 - \varepsilon_2'(x)) \frac{g_1(x)}{g_2(x)}. \end{aligned}$$

On passe facilement aux puissances $n \in \mathbb{Z}$ en itérant le résultat ci-dessus, du coup on passe aux puissances rationnelles (si les fonctions sont > 0). Pour les puissances quelconques, si f et g sont > 0 et $f \sim_a g$ alors

$$f(x)^\alpha = e^{\alpha \ln(f(x))} = e^{\alpha \ln((1+\varepsilon(x))g(x))} = e^{\alpha \ln(1+\varepsilon(x))} e^{\alpha \ln(g(x))}.$$

La limite de $e^{\alpha \ln(1+\varepsilon(x))}$ quand $x \rightarrow a$ est clairement 1, d'où le résultat. \square

Les erreurs à ne pas faire avec les équivalents :

– En général on ne peut pas additionner (ou soustraire) des équivalents. Par exemple $x^2 - x \sim_0 -x$ et $x \sim_0 x$, par contre $x^2 \not\sim_0 0$.

– On ne peut pas composer les équivalents par une fonction à gauche. Par exemple $x^2 + x \sim_{+\infty} x^2$, mais

$$\lim_{x \rightarrow +\infty} \frac{e^{x^2+x}}{e^{x^2}} = \lim_{x \rightarrow +\infty} e^x = +\infty$$

et donc $e^{x^2+x} \not\sim_{+\infty} e^{x^2}$.

Théorème 5.2.2 (Équivalents importants en 0)

$$\begin{array}{ll}
e^x \sim 1 + x, & e^x = 1 + x + o(x) \\
\ln(1+x) \sim x, & \ln(1+x) = x + o(x) \\
\sin(x) \sim x, & \sin(x) = x + o(x^2) \\
\cos(x) \sim 1 - \frac{x^2}{2}, & \cos(x) = 1 - \frac{x^2}{2} + o(x^3) \\
\tan(x) \sim x, & \tan(x) = x + o(x^2) \\
\frac{1}{(1+x)^\alpha} \sim 1 + \alpha x, & \frac{1}{(1+x)^\alpha} = 1 + \alpha x + o(x).
\end{array}$$

Pour l'instant nous donnons ces résultats sans preuve, ces formules seront nettement améliorées (et prouvées) dans la suite.

5.3 Les formules de Taylor

La formule de Taylor et ses variantes, sont des formules qui disent que les bonnes fonctions (C^n , par exemple) peuvent être correctement approchées par des polynômes, en tout cas localement près d'un point, et qu'en plus on sait mesurer assez précisément l'erreur que l'on commet en faisant cette approximation.

5.3.1 Un rappel concernant la dérivation

Rappelons un résultat du premier semestre.

Théorème 5.3.1 *Une fonction f définie au voisinage de x_0 est dérivable en x_0 si et seulement si*

$$f(x_0 + h) = f(x_0) + hf'(x_0) + o(h),$$

ou encore

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + o(x - x_0).$$

Par exemple :

$$\begin{aligned}
\left(1 - \frac{x}{n}\right)^n &= e^{n \log(1 - \frac{x}{n})} \\
&= e^{n(-\frac{x}{n} + o(\frac{1}{n}))} \\
&= e^{-x + o(1)}
\end{aligned}$$

et donc

$$\lim_{n \rightarrow +\infty} \left(1 - \frac{x}{n}\right)^n = e^{-x}.$$

5.3.2 La formule de Taylor-Young

Lemme 5.3.2 Soit $k \in \mathbb{N}$ et g une fonction dérivable au voisinage de 0. Si $g'(h) = o(h^k)$ alors $g(h) - g(0) = o(h^{k+1})$.

Démonstration Par hypothèse $g'(h) = \mathcal{E}(h)h^k$ au voisinage de 0, où \mathcal{E} est une fonction qui tend vers 0 en 0. En particulier, pour tout $\varepsilon > 0$, il existe $h_0 > 0$ tel que pour $|h| \leq h_0$ on a $|g'(h)| \leq \varepsilon |h|^k$. Notez qu'alors, pour tout $0 \leq |h'| \leq |h|$ on a du coup $|g'(h')| \leq \varepsilon |h|^k$. Par l'inégalité des accroissements finis on a alors

$$|g(h) - g(0)| \leq \varepsilon |h|^{k+1}.$$

En d'autres termes, la fonction $\widehat{\mathcal{E}}(h) = (g(h) - g(0))/h^{k+1}$ tend vers 0 et donc $g(h) - g(0) = o(h^{k+1})$. \square

Théorème 5.3.3 (Formule de Taylor-Young) Soit f une fonction définie et n fois dérivable au voisinage de x_0 . Alors au voisinage de x_0 on

$$f(x) = f(x_0) + \sum_{k=1}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + o((x - x_0)^n).$$

Démonstration Nous le montrons par récurrence sur n . On a vu qu'elle est vraie pour $n = 1$. Supposons-la vraie au rang $n - 1$, on l'applique à f' :

$$\begin{aligned} f'(x) &= f'(x_0) + \sum_{k=1}^{n-1} \frac{(f')^{(k)}(x_0)}{k!} (x - x_0)^k + o((x - x_0)^{n-1}) \\ &= f'(x_0) + \sum_{k=1}^{n-1} \frac{f^{(k+1)}(x_0)}{k!} (x - x_0)^k + o((x - x_0)^{n-1}). \end{aligned}$$

On applique le lemme ci-dessus à la fonction

$$g(h) = f(x_0 + h) - f(x_0) - \sum_{k=1}^n \frac{f^{(k)}(x_0)}{k!} h^k.$$

On a

$$\begin{aligned} g'(h) &= f'(x_0 + h) - \sum_{k=1}^n \frac{f^{(k)}(x_0)}{(k-1)!} h^{k-1} \\ &= f'(x_0 + h) - \sum_{k=0}^{n-1} \frac{f^{(k+1)}(x_0)}{k!} h^k \\ &= o(h^{n-1}). \end{aligned}$$

et comme $g(0) = 0$ on conclut que $g(h) = o(h^n)$ par le lemme précédent. \square

5.3.3 Développements limités

On dit qu'une fonction f définie au voisinage de x_0 admet un *développement limité à l'ordre n en x_0* si il existe des coefficients a_0, \dots, a_n tels que

$$f(x) = \sum_{k=0}^n a_k (x - x_0)^k + o((x - x_0)^n)$$

au voisinage de x_0 .

Attention ! On a vu que l'existence d'un développement limité à l'ordre 1 est équivalent à la dérivabilité. Cette propriété ne s'étend pas aux ordres supérieurs à 1 : ce n'est pas parce qu'une fonction admet un développement limité à l'ordre n qu'elle est n fois dérivable. Voyons cela sur un exemple.

La fonction $f(x) = |x|^{5/2} \sin(1/x)$ est définie sur \mathbb{R}^* , elle est C^∞ sur \mathbb{R}^* et elle vérifie $|f(x)| \leq |x|^{5/2}$. Elle est donc prolongeable par continuité en 0 par une fonction f telle que $f(0) = 0$. Regardons la dérivabilité en 0. On a

$$\frac{f(x) - f(0)}{x} = \frac{|x|^{5/2}}{x} \sin(1/x)$$

dont la limite est 0 quand x tend vers 0. Donc f est dérivable en 0 et $f'(0) = 0$.

Pour les autres valeurs de x on a

$$f'(x) = \begin{cases} \frac{5}{2} |x|^{3/2} \sin(1/x) - |x|^{1/2} \cos(1/x) & x > 0, \\ -\frac{5}{2} |x|^{3/2} \sin(1/x) - |x|^{1/2} \cos(1/x) & x < 0. \end{cases}$$

En particulier

$$\frac{f'(x) - f'(0)}{x} = \frac{5}{2} |x|^{1/2} \sin(1/x) - \frac{|x|^{1/2}}{x} \cos(1/x).$$

A cause du second terme cette expression n'a pas de limite en 0. Notre fonction n'est pas deux fois dérivable en 0. Pourtant

$$|f(x)| \leq |x|^{5/2}$$

donc

$$f(x) = o(x^2).$$

La fonction f admet bien un développement limité d'ordre 2 en 0.

Revenons sur les développements limités en général. Nous allons démontrer l'unicité du développement limité.

Lemme 5.3.4 Soit P un polynôme de degré n . Si $P(x) = o(x^n)$ au voisinage de 0 alors $P = 0$.

Démonstration Soit $P(x) = \sum_{k=0}^n a_k x^k$ le développement de P . Supposons qu'au moins un des a_k est non nul. Soit p le plus petit indice tel que $a_k \neq 0$. Alors en divisant par x^p on a

$$a_p = - \sum_{k=p+1}^n a_k x^{k-p} + o(1).$$

Le membre de droite tend vers 0 quand x tend vers 0 , donc $a_p = 0$. Contradiction. Tous les a_k sont donc nuls. \square

Proposition 5.3.5 Soit f une fonction définie au voisinage de x_0 . Si il existe des coefficients a_0, \dots, a_n et b_0, \dots, b_n tels que

$$f(x) = \sum_{k=0}^n a_k (x - x_0)^k + o((x - x_0)^n) = \sum_{k=0}^n b_k (x - x_0)^k + o((x - x_0)^n)$$

alors $a_k = b_k$ pour tout $k = 0, \dots, n$.

Démonstration On fait la différence des deux développements et on applique le lemme précédent. \square

Le résultat suivant montre que le développement limité de f à l'ordre n est la meilleure approximation de f au voisinage de 0 par un polynôme de degré n .

Proposition 5.3.6 Soit f une fonction définie au voisinage de x_0 et possédant un développement limité à l'ordre n en x_0 , de la forme $f(x) = P(x - x_0) + o((x - x_0)^n)$. Alors, pour tout polynôme Q de degré n , il existe $\delta > 0$ tel que

$$|f(x) - P(x - x_0)| \leq |f(x) - Q(x - x_0)|$$

pour tout x tel que $|x - x_0| \leq \delta$.

Démonstration Si $P = Q$ alors le résultat est évident. On suppose donc $P \neq Q$.

On sait que $|f(x) - P(x - x_0)| = o((x - x_0)^n)$. D'autre part, comme P et Q sont différents on a

$$P(x - x_0) - Q(x - x_0) = a_p (x - x_0)^p + o((x - x_0)^p)$$

pour un $a_p \in \mathbb{R}^*$ et un $p \leq n - 1$. En particulier, on a

$$\begin{aligned} f(x) - Q(x - x_0) - a_p(x - x_0)^p &= \\ &= f(x) - P(x - x_0) + P(x - x_0) - Q(x - x_0) - a_p(x - x_0)^p \\ &= o((x - x_0)^p). \end{aligned}$$

Donc il existe $\delta_1 > 0$ tel que

$$|f(x) - Q(x - x_0)| \geq |a_p(x - x_0)^p| - |o((x - x_0)^p)| \geq \frac{|a_p|}{2} |x - x_0|^p$$

pour tout x tel que $|x - x_0| \leq \delta_1$. D'autre part

$$|f(x) - P(x - x_0)| = o((x - x_0)^n) = \mathcal{E}(x) |x - x_0|^{n-p} |x - x_0|^p$$

donc il existe aussi $\delta_2 > 0$ tel que

$$|f(x) - P(x - x_0)| \leq \frac{|a_p|}{2} |x - x_0|^p$$

pour tout x tel que $|x - x_0| \leq \delta_2$. En prenant $\delta = \min\{\delta_1, \delta_2\}$ on a le résultat annoncé. \square

5.3.4 Développement des fonctions usuelles

En application directe de la Formule de Taylor-Young on assez facilement les développements suivants, qu'il faut connaître par coeur.

Théorème 5.3.7 *Au voisinage de 0 on a*

$$e^x = \sum_{k=0}^n \frac{x^k}{k!} + o(x^n),$$

autrement dit

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + o(x^n).$$

On a

$$\ln(1 + x) = \sum_{k=1}^n \frac{(-1)^{k-1}}{k} x^k + o(x^n),$$

autrement dit

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots + \frac{(-1)^{n-1}}{n} x^n + o(x^n).$$

On a

$$(1+x)^\alpha = \sum_{k=0}^n \frac{\alpha(\alpha-1)(\alpha-2)\dots(\alpha-k+1)}{k!} x^k + o(x^n),$$

autrement dit

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2} x^2 + \dots + \frac{\alpha(\alpha-1)(\alpha-2)\dots(\alpha-n+1)}{n!} x^n + o(x^n).$$

On a

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

autrement dit

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{4!} + \dots + \frac{(-1)^n}{(2n)!} x^{2n} + o(x^{2n+1}).$$

On a

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

autrement dit

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots + \frac{(-1)^n}{(2n+1)!} x^{2n+1} + o(x^{2n+2}).$$

On a

$$\operatorname{ch}(x) = \sum_{k=0}^n \frac{1}{(2k)!} x^{2k} + o(x^{2n+1}),$$

autrement dit

$$\operatorname{ch}(x) = 1 + \frac{x^2}{2} + \frac{x^4}{4!} + \dots + \frac{1}{(2n)!} x^{2n} + o(x^{2n+1}).$$

On a

$$\text{sh}(x) = \sum_{k=0}^n \frac{1}{(2k+1)!} x^{2k+1} + o(x^{2n+2}),$$

autrement dit

$$\text{sh}(x) = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots + \frac{1}{(2n+1)!} x^{2n+1} + o(x^{2n+2}),$$

5.3.5 Taylor-Lagrange, reste intégral

Sous certaines conditions un peu plus fortes on peut dire plus de choses sur le reste dans la formule de Taylor-Young.

Théorème 5.3.8 (Formule de Taylor-Lagrange) *Si f est C^n sur $[a, b]$ et $n+1$ fois dérivable sur $]a, b[$, alors il existe $c \in]a, b[$ tel que*

$$f(b) = f(a) + \sum_{k=1}^n \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(c).$$

Démonstration Soit

$$g(x) = f(b) - f(x) - \sum_{k=1}^n \frac{(b-x)^k}{k!} f^{(k)}(x) - C \frac{(b-x)^{n+1}}{(n+1)!},$$

où C est une constante choisie telle que $g(a) = 0$ (ce qui est possible car le facteur devant C est non nul en $x = a$).

On a aussi $g(b) = 0$ clairement. La fonction g est continue sur $[a, b]$ et dérivable sur $]a, b[$. On peut donc appliquer le théorème de Rolle : il existe $c \in]a, b[$ tel que $g'(c) = 0$. Mais on a aussi

$$\begin{aligned} g'(x) &= -f'(x) - \sum_{k=1}^n \left(\frac{-(b-x)^{k-1}}{(k-1)!} f^{(k)}(x) + \frac{(b-x)^k}{k!} f^{(k+1)}(x) \right) + \\ &\quad + C \frac{(b-x)^n}{n!} \\ &= -\frac{(b-x)^n}{n!} f^{(n+1)}(x) + C \frac{(b-x)^n}{n!}. \end{aligned}$$

En particulier

$$0 = g'(c) = -\frac{(b-c)^n}{n!} f^{(n+1)}(c) + C \frac{(b-c)^n}{n!}$$

ce qui donne

$$C = f^{(n+1)}(c).$$

Du coup quand on écrit à nouveau $g(a) = 0$ on obtient la formule annoncée.
□

On a du coup une jolie estimation de l'erreur.

Théorème 5.3.9 (Inégalité de Taylor-Lagrange) *Si f est C^n sur $[a, b]$ et $n + 1$ fois dérivable sur $]a, b[$, alors si $f^{(n+1)}$ est bornée sur $]a, b[$ on a*

$$\left| f(b) - f(a) - \sum_{k=1}^n \frac{(b-a)^k}{k!} f^{(k)}(a) \right| \leq \frac{(b-a)^{n+1}}{(n+1)!} \sup_{x \in [a, b]} |f^{(n+1)}(x)|.$$

Enfin on termine par une autre formule exacte pour le reste.

Théorème 5.3.10 (Formule de Taylor avec reste intégral) *Si f est de classe C^{n+1} sur $[a, b]$ alors*

$$f(b) = f(a) + \sum_{k=1}^n \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t) dt.$$

Démonstration La démonstration se fait par récurrence sur n . Pour $n = 0$, la fonction f est C^1 sur $[a, b]$ et donc

$$f(b) = f(a) + \int_a^b f'(t) dt.$$

Donc l'assertion est vraie.

Supposons-la vraie au rang $n - 1$. Nous avons donc

$$f(b) = f(a) + \sum_{k=1}^{n-1} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-t)^{n-1}}{(n-1)!} f^{(n)}(t) dt.$$

On fait une intégration par parties dans l'intégrale :

$$\begin{aligned} \int_a^b \frac{(b-t)^{n-1}}{(n-1)!} f^{(n)}(t) dt &= \left[-\frac{(b-t)^n}{n!} f^{(n)}(t) \right]_a^b + \int_a^b \frac{(b-t)^n}{(n)!} f^{(n+1)}(t) dt \\ &= \frac{(b-a)^n}{n!} f^{(n)}(a) + \int_a^b \frac{(b-t)^n}{(n)!} f^{(n+1)}(t) dt. \end{aligned}$$

Ce qui permet de terminer la récurrence. □

5.4 Propriétés des développements limités

5.4.1 Opérations sur les développements limités

Proposition 5.4.1 *Si f et g sont deux fonctions définies au voisinage de x_0 qui admettent des développements limités à l'ordre n et p*

$$f(x_0 + h) = \sum_{k=0}^n a_k h^k + o(h^n), \quad g(x_0 + h) = \sum_{k=0}^p b_k h^k + o(h^p),$$

alors pour tous $\lambda, \mu \in \mathbb{R}$ les fonctions $\lambda f + \mu g$ et fg admettent un développement limité à l'ordre $q = \min\{n, p\}$

$$(\lambda f + \mu g)(x_0 + h) = \sum_{k=0}^q (\lambda a_k + \mu b_k) h^k + o(h^q),$$

$$(fg)(x_0 + h) = \sum_{k=0}^q c_k h^k + o(h^q)$$

où

$$c_k = \sum_{j=0}^k a_j b_{k-j}.$$

Démonstration Supposons pour simplifier que $n \leq p$. On a donc

$$(\lambda f + \mu g)(x_0 + h) = \sum_{k=0}^n (\lambda a_k + \mu b_k) h^k + \sum_{k=n+1}^p \mu b_k h^k + o(h^n) + o(h^p).$$

Les trois derniers termes du membre de droite sont des $o(h^n)$. D'où le résultat pour la somme.

De même pour le produit :

$$f(x_0+h)g(x_0+h) = \left(\sum_{k=0}^n a_k h^k \right) \left(\sum_{l=0}^p b_l h^l \right) + o(h^n)g(x_0+h) + o(h^p)f(x_0+h).$$

Les fonctions f et g sont continues en x_0 donc bornées au voisinage de x_0 , d'où

$$f(x_0 + h)g(x_0 + h) = \sum_{k=0}^n \sum_{l=0}^n a_k b_l h^{k+l} + o(h^n).$$

Dans les puissances h^{k+l} qui apparaissent dans la double somme ci-dessus on ne garde que les puissances $\leq n$ et on regroupe

$$f(x_0 + h)g(x_0 + h) = \sum_{k=0}^n \sum_{j=0}^k a_j b_{k-j} h^k + o(h^n).$$

□

Proposition 5.4.2 *Soit g une fonction définie au voisinage de x_0 et f une fonction définie au voisinage de $b = g(x_0)$. Si g admet un développement limité à l'ordre p en x_0 et si f admet un développement limité à l'ordre n en b alors $f \circ g$ admet un développement limité à l'ordre $q = \min\{n, p\}$ obtenu en substituant les développements limités et en tronquant à l'ordre q .*

Démonstration Comme g est continue en x_0 on a $\lim_{h \rightarrow 0} g(x_0 + h) = g(x_0) = b$. Donc

$$\begin{aligned} f(g(x_0 + h)) &= f(b + (g(x_0 + h) - b)) \\ &= \sum_{k=0}^q a_k (g(x_0 + h) - b)^k + o((g(x_0 + h) - b)^q) \\ &= \sum_{k=0}^q a_k \left(\sum_{l=1}^q b_l h^l \right)^k + o((h + o(h))^q) \end{aligned}$$

d'où le résultat. □

Proposition 5.4.3 *Soient f et g deux fonctions définies au voisinage de x_0 et admettant un développement limité à l'ordre n et p respectivement. Si $g(x_0) \neq 0$ alors f/g admet un développement limité à l'ordre $q = \min\{n, p\}$.*

Démonstration La fonction $x \mapsto 1/x$ admet un développement limité de tout ordre en tout point $b_0 \neq 0$. Donc par le théorème de composition des développements limités la fonction $1/g$ admet un développement limité en 0, à l'ordre p . Ensuite on applique le résultat sur les produits de développements limités. □

Proposition 5.4.4 *Soit f une fonction dérivable définie au voisinage de x_0 et telle que f' admette un développement limité à l'ordre n*

$$f'(x_0 + h) = \sum_{k=0}^n a_k h^k + o(h^n).$$

Alors f admet un développement limité à l'ordre $n + 1$

$$f(x_0 + h) = f(x_0) + \sum_{k=1}^{n+1} \frac{a_{k-1}}{k} h^k + o(h^{n+1}).$$

Démonstration Soit

$$g(h) = f(x_0 + h) - f(x_0) - \sum_{k=1}^{n+1} \frac{a_{k-1}}{k} h^k.$$

Par hypothèse

$$g'(h) = f'(x_0 + h) - \sum_{k=0}^n a_k h^k = o(h^n).$$

On applique alors le lemme 5.3.2. □

Proposition 5.4.5 Soit f une fonction dérivable définie au voisinage de x_0 admettant un développement limité à l'ordre n

$$f(x_0 + h) = \sum_{k=0}^n a_k h^k + o(h^n).$$

Si f' admet un développement limité à l'ordre $n - 1$ alors ce développement est de la forme

$$f'(x_0 + h) = \sum_{k=1}^n k a_k h^{k-1} + o(h^{n-1}).$$

Démonstration On applique la proposition précédente à f' . □

Attention! Notez bien qu'ici on doit supposer l'existence d'un développement limité pour f' car celui-ci n'est pas garanti par l'existence du développement limité de f . Par exemple $f(x) = |x|^{5/2} \sin(1/x)$ est dérivable et admet un développement limité à l'ordre 2. Mais sa dérivée n'admet pas de développement limité à l'ordre 1.

5.4.2 Comportement local près des points critiques

On rappelle que pour une fonction réelle f dérivable, un *point critique* est un point x_0 tel que $f'(x_0) = 0$. Ces points peuvent être de 3 natures différentes : un minimum local, un maximum local, ni l'un ni l'autre (du type x^3 en 0).

Proposition 5.4.6 *Soit f définie au voisinage de x_0 . Si f admet un développement limité au voisinage de x_0 de la forme*

$$f(x) = f(x_0) + \alpha(x - x_0)^p + o((x - x_0)^p)$$

avec $\alpha \neq 0$ et $p \geq 2$. Alors

- si p est impair le point x_0 n'est ni un minimum, ni un maximum,
- si p est pair et $\alpha > 0$ alors x_0 est un minimum local strict,
- si p est pair et $\alpha < 0$ alors x_0 est un maximum local strict.

Démonstration Si p est pair et $\alpha > 0$, alors on a

$$\frac{\alpha}{2}(x - x_0)^p + o((x - x_0)^p) \geq 0$$

dès que $|x - x_0|$ est assez petit pour que

$$|o((x - x_0)^p)| \leq \frac{\alpha}{4}(x - x_0)^p.$$

Donc on a

$$f(x) \geq f(x_0) + \frac{\alpha}{2}(x - x_0)^p > f(x_0)$$

pour x suffisamment proche de x_0 et différent de x_0 . Cela prouve que x_0 est un minimum local.

Le cas $\alpha < 0$ se traite exactement de la même façon.

Dans le cas p impair, supposons $\alpha > 0$ (l'autre cas se traite de manière analogue). De la même façon que ci-dessus on a

$$f(x) \geq f(x_0) + \frac{\alpha}{2}(x - x_0)^p > f(x_0)$$

pour x suffisamment proche de x_0 et $x > x_0$. Mais lorsque $x < x_0$ on a

$$f(x) \geq f(x_0) + \frac{\alpha}{2}(x - x_0)^p < f(x_0).$$

Ainsi on voit bien que x_0 n'est ni un minimum local, ni un maximum local.

□

Proposition 5.4.7 *Soit f définie au voisinage de x_0 . Si f admet un développement limité au voisinage de x_0 de la forme*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \alpha(x - x_0)^p + o((x - x_0)^p)$$

avec $\alpha \neq 0$ et $p \geq 2$. Alors

- si p est impair le graphe de f traverse sa tangente en $(x_0, f(x_0))$ (x_0 est un point d'inflexion),
- si p est pair et $\alpha > 0$ le graphe de f reste localement au-dessus de sa tangente en $(x_0, f(x_0))$,
- si p est pair et $\alpha < 0$ le graphe de f reste localement en-dessous de sa tangente en $(x_0, f(x_0))$.

Chapitre 6

Intégration

L'intégrale est un des plus beaux et des plus puissants objet mathématique. Il s'agit sans aucun doute d'une des plus belles inventions de l'esprit humain. En effet, il s'agit tout d'abord d'une pure création de l'esprit au sens où c'est un objet limite, obtenu en passant à la limite sur des subdivisions etc, pas un objet qui existe dans la nature (ou du moins ça se discute!). Ensuite c'est un objet qui permet de calculer des choses très compliquées : pratiquement toutes les surfaces. Imaginez comment on s'y prenait il y a plusieurs siècles pour calculer des surfaces compliquées : on ne dispose en général que de peu d'outils ou de formules, on est obligé de passer par des approximations par des figures simples (comme les grecs avec le disque qu'ils encadraient par des polygones réguliers). L'intégrale fournit une réponse très puissante et extrêmement simple : il suffit de calculer une primitive et de prendre la différence de sa valeur en 2 points ! C'est vraiment surprenant de simplicité.

Il n'y a pas un seul pan des sciences, qu'elles soient physiques, chimiques, biologiques, économiques, informatiques etc. qui ne fasse pas aujourd'hui un usage intensif de l'intégrale. C'est clairement un objet-clef de l'analyse mathématique et des sciences en général.

Il existe des théories plus ou moins fines de l'intégration. Des théories qui permettent de calculer l'intégrale de fonctions plus ou moins compliquées, ou de fonctions qui vivent sur des espaces plus ou moins bizarres (mais nécessaires à un certain niveau). Cette année nous étudierons l'intégrale dite de Riemann, qui est déjà très puissante et générale. La forme la plus générale de l'intégrale est celle de Lebesgue, étudiée en L3 de Mathématiques.

6.1 Fonctions en escalier

6.1.1 Subdivisions

Une *subdivision* d'un intervalle $[a, b]$ est une famille finie $\mathcal{S} = \{x_0, x_1, \dots, x_n\}$ de réels tels que

$$a = x_0 < x_1 < \dots < x_n = b.$$

On appelle *diamètre de la subdivision* la quantité

$$\delta(\mathcal{S}) = \min\{x_{i+1} - x_i; i = 0, \dots, n - 1\}.$$

Une subdivision \mathcal{S} de $[a, b]$ est dite *plus fine* qu'une subdivision \mathcal{S}' de $[a, b]$ si $\mathcal{S} \subset \mathcal{S}'$. Cela veut dire que \mathcal{S}' découpe $[a, b]$ en plus de morceaux. En particulier dans ce cas, on a évidemment $\delta(\mathcal{S}) \geq \delta(\mathcal{S}')$.

Si \mathcal{S} et \mathcal{S}' sont deux subdivisions quelconques de $[a, b]$ on note $\mathcal{S} \vee \mathcal{S}'$ la subdivision obtenue en prenant l'union des deux ensembles $\mathcal{S} \cup \mathcal{S}'$. La subdivision $\mathcal{S} \vee \mathcal{S}'$ est alors plus fine que \mathcal{S} et que \mathcal{S}' ; c'est d'ailleurs la plus *grossière* subdivision qui est plus fine que \mathcal{S} et \mathcal{S}' .

6.1.2 Fonctions en escalier

On dit qu'une fonction f sur $[a, b]$ est *en escalier*, s'il existe une subdivision $\mathcal{S} = \{x_0, x_1, \dots, x_n\}$ de $[a, b]$ telle que f soit constante sur chaque intervalle $]x_i, x_{i+1}[$, $i = 0, \dots, n - 1$.

Notez qu'on ne parle que des intervalles ouverts, rien n'est dit et imposé sur les points x_i , on peut très bien avoir de la continuité à droite ou à gauche en certain points, des points isolés, etc.

Si f est une fonction en escalier, on dit qu'une subdivision \mathcal{S} est *adaptée* à f si f est constante sur chaque intervalle de \mathcal{S} . Notez que \mathcal{S} peut très bien être *trop fine* pour f .

Proposition 6.1.1 *Soient f et g deux fonctions en escalier sur $[a, b]$ et $\lambda \in \mathbb{R}$. Alors $|f|$, $f + g$, λf et fg sont des fonctions en escalier sur $[a, b]$.*

Démonstration Pour $|f|$ et λf c'est vraiment évident.

Si \mathcal{S} et \mathcal{S}' sont des subdivisions adaptées à f et g respectivement alors $\mathcal{S} \vee \mathcal{S}'$ est adaptée à f et à g . On peut donc supposer que f et g sont en escalier sur la même subdivision $\mathcal{S}'' = (x_i)$. Ainsi f et g sont constantes, égales respectivement à c_k et d_k sur chaque intervalle $]x_k, x_{k+1}[$. Donc $f + g$ et fg sont égales à $c_k + d_k$ et $c_k d_k$ sur ces mêmes intervalles. Donc $f + g$ et fg sont aussi en escalier. \square

6.1.3 Intégrales de fonctions en escalier

Si f est une fonction en escalier sur $\mathcal{S} = (x_k)_{i=0}^n$, qui est égale à c_k sur chaque intervalle $]x_k, x_{k+1}[$, on note $I_{\mathcal{S}}(f)$ la surface algébrique de la famille de rectangles sous la courbe de f :

$$I_{\mathcal{S}}(f) = \sum_{i=0}^{n-1} c_k(x_{k+1} - x_k).$$

Proposition 6.1.2 *La quantité $I_{\mathcal{S}}(f)$ ne dépend pas du choix de la subdivision \mathcal{S} adaptée à f , elle ne dépend que de f et de $[a, b]$.*

Démonstration Prenons deux subdivisions $\mathcal{S} = (x_k)_{k=0}^n$ et $\mathcal{S}' = (y_l)_{l=0}^m$ adaptées à f et commençons par le cas où \mathcal{S}' est plus fine que \mathcal{S} . Sur chaque intervalle $]x_k, x_{k+1}[$ la fonction f est constante égale à c_k . Mais cet intervalle se découpe en union de certains intervalles $]y_l, y_{l+1}[$ $l = l_0, \dots, l_1$ où f prend des valeurs d_l qui sont forcément toutes égales à c_k . Donc

$$\sum_{l=l_0}^{l_1-1} d_l(y_{l+1} - y_l) = \sum_{l=l_0}^{l_1-1} c_k(y_{l+1} - y_l) = c_k(x_{k+1} - x_k).$$

On voit bien dans ce cas $I_{\mathcal{S}}(f) = I_{\mathcal{S}'}(f)$.

Maintenant si \mathcal{S} et \mathcal{S}' sont quelconques, adaptées à f alors $\mathcal{S} \vee \mathcal{S}'$ est plus fine que chacune et encore adaptée à f . Donc

$$I_{\mathcal{S}}(f) = I_{\mathcal{S} \vee \mathcal{S}'}(f) = I_{\mathcal{S}'}(f).$$

□

Cette quantité qui ne dépend donc que de f et de $[a, b]$ est notée

$$\int_a^b f(x) dx.$$

On l'appelle *intégrale de f entre a et b* .

Proposition 6.1.3 *Soient f et g deux fonctions en escalier sur $[a, b]$ et $\lambda \in \mathbb{R}$. Alors*

$$1) \quad \int_a^b \lambda f(x) dx = \lambda \int_a^b f(x) dx$$

et

$$2) \quad \int_a^b f(x) + g(x) dx = \int_a^b f(x) dx + \int_a^b g(x) dx.$$

Démonstration 1) Si \mathcal{S} est une subdivision adaptée à f alors elle l'est à λf aussi. Si f prenait les valeurs c_k sur les intervalles $]x_k, x_{k+1}[$ alors λf prend les valeurs λc_k sur ces mêmes intervalles. On obtient donc facilement le résultat annoncé.

2) On prend une subdivision adaptée à f et à g . Chacune de ces fonctions vaut c_k et d_k respectivement sur les intervalles $]x_k, x_{k+1}[$ de cette subdivision. Ainsi $f + g$ vaut $c_k + d_k$ sur ces intervalles et on conclut facilement. \square

Proposition 6.1.4 Soient f et g sont deux fonctions en escalier sur $[a, b]$.

1) Si f est positive sur tout $[a, b]$ alors

$$\int_a^b f(x) dx \geq 0.$$

2) Si $f \geq g$ sur tout $[a, b]$ alors

$$\int_a^b f(x) dx \geq \int_a^b g(x) dx.$$

3) On a

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

Démonstration

1) Toutes les valeurs c_k de f sur $]x_k, x_{k+1}[$ sont positives. Comme les $x_{k+1} - x_k$ sont tous positifs aussi on a

$$\int_a^b f(x) dx = \sum_{k=0}^{n-1} c_k (x_{k+1} - x_k) \geq 0.$$

2) On applique 1) à la fonction $f - g$ qui est positive. Donc

$$\int_a^b (f(x) - g(x)) dx \geq 0$$

mais on a vu précédemment que

$$\int_a^b f(x) - g(x) dx = \int_a^b f(x) dx - \int_a^b g(x) dx,$$

d'où le résultat.

3) Pour tout $x \in [a, b]$ on a $-|f(x)| \leq f(x) \leq |f(x)|$. Donc on a

$$-\int_a^b |f(x)| dx \leq \int_a^b f(x) dx \leq \int_a^b |f(x)| dx .$$

D'où

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx .$$

□

6.2 Fonctions Riemann-intégrables

Nous allons quitter les fonctions en escalier pour essayer de calculer l'intégrale de fonctions plus compliquées. En fait, nous allons utiliser l'idée intuitive de l'intégrale, la même que les grecs, c'est à dire que l'on peut calculer l'intégrale des fonctions qui se laissent bien approcher par des fonctions en escalier.

Au début on va travailler un peu à l'aveugle, c'est à dire qu'on va regarder les propriétés de l'intégrale sans savoir vraiment sur quelles fonctions ça s'applique. Il faudra être un peu patient, on donnera ensuite des classes de fonctions qui marchent et auxquelles toutes les propriétés déjà obtenues s'appliquent.

6.2.1 Construction de l'intégrale de Riemann

Soit f une fonction définie et bornée sur $[a, b]$. On note $\mathcal{E}_-(f)$ l'ensemble de toutes les fonctions en escalier ϕ sur $[a, b]$ qui vérifient $\phi \leq f$ sur $[a, b]$. On note $\mathcal{E}_+(f)$ l'ensemble de toutes les fonctions en escalier ψ sur $[a, b]$ qui vérifient $f \leq \psi$ sur $[a, b]$. On note

$$\mathcal{I}_-(f) = \left\{ \int_a^b \phi(x) dx ; \phi \in \mathcal{E}_-(f) \right\}$$

et

$$\mathcal{I}_+(f) = \left\{ \int_a^b \psi(x) dx ; \psi \in \mathcal{E}_+(f) \right\} .$$

Comme f est bornée, par exemple $m \leq f \leq M$ sur $[a, b]$, alors $\mathcal{E}_-(f)$ est non vide car il contient la fonction constante égale à m et $\mathcal{E}_+(f)$ est non vide car il contient la fonction constante égale à M . Donc $\mathcal{I}_-(f)$ et $\mathcal{I}_+(f)$ sont non

vides. Clairement $M(b-a)$ est un majorant de $\mathcal{I}_-(f)$ et $(b-a)m$ est un minorant de $\mathcal{I}_+(f)$. Donc

$$i_a^b(f) = \sup \mathcal{I}_-(f)$$

et

$$I_a^b(f) = \inf \mathcal{I}_+(f)$$

existent.

De toute évidence on a toujours

$$i_a^b(f) \leq I_a^b(f).$$

Mais rien ne dit pour une fonction quelconque que ces deux quantités vont être égales. On va d'ailleurs voir un contre-exemple.

Soit f la fonction sur l'intervalle $[0, 1]$ qui est définie par

$$f(x) = \begin{cases} 1 & \text{si } x \in \mathbb{Q}, \\ 0 & \text{si } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Si ψ est une fonction en escalier qui majore f sur $[0, 1]$, alors sur tout intervalle où ψ est constante il y a un rationnel, donc ψ doit être supérieure ou égale à 1 sur cet intervalle. Donc $\psi \geq 1$ sur $[0, 1]$, en particulier $I_0^1(f) \geq 1$.

Avec le même raisonnement, on voit bien que l'on a aussi $i_a^b(f) \leq 0$. Donc on n'a sûrement pas égalité entre $i_a^b(f)$ et $I_a^b(f)$.

L'idée de base pour la construction de l'intégrale c'est de définir les fonctions intégrables comme étant celles pour lesquelles les deux quantités coïncident. Dans ce cas cette valeur commune serait par définition l'intégrale de f sur $[a, b]$.

Faisons ensemble en détail le cas d'une fonction pour laquelle ça marche. Prenons la fonction $f(x) = x^2$ sur $[0, 1]$. Pour $n \in \mathbb{N}^*$ on considère la subdivision $x_k = k/n$, $k = 0, \dots, n$. Si, sur chaque intervalle $]x_k, x_{k+1}[$ on définit la fonction $\psi(x) = (k+1)^2/n^2$ et la fonction $\phi(x) = k^2/n^2$, $k = 0, \dots, n-1$, on a défini 2 fonctions en escalier telles que $\phi \leq f \leq \psi$. Calculons les intégrales des ces deux fonctions :

$$\int_0^1 \phi(x) dx = \sum_{k=0}^{n-1} \left(\frac{k}{n}\right)^2 \frac{1}{n} = \frac{1}{n^3} \frac{(n-1)n(2n-1)}{6},$$

$$\int_0^1 \psi(x) dx = \sum_{k=0}^{n-1} \left(\frac{k+1}{n}\right)^2 \frac{1}{n} = \frac{1}{n^3} \frac{n(n+1)(2n+1)}{6},$$

Donc on a en particulier

$$\frac{(n-1)n(2n-1)}{6n^3} \leq i_0^1(f) \leq I_0^1(f) \leq \frac{n(n+1)(2n+1)}{6n^3}$$

pour tout $n \in \mathbb{N}^*$. En faisant tendre n vers $+\infty$ on voit bien que

$$i_0^1(f) = I_0^1(f) = \frac{1}{3}.$$

On dit qu'une fonction bornée f sur $[a, b]$ est *intégrable* (au sens de Riemann) si $i_a^b(f) = I_a^b(f)$. Cette valeur commune est notée

$$\int_a^b f(x) dx$$

et appelée *intégrale de f entre a et b* .

Théorème 6.2.1 Une fonction f définie et bornée sur $[a, b]$ est intégrable sur $[a, b]$ si et seulement si il existe des suites (ϕ_n) et (ψ_n) de fonctions en escalier telles que $\phi_n \leq f \leq \psi_n$ pour tout n et

$$\lim_{n \rightarrow +\infty} \int_a^b \phi_n(x) - \psi_n(x) dx = 0.$$

Dans ce cas on a $\lim_{n \rightarrow +\infty} \int_a^b \phi_n(x) dx = \lim_{n \rightarrow +\infty} \int_a^b \psi_n(x) dx$ et cette limite commune est $\int_a^b f(x) dx$.

Démonstration Si f est intégrable alors $i_a^b(f) = I_a^b(f) = \int_a^b f(x) dx$. Par la caractérisation du sup, il existe une suite (ϕ_n) dans $\mathcal{E}_-(f)$ et une suite (ψ_n) dans $\mathcal{E}_+(f)$ telles que

$$\lim_{n \rightarrow +\infty} \int_a^b \phi_n(x) dx = i_a^b(f) \quad \lim_{n \rightarrow +\infty} \int_a^b \psi_n(x) dx = I_a^b(f).$$

D'où le résultat dans un sens.

Inversement, si il existe deux suites (ϕ_n) et (ψ_n) de fonctions en escalier telles que $\phi_n \leq f \leq \psi_n$ pour tout n et

$$\lim_{n \rightarrow +\infty} \int_a^b \phi_n(x) - \psi_n(x) dx = 0,$$

alors comme on a

$$\int_a^b \phi_n(x) \leq i_a^b(f) \leq I_a^b(f) \leq \int_a^b \psi_n(x) dx,$$

on en déduit que

$$0 \leq I_a^b(f) - i_a^b(f) \leq \int_a^b \psi_n(x) - \phi_n(x) dx .$$

Ce qui prouve que $I_a^b(f) = i_a^b(f)$. De plus on a aussi

$$0 \leq i_a^b(f) - \int_a^b \phi_n(x) dx \leq \int_a^b \psi_n(x) - \phi_n(x) dx$$

et

$$\int_a^b \psi_n(x) dx - I_a^b(f) \leq \int_a^b \psi_n(x) - \phi_n(x) dx .$$

On conclut facilement maintenant. □

6.2.2 Opérations sur les fonctions intégrables

Proposition 6.2.2 *Si f et g sont deux fonctions bornées et intégrables sur $[a, b]$ et si $\lambda \in \mathbb{R}$ alors λf et $f + g$ sont intégrables sur $[a, b]$ et*

$$\int_a^b \lambda f(x) dx = \lambda \int_a^b f(x) dx$$

$$\int_a^b f(x) + g(x) dx = \int_a^b f(x) dx + \int_a^b g(x) dx .$$

Démonstration Comme f et g sont intégrables on sait qu'il existe des suites (ϕ_n) , (ψ_n) , (θ_n) et (η_n) de fonctions en escalier telles que

$$\phi_n \leq f \leq \psi_n \quad \theta_n \leq g \leq \eta_n$$

pour tout n et

$$\lim_{n \rightarrow +\infty} \int_a^b \phi_n(x) dx = \lim_{n \rightarrow +\infty} \int_a^b \psi_n(x) dx = \int_a^b f(x) dx ,$$

$$\lim_{n \rightarrow +\infty} \int_a^b \theta_n(x) dx = \lim_{n \rightarrow +\infty} \int_a^b \eta_n(x) dx = \int_a^b g(x) dx .$$

Si $\lambda > 0$, alors

$$\lambda \phi_n \leq \lambda f \leq \lambda \psi_n$$

pour tout n et

$$\lim_{n \rightarrow +\infty} \int_a^b \lambda \phi_n(x) dx = \lim_{n \rightarrow +\infty} \int_a^b \lambda \psi_n(x) dx = \int_a^b \lambda f(x) dx .$$

Donc λf est intégrable sur $[a, b]$.

Mais comme on a déjà vu que $\int_A \lambda \phi_n(x) dx = \lambda \int_a^b \phi_n(x) dx$ et la même chose pour ψ_n on a $\int_a^b \lambda f(x) dx = \lambda \int_a^b f(x) dx$.

Pour $\lambda > 0$ on inverse les inégalités, mais les arguments sont les mêmes. Pour $\lambda = 0$ c'est trivial.

On a aussi

$$\phi_n + \theta_n \leq f + g \leq \psi_n + \eta_n$$

pour tout n et

$$\begin{aligned} \lim_{n \rightarrow +\infty} \int_a^b \phi_n(x) + \theta_n(x) dx &= \lim_{n \rightarrow +\infty} \int_a^b \phi_n(x) dx + \int_a^b \theta_n(x) dx \\ &= \lim_{n \rightarrow +\infty} \int_a^b \phi_n(x) dx + \lim_{n \rightarrow +\infty} \int_a^b \theta_n(x) dx \\ &= \int_a^b f(x) dx + \int_a^b g(x) dx \\ \lim_{n \rightarrow +\infty} \int_a^b \psi_n(x) + \eta_n(x) dx &= \lim_{n \rightarrow +\infty} \int_a^b \psi_n(x) dx + \int_a^b \eta_n(x) dx \\ &= \lim_{n \rightarrow +\infty} \int_a^b \psi_n(x) dx + \lim_{n \rightarrow +\infty} \int_a^b \eta_n(x) dx \\ &= \int_a^b f(x) dx + \int_a^b g(x) dx. \end{aligned}$$

On conclut facilement maintenant. □

6.2.3 Intégrales et inégalités

Les propriétés et en particulier les inégalités liées aux intégrales de fonctions en escalier vont s'étendre sans difficulté.

Proposition 6.2.3 *Soient f et g deux fonctions bornées et intégrables sur $[a, b]$.*

1) *Si $f \geq 0$ sur $[a, b]$ alors*

$$\int_a^b f(x) dx \geq 0.$$

2) *Si $f \geq g$ sur $[a, b]$ alors*

$$\int_a^b f(x) dx \geq \int_a^b g(x) dx.$$

Démonstration

1) Comme f est intégrable, on sait qu'il existe une suite (ψ_n) de fonction en escalier telles que $f_n \leq \psi_n$ pour tout n et $\int_a^b f(x) dx = \lim_{n \rightarrow +\infty} \int_a^b \psi_n(x) dx$. Comme f est positive, toutes les fonctions ψ_n le sont aussi, donc les intégrales des ψ_n sont positives et leur limite aussi.

2) On raisonne avec la fonction $f - g$ qui est positive sur $[a, b]$ et on applique 1). \square

Pour toute fonction f bornée sur $[a, b]$, pour tout $x \in [a, b]$ on pose $f_+(x) = \max\{f(x), 0\}$ et $f_-(x) = \max\{-f(x), 0\}$. Ces deux fonctions sont positives. On remarque que

$$f(x) = f_+(x) - f_-(x) \quad |f(x)| = f_+(x) + f_-(x).$$

Théorème 6.2.4 *Soit f une fonction bornée intégrable sur $[a, b]$, alors f_+ , f_- et $|f|$ sont aussi intégrables. On a*

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

Démonstration Si f est intégrable alors on a comme d'habitude des fonctions en escalier $\phi_n \leq f \leq \psi_n$ dont les intégrales convergent vers celle de f . On vérifie alors facilement que $(\phi_n)_+ \leq f_+ \leq (\psi_n)_+$ et que $(\psi_n)_+ - (\phi_n)_+ \leq \psi_n - \phi_n$. Donc f_+ est intégrable sur $[a, b]$.

Avec un raisonnement analogue on a l'intégrabilité de f_- .

Comme $|f|$ est la somme de deux fonctions intégrables elle est elle-même intégrable. Comme de plus on a toujours $-|f| \leq f \leq |f|$ on en déduit l'inégalité des intégrales. \square

Proposition 6.2.5 (Formule de la moyenne) *Soit f une fonction bornée et intégrable sur $[a, b]$ avec $a < b$. Soient m et M la borne inférieure et la borne supérieure de f sur $[a, b]$. Alors la quantité*

$$\mu = \frac{1}{b-a} \int_a^b f(x) dx$$

appartient à l'intervalle $[m, M]$.

Démonstration Comme on a $m \leq f \leq M$ on en déduit

$$m(b-a) \leq \int_a^b f(x) dx \leq M(b-a).$$

D'où le résultat. \square

Cette quantité μ obtenue ci-dessus est la *valeur moyenne* de f sur $[a, b]$.

6.2.4 Intégrales et produits

Nous allons maintenant montrer que le produit de deux fonctions intégrables est une fonction intégrable. C'est déjà un travail un peu plus subtil que ce qu'on a fait précédemment. Par contre il ne faut surtout pas s'attendre à ce que l'intégrale du produit soit égale au produit des intégrales. Prenez par exemple une fonction ϕ constante égale à γ sur $[0, 2]$ on a

$$\int_0^2 \phi(x)^2 dx = 2\gamma^2 \quad \left(\int_0^2 \phi(x) dx \right)^2 = 4\gamma^2.$$

Proposition 6.2.6 *Si f et g sont deux fonctions bornées et intégrables sur $[a, b]$ alors fg est bornée et intégrable sur $[a, b]$.*

Démonstration On commence la démonstration par le cas où f et g sont toutes les deux positives. On note M et N un majorant de f et de g respectivement. Soient (ϕ_n) , (ψ_n) , (θ_n) , (η_n) des suites de fonctions en escalier telles que $\phi_n \leq f \leq \psi_n$ et $\theta_n \leq g \leq \eta_n$ et avec les propriétés de convergence usuelles.

On pose

$$\phi'_n = (\phi_n)_+, \quad \psi'_n(x) = \min\{M, \psi_n(x)\}, \quad \theta'_n = (\theta_n)_+, \quad \eta'_n(x) = \min\{N, \eta_n(x)\}.$$

Ce sont toutes des fonctions en escalier sur $[a, b]$ et on a

$$0 \leq \phi'_n \leq f \leq \psi'_n \leq M, \quad 0 \leq \theta'_n \leq g \leq \eta'_n \leq N.$$

En particulier on a $\phi'_n \theta'_n \leq fg \leq \psi'_n \eta'_n$ et les fonctions qui encadrent fg sont en escalier. On va maintenant regarder la convergence de $\int_a^b \psi'_n(x) \eta'_n(x) - \phi'_n(x) \theta'_n(x) dx$. On a, en utilisant $\psi'_n \leq \psi_n$, $\phi'_n \geq \phi_n$, etc.

$$\begin{aligned} \int_a^b \psi'_n(x) \eta'_n(x) - \phi'_n(x) \theta'_n(x) dx &\leq \\ &\leq \int_a^b (\psi'_n(x) - \phi'_n(x)) \eta'_n(x) + \phi'_n(x) (\eta'_n(x) - \theta'_n(x)) dx \\ &\leq N \int_a^b \psi'_n(x) - \phi'_n(x) dx + M \int_a^b \eta'_n(x) - \theta'_n(x) dx \\ &\leq N \int_a^b \psi_n(x) - \phi_n(x) dx + M \int_a^b \eta_n(x) - \theta_n(x) dx. \end{aligned}$$

Cette dernière quantité tend bien vers 0. On a prouvé l'intégrabilité de fg .

Si f et g sont maintenant quelconques (bornées et intégrables sur $[a, b]$), si m et n sont des minorants de f et g respectivement alors $f - m$ et $g - n$ sont positives, bornées et intégrables sur $[a, b]$. Donc, on a vu que $(f - m)(g - n)$ est intégrable sur $[a, b]$. mais comme

$$fg = (f - m)(g - n) + mg + nf - mn$$

on a que fg est intégrable, comme somme de fonctions intégrables. \square

Une inégalité très importante pour les intégrales de produits.

Théorème 6.2.7 (Inégalité de Cauchy-Schwartz) *Si f et g sont deux fonctions bornées et intégrables sur $[a, b]$ alors*

$$\left(\int_a^b f(x)g(x) dx \right)^2 \leq \int_a^b f(x)^2 dx \int_a^b g(x)^2 dx .$$

Démonstration

Soit λ un réel quelconque. La fonction $f + \lambda g$ est intégrable, donc $(f + \lambda g)^2$ aussi. Cette dernière est une fonction positive, donc son intégrale est positive aussi. On a donc

$$\int_a^b f(x)^2 dx + 2\lambda \int_a^b f(x)g(x) dx + \lambda^2 \int_a^b g(x)^2 dx \geq 0$$

pour tout $\lambda \in \mathbb{R}$. C'est un polynôme de degré 2 en λ . S'il est toujours positif c'est que son discriminant est négatif ou nul :

$$4 \left(\int_a^b f(x)g(x) dx \right)^2 - 4 \int_a^b g(x)^2 dx \int_a^b f(x)^2 dx \leq 0 .$$

Ce qui donne l'inégalité de Cauchy-Schwartz. \square

6.3 Familles de fonctions intégrables

6.3.1 Manipulation de fonctions intégrables

Théorème 6.3.1 *Si f est une fonction bornée et intégrable et si g est une fonction définie sur $[a, b]$ est égale à f sauf sur un nombre fini de points, alors g est intégrable et $\int_a^b f(x) dx = \int_a^b g(x) dx$.*

Démonstration Par hypothèse il existe une subdivision $\mathcal{S} = (x_k)$ de $[a, b]$ telle que $f = g$ sur chacun des intervalles $]x_k, x_{k+1}[$. La fonction $f - g$ est donc nulle sur chacun des intervalles $]x_k, x_{k+1}[$. En particulier la fonction $f - g$ est en escalier ! Elle est donc intégrable et son intégrale est clairement nulle. La fonction $g = f - (f - g)$ est donc intégrable et son intégrale est égale à celle de f . \square

6.3.2 Monotonie

Maintenant qu'on a établi les propriétés principales des fonctions intégrables et des intégrales, on va revenir à la définition des fonctions intégrables et exhiber des ensembles explicites de fonctions intégrables.

On dit qu'une fonction f sur $[a, b]$ a une certaine propriété (comme être continue, monotone, dérivable, etc.) *par morceaux* si il existe une subdivision $\mathcal{S} = (x_k)_{k=0}^n$ de $[a, b]$ telle que f a cette propriété sur chacun des intervalles $]x_k, x_{k+1}[$ et si elle admet un prolongement ayant cette même propriété sur $[x_k, x_{k+1}]$.

Les fonctions étagées sont les fonctions constantes par morceaux.

Théorème 6.3.2 *Toute fonction monotone par morceaux sur $[a, b]$ est intégrable.*

Démonstration Une fonction monotone par morceaux est monotone sur chacun des intervalles $]x_k, x_{k+1}[$ et admet un prolongement monotone sur $[x_k, x_{k+1}]$. En particulier elles sont bornées sur chacun des intervalles $]x_k, x_{k+1}[$ et donc bornées sur $[a, b]$.

Fixons-nous maintenant sur un des intervalles de monotonie, appelons-le $[a, b]$ aussi et supposons f croissante. Cela ne change rien dans la suite.

On subdivise l'intervalle $[a, b]$ en n morceaux identiques de taille $(b-a)/n$, à savoir $x_k = a + k(b-a)/n$, $n = 0, \dots, n-1$. On construit les fonctions étagées

$$\phi(t) = f(x_k), \quad \psi(t) = f(x_{k+1})$$

sur $]x_k, x_{k+1}[$. On a donc $\phi \leq f \leq \psi$ et

$$\begin{aligned} \int_a^b (\psi(t) - \phi(t)) dt &= \sum_{k=0}^{n-1} \frac{(b-a)}{n} (f(x_{k+1}) - f(x_k)) \\ &= \frac{(b-a)}{n} (f(b) - f(a)). \end{aligned}$$

On a prouvé l'intégrabilité de f . □

6.3.3 Continuité

Rappelez-vous qu'une fonction f est continue sur un ensemble E si

$$\forall x \in E, \forall \varepsilon > 0, \exists \delta > 0; \forall y \in E, |x - y| \leq \delta \Rightarrow |f(x) - f(y)| \leq \varepsilon.$$

Dans cette définition, il faut noter que le $\delta > 0$ ci-dessus dépend de ε bien sûr mais aussi de x . Prenons un exemple, la fonction $x \mapsto x^2$ sur \mathbb{R} . Pour x

fixé, pour $\varepsilon > 0$ fixé, la condition $|f(x) - f(y)| \leq \varepsilon$ revient à $|x^2 - y^2| \leq \varepsilon$ ou encore, en écrivant $y = x + (y - x)$, on obtient $|2x(y - x) + (y - x)^2| \leq \varepsilon$. Prenons pour simplifier le cas $y - x > 0$ et $x > 0$. Si $y - x \leq \delta$ alors $|2x(y - x) + (y - x)^2| \leq (2x + \delta)\delta$. Si on veut que ce soit inférieur à ε il faut $\delta \leq \varepsilon/(2x + \delta) \leq \varepsilon/2x$. On voit bien que lorsque x est grand il faut une condition bien plus forte sur $|x - y|$ pour que cette inégalité soit réalisée. Le δ dépend de x .

Une fonction f est dite *uniformément continue* sur un ensemble E si

$$\forall \varepsilon > 0, \exists \delta > 0; \forall x, y \in E, |x - y| \leq \delta \Rightarrow |f(x) - f(y)| \leq \varepsilon.$$

Là, le δ est le même pour tout x dans E . On a vu ci-dessus que $x \mapsto x^2$ n'est pas uniformément continue sur \mathbb{R} .

Les fonctions linéaires, les fonctions lipschitziennes

$$\exists K > 0; |f(x) - f(y)| \leq K|x - y|, \forall x, y$$

sont uniformément continues sur \mathbb{R} .

Théorème 6.3.3 (Théorème de Heine) *Toute fonction continue sur un intervalle fermé borné $[a, b]$ est uniformément continue sur $[a, b]$.*

Démonstration Raisonnons par l'absurde et supposons que f n'est pas uniformément continue sur $[a, b]$. Ainsi, il existe $\varepsilon > 0$ tel que pour tout $\delta > 0$, il existe $x, y \in [a, b]$ avec $|x - y| \leq \delta$ et tels que $|f(x) - f(y)| > \varepsilon$. Ainsi, il existe deux suites (x_n) et (y_n) dans $[a, b]$ telles que $|x_n - y_n| \leq 1/n$ et $|f(x_n) - f(y_n)| > \varepsilon$. Les suites (x_n) et (y_n) sont bornées, elles admettent donc une sous-suite convergente, vers des limites $x, y \in [a, b]$. Comme on a $|x_n - y_n| \leq 1/n$, on a forcément $x = y$. Comme f est continue alors $\lim_n |f(x_n) - f(y_n)| = 0$ et ça contredit le fait que $|f(x_n) - f(y_n)| \geq \varepsilon$ pour tout n . \square

Théorème 6.3.4 *Toute fonction continue par morceaux sur $[a, b]$ est intégrable.*

Démonstration Sur chacun des intervalles $]x_k, x_{k+1}[$ la fonction est bornée car prolongeable par continuité sur $[x_k, x_{k+1}]$. On peut supposer, sans perte de généralité que f est continue sur $[a, b]$. On a vu qu'alors f est uniformément continue sur $[a, b]$. Donc pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que pour tout $x, y \in [a, b]$ vérifiant $|x - y| \leq \delta$, on ait $|f(x) - f(y)| \leq \varepsilon$. Découpons $[a, b]$ en intervalles réguliers de diamètre inférieur à δ . On définit deux fonctions étagées ϕ et ψ sur cette subdivision par

$$\phi(t) = f(x_k) - \varepsilon, \quad \psi(t) = f(x_k) + \varepsilon$$

sur $]x_k, x_{k+1}[$. On a $\phi \leq f \leq \psi$ grace à la continuité uniforme de f . De plus

$$\begin{aligned} \int_a^b \psi(t) - \phi(t) dt &= \sum_{k=0}^{n-1} (x_{k+1} - x_k)(f(x_k) + \varepsilon - f(x_k) + \varepsilon) \\ &= 2\varepsilon(b - a). \end{aligned}$$

On peut donc choisir ϕ et ψ de telle sorte que cette différence d'intégrale soit aussi petite que voulue. Cela montre l'intégrabilité de f . \square

6.3.4 Convention et relation de Chasles

Proposition 6.3.5 *Soit f une fonction bornée sur $[a, b]$ et $c \in [a, b]$.*

- 1) *Si f est intégrable sur $[a, b]$ alors f est intégrable sur $[a, c]$ et sur $[c, b]$.*
- 2) *Si f est intégrable sur $[a, c]$ et sur $[c, b]$ alors f est intégrable sur $[a, b]$.*
- 3) *Si f est intégrable sur $[a, b]$ alors on a la relation de Chasles*

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Démonstration

1) Soient ϕ et ψ sont deux fonctions en escaliers sur $[a, b]$ telles que $\phi \leq f \leq \psi$ et $\int_a^b \psi(x) - \phi(x) dx \leq \varepsilon$. On note ϕ_1 et ψ_1 les restrictions de ϕ et ψ sur $[a, c]$. Elles sont encore en escalier et vérifient $\phi_1 \leq f \leq \psi_1$ sur $[a, c]$. Comme on voit facilement que $\int_a^c \psi_1(x) - \phi_1(x) dx \leq \int_a^b \psi(x) - \phi(x) dx \leq \varepsilon$, on a l'intégrabilité de f sur $[a, c]$.

2) Facile (laissé en exercice).

3) La relation de Chasles est claire pour les fonctions en escalier. Maintenant si ϕ_1, ψ_1 sont comme d'habitude pour f sur $[a, c]$ et ϕ_2, ψ_2 pour f sur $[c, b]$, alors $\phi_1 + \phi_2$ et $\psi_1 + \psi_2$ sont associées à f sur $[a, b]$. \square

On adopte une convention très pratique pour la suite : si $a \leq b$ alors

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

En particulier, la relation

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$$

devient vraie quel que soient les relations d'ordre entre a, b et c (si au moins deux des trois intégrales sont bien définies, l'autre l'est forcément).

6.4 Primitives et intégrales

Jusque là l'intégrale peut paraître comme un objet assez abstrait et avec lequel il est pas plus facile de faire des calculs que de faire des approximations de surface par des rectangles. Le miracle va se faire quand on va établir le lien avec la primitive.

6.4.1 Le théorème fondamental

Pour le moment on va se concentrer sur l'étude des propriétés de la fonction

$$F(x) = \int_c^x f(t) dt.$$

On va voir que F améliore toujours la régularité de f . Mais il faut faire un peu attention aux détails !

Proposition 6.4.1 *Soit f une fonction bornée et intégrable sur $[a, b]$, avec $|f| \leq M$. Soit $c \in [a, b]$. Alors, si on pose $F(x) = \int_c^x f(x) dx$ on a*

$$|F(x) - F(y)| \leq M |x - y|$$

pour tout $x, y \in [a, b]$.

Démonstration Supposons que $x \leq y$ (l'autre cas se traite de la même façon). On a

$$F(y) - F(x) = \int_x^y f(t) dt.$$

En particulier

$$|F(y) - F(x)| \leq \int_x^y |f(t)| dt \leq \int_x^y M dt = M(y - x).$$

□

Corollaire 6.4.2 *Soit f une fonction bornée et intégrable sur $[a, b]$. Soit $c \in [a, b]$. On pose $F(x) = \int_c^x f(x) dx$, pour tout $x \in [a, b]$. Alors F est continue sur $[a, b]$.*

Démonstration On a vu ci-dessus que F est lipschitzienne. La conclusion est facile maintenant (exercice). □

Proposition 6.4.3 Soit f une fonction bornée et intégrable sur $[a, b]$. Soit $c \in [a, b]$. On pose $F(x) = \int_c^x f(x) dx$, pour tout $x \in [a, b]$. Si f admet une limite à droite en $x_0 \in [a, b]$ alors F est dérivable à droite en x_0 et $F'_d(x_0) = \lim_{x \rightarrow x_0^+} f(x)$. De même, si f admet une limite à gauche en $x_0 \in [a, b]$ alors F est dérivable à gauche en x_0 et $F'_g(x_0) = \lim_{x \rightarrow x_0^-} f(x)$.

Démonstration Supposons que f admet une limite à droite en $x_0 \in [a, b]$, notée l . Alors pour $h > 0$ on a

$$F(x_0 + h) - F(x_0) = \int_{x_0}^{x_0+h} f(t) dt.$$

Si h est suffisamment petit alors $|f(t) - l| \leq \varepsilon$ pour tout $t \in [x_0, x_0 + h]$. On a donc

$$\begin{aligned} \left| \frac{F(x_0 + h) - F(x_0)}{h} - l \right| &= \frac{1}{h} \left| \int_{x_0}^{x_0+h} f(t) dt - hl \right| \\ &= \frac{1}{h} \left| \int_{x_0}^{x_0+h} f(t) - l dt \right| \\ &\leq \frac{1}{h} \int_{x_0}^{x_0+h} |f(t) - l| dt \\ &\leq \frac{1}{h} \int_{x_0}^{x_0+h} \varepsilon dt \\ &= \frac{\varepsilon h}{h} = \varepsilon. \end{aligned}$$

Ca montre que F est dérivable à droite, de dérivée à droite l . L'autre cas se fait de la même manière. \square

Le théorème qui suit découle maintenant immédiatement.

Théorème 6.4.4 Si f est une fonction continue sur $[a, b]$, si $c \in [a, b]$, alors la fonction $F(x) = \int_c^x f(t) dt$ est dérivable sur $[a, b]$, de dérivée f .

Faites attention, si f n'est pas continue, la fonction F n'est plus dérivable. Prenons un exemple. Soit f la fonction sur $[0, 2]$ qui vaut 1 sur $[0, 1]$ et 2 sur $]1, 2]$. On pose $F(x) = \int_0^x f(t) dt$. Pour tout $x \in [0, 1]$ on a $F(x) = x$. En particulier $F(1) = 1$. Ensuite, pour $x \in]1, 2]$ on a $F(x) = 1 + 2x$.

La fonction F est continue en 1, comme on le savait déjà, mais elle n'est pas dérivable (dérivée 1 à gauche, 2 à droite).

On dit qu'une fonction f définie sur $[a, b]$ admet une primitive F sur $[a, b]$ si F est dérivable sur $[a, b]$ et si $F' = f$ sur $[a, b]$.

Proposition 6.4.5 *Si F et G sont deux primitives de f sur $[a, b]$ alors $F - G$ est constante sur $[a, b]$.*

Démonstration On a déjà vu cet argument : $F - G$ est dérivable, de dérivée nulle sur $[a, b]$, donc constante sur $[a, b]$. \square

Proposition 6.4.6 *Toute fonction continue sur $[a, b]$ admet une primitive sur $[a, b]$.*

Démonstration On l'a vu. Si f est continue alors $F(x) = \int_c^x f(t) dt$ est dérivable, de dérivée f . \square

Attention, il existe des fonctions non continues qui ont des primitives. Par exemple $F(x) = x^2 \sin(1/x)$ pour $x \neq 0$ et $F(0) = 0$ est dérivable. Sa dérivée est $f(x) = 2x \sin(1/x) - \cos(1/x)$ pour $x \neq 0$ et $f(0) = 1$. Sa dérivée n'est pas continue en 0.

Théorème 6.4.7 (Théorème fondamental de l'analyse) *Si f est une fonction continue sur $[a, b]$ et si F est une primitive de f alors*

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

Démonstration Posons $G(x) = \int_a^x f(x) dx$, pour tout $x \in [a, b]$. On a vu que G est dérivable et que $G' = f$, donc G est une primitive de f . On a de plus $\int_a^b f(x) dx = G(b) - G(a)$. Maintenant si F est une primitive quelconque de f alors, comme F et G diffèrent seulement par une constante additive, on a $F(b) - F(a) = G(b) - G(a) = \int_a^b f(x) dx$. \square

Une fonction f est dite C^1 sur $[a, b]$ si elle est dérivable sur $[a, b]$ et si sa dérivée est continue sur $[a, b]$.

Théorème 6.4.8 *Si f est une fonction C^1 sur $[a, b]$ alors*

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

Démonstration Si f est C^1 alors sa dérivée f' est continue et f est une primitive de f' . On applique alors le théorème ci-dessus. \square

6.4.2 Intégration par parties

Théorème 6.4.9 Soient u et v deux fonctions C^1 sur $[a, b]$, alors

$$\int_a^b u'(x)v(x) dx = [u(x)v(x)]_a^b - \int_a^b u(x)v'(x) dx.$$

Démonstration La preuve est simple. On a $(uv)'(x) = u'(x)v(x) + u(x)v'(x)$. Comme uv est C^1 on applique le théorème fondamental de l'analyse :

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

Les fonctions $u'v$ et uv' sont intégrables sur $[a, b]$ donc

$$\int_a^b (uv)'(x) dx = \int_a^b u'(x)v(x) dx + \int_a^b u(x)v'(x) dx.$$

D'où le résultat. □

6.4.3 Changement de variables

Théorème 6.4.10 (Changement de variable) Soit f une fonction continue sur $[a, b]$ et soit φ une fonction C^1 sur $[\alpha, \beta]$ telle que $\varphi([\alpha, \beta]) \subset [a, b]$. Alors

$$\int_{\alpha}^{\beta} f(\varphi(t)) \varphi'(t) dt = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(x) dx.$$

Démonstration La fonction f est continue donc elle admet une primitive F qui est C^1 . La fonction $F \circ \varphi$ est donc C^1 et $(F \circ \varphi)'(x) = \varphi'(x) f \circ \varphi(x)$. Par le théorème fondamental on a

$$\int_{\alpha}^{\beta} \varphi'(x) f \circ \varphi(x) dx = (F \circ \varphi)(\beta) - (F \circ \varphi)(\alpha).$$

Mais cette dernière quantité est aussi égale à

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(x) dx.$$

□

Théorème 6.4.11 (Changement de variable, variante 1)

Soit f une fonction continue sur $[a, b]$ et soit ϕ une fonction C^1 sur $[\alpha, \beta]$ telle que $\phi([\alpha, \beta]) \subset [a, b]$ et telle que ϕ soit strictement monotone sur $]\alpha, \beta[$. Alors

$$\int_{\alpha}^{\beta} f(t) dt = \int_{\varphi^{-1}(\alpha)}^{\varphi^{-1}(\beta)} f(\phi(x)) \varphi'(x) dx.$$

Démonstration Comme φ est strictement monotone sur $[\alpha, \beta]$, c'est une bijection. Si on pose $\alpha' = \varphi^{-1}(a)$ et $\beta' = \varphi^{-1}(b)$, ce n'est plus qu'une application du théorème ci-dessus. \square

Théorème 6.4.12 (Changement de variable, variante 2)

Soit f une fonction continue sur $[a, b]$ et soit ϕ une fonction C^1 sur $[\alpha, \beta]$ telle que $\phi([\alpha, \beta]) \subset [a, b]$ et telle que ϕ' ne s'annule pas sur $] \alpha, \beta [$. Alors

$$\int_{\alpha}^{\beta} f(\varphi(t)) dt = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(x) \frac{1}{\varphi'(\varphi^{-1}(x))} dx.$$

Démonstration Comme la dérivée de φ ne s'annule pas sur $] \alpha, \beta [$ alors φ est strictement monotone sur $[\alpha, \beta]$, en particulier c'est une bijection. On écrit

$$\begin{aligned} \int_{\alpha}^{\beta} f(\varphi(t)) dt &= \int_{\alpha}^{\beta} \frac{f(\varphi(t))}{\varphi'(t)} \varphi'(t) dt \\ &= \int_{\alpha}^{\beta} \frac{f(\varphi(t))}{\varphi'(\varphi^{-1}(\varphi(t)))} \varphi'(t) dt. \end{aligned}$$

La fonction $g(t) = f(\varphi(t))/\varphi'(\varphi^{-1}(\varphi(t)))$ est continue sur $[a, b]$, on applique le théorème du changement de variable, on obtient

$$\int_{\alpha}^{\beta} f(\varphi(t)) dt = \int_{\varphi(\alpha)}^{\varphi(\beta)} \frac{f(x)}{\varphi'(\varphi^{-1}(x))} dx.$$

\square

Grâce à deux exemple on va voir comment concrètement on utilise ces formules. On commence par

$$\int_0^1 \sqrt{1-x^2} dx.$$

On veut faire le changement de variable $x = \cos(u)$. L'application $u \mapsto \cos(u)$

est C^1 et c'est une bijection de $[0, \pi/2]$ dans $[0, 1]$. On applique la variante 1,

$$\begin{aligned} \int_0^1 \sqrt{1-x^2} dx &= \int_{\arccos(0)}^{\arccos(1)} \sqrt{1-\cos(u)^2} (-\sin(u)) du \\ &= \int_{\pi/2}^0 -\sin(u)^2 du \\ &= \int_0^{\pi/2} \frac{1}{2} (1 - \cos(2u)) du \\ &= \left[\frac{1}{2} \left(u - \frac{1}{2} \sin(2u) \right) \right]_0^{\pi/2} \\ &= \frac{\pi}{4}. \end{aligned}$$

Le moyen pour se souvenir de cette variante ainsi que de l'autre est la suivante : on a choisi le changement de variable $x = \cos(u)$, on doit remplacer tout ce qui est en x par des u . Ca veut dire 3 choses :

- les bornes : si x vaut 0 alors u vaut $\arccos(0)$, si x vaut 1 alors u vaut $\arccos(1)$;
- la variable x quand elle apparaît dans $f : \sqrt{1-\cos(u)^2}$;
- le dx doit être remplacé par du : on a $x = \cos(u)$, on écrit

$$\frac{dx}{du} = -\sin(u)$$

ou encore

$$dx = -\sin(u) du$$

et on remplace.

Faisons un exemple avec la variante 2. On veut calculer, pour $x > 0$

$$\int_0^x \sin(\sqrt{t}) dt.$$

La fonction $t \mapsto \sqrt{t}$ est C^1 sur $[0, x]$ et sa dérivée ne s'annule pas sur l'intervalle ouvert. On applique la variante 2 :

$$\begin{aligned} \int_0^x \sin(\sqrt{t}) dt &= \int_0^{\sqrt{x}} \sin(y) 2y dy \\ &= [-2y \cos(y)]_0^{\sqrt{x}} + \int_0^{\sqrt{x}} 2 \cos(y) dy \\ &= -2\sqrt{x} \cos(\sqrt{x}) + 2 \sin(\sqrt{x}). \end{aligned}$$

Comme précédemment détaillons le changement de variable avec la méthode mnémotechnique. On veut poser $y = \sqrt{t}$:

- les bornes : quand t vaut 0 alors y vaut 0, quand t vaut x alors y vaut \sqrt{x} ;
- la variable : on remplace \sqrt{t} par y ;
- la différentielle : on a $y = \sqrt{t}$ donc

$$dy = \frac{1}{2\sqrt{t}} dt = \frac{1}{2y} dt,$$

ce qui donne

$$dt = 2y dy.$$

6.5 Quelques résultats

Dans cette section nous allons voir plusieurs petits résultats assez utiles liés à la théorie de l'intégration.

Théorème 6.5.1 *Soit f une fonction continue sur $[a, b]$ et de signe constant sur $[a, b]$. Si $\int_a^b f(x) dx = 0$, alors f est identiquement nulle sur $[a, b]$.*

Démonstration Supposons par exemple que f est positive sur $[a, b]$. Pour tout $x \in [a, b]$, on pose $F(x) = \int_a^x f(t) dt$. Cette quantité est positive pour tout x car f est positive, elle est aussi $\leq \int_a^b f(t) dt$, puisqu'il faut rajouter $\int_x^b f(t) dt$ qui est positif. Donc $F(x) = 0$ pour tout $x \in [a, b]$. Mais, comme f est continue, par le théorème fondamental de l'analyse on sait que F est dérivable, de dérivée $F'(x) = f(x)$. Donc $f(x) = F'(x) = 0$ pour tout x . \square

Ce théorème permet d'améliorer l'inégalité de Cauchy-Schwarz dans le cas des fonctions continues.

Théorème 6.5.2 *Soient f et g deux fonctions continues sur $[a, b]$. Alors*

$$\left(\int_a^b f(x)g(x) dx \right)^2 \leq \left(\int_a^b f(x)^2 dx \right) \left(\int_a^b g(x)^2 dx \right).$$

De plus, on a égalité ci-dessus si et seulement si il existe $\lambda \in \mathbb{R}$ tel que $f(x) = \lambda g(x)$ pour tout $x \in [a, b]$.

Démonstration On a déjà vu la démonstration de l'inégalité. Souvenez-vous, on avait

$$\begin{aligned} \int_a^b (f(x) + \lambda g(x))^2 dx &= \\ &= \int_a^b f(x)^2 dx + 2\lambda \int_a^b f(x)g(x) dx + \lambda^2 \int_a^b g(x)^2 dx \geq 0 \end{aligned}$$

pour tout $\lambda \in \mathbb{R}$. C'est un polynôme de degré 2 en λ . Son discriminant est négatif ou nul :

$$4 \left(\int_a^b f(x)g(x) dx \right)^2 - 4 \int_a^b g(x)^2 dx \int_a^b f(x)^2 dx \leq 0.$$

Si on a l'égalité dans l'inégalité de Cauchy-Schwarz, alors le discriminant est nul. Cela veut dire qu'il existe $\lambda_0 \in \mathbb{R}$ tel que $\int_a^b (f(x) + \lambda_0 g(x))^2 dx = 0$. Mais la fonction $(f + \lambda_0 g)^2$ est continue et positive, donc par le théorème précédent elle est nulle sur $[a, b]$. \square

Théorème 6.5.3 (Formule de la moyenne) *Soit f continue sur $[a, b]$ et g intégrable sur $[a, b]$ avec g de signe constant. Alors il existe $c \in [a, b]$ tel que*

$$\int_a^b f(x)g(x) dx = f(c) \int_a^b g(x) dx.$$

En particulier si $g = 1$ on a

$$f(c) = \frac{1}{b-a} \int_a^b f(x) dx.$$

Démonstration Nous supposons ici que g est positive. L'autre cas se traite de manière analogue.

La fonction f est continue sur $[a, b]$, donc elle est bornée et atteint ses bornes :

$$m = \inf\{f(x); x \in [a, b]\}, \quad M = \sup\{f(x); x \in [a, b]\}.$$

On a

$$m \int_a^b g(x) dx \leq \int_a^b f(x)g(x) dx \leq M \int_a^b g(x) dx.$$

Si $\int_a^b g(x) dx = 0$, alors la double inégalité ci-dessus montre que $\int_a^b f(x)g(x) dx = 0$ et donc que l'énoncé du théorème est trivialement vérifié.

Si $\int_a^b g(x) dx \neq 0$ alors on a

$$m \leq \frac{\int_a^b f(x)g(x) dx}{\int_a^b g(x) dx} \leq M.$$

Par le théorème des valeurs intermédiaires on sait que comme f est continue elle passe par toutes les valeurs de $[m, M]$. Donc il existe $c \in [a, b]$ tel que

$$f(c) = \frac{\int_a^b f(x)g(x) dx}{\int_a^b g(x) dx}.$$

\square

6.6 Sommes de Riemann, de Darboux, surfaces etc.

Dans cette section on va concrétiser le fait qu'une fonction intégrable est une fonction dont on peut approcher l'intégrale par des intégrales de fonctions en escalier, donc par des sommes particulières.

6.6.1 Sommes de Darboux

Soit f une fonction bornée sur $[a, b]$ et $\mathcal{S} = \{a = x_0 < x_1 < \dots < x_n = b\}$ une subdivision de $[a, b]$. On note

$$m_k = \inf\{f(x); x \in [x_k, x_{k+1}]\} \quad \text{et} \quad M_k = \sup\{f(x); x \in [x_k, x_{k+1}]\}$$

pour tout $k = 0, \dots, n-1$. On définit les *sommes de Darboux*

$$s(\mathcal{S}, f) = \sum_{k=0}^{n-1} m_k(x_{k+1} - x_k) \quad \text{et} \quad S(\mathcal{S}, f) = \sum_{k=0}^{n-1} M_k(x_{k+1} - x_k).$$

On note

$$s_a^b(f) = \sup\{s(\mathcal{S}, f); \mathcal{S} \text{ subdivision de } [a, b]\}$$

$$S_a^b(f) = \inf\{S(\mathcal{S}, f); \mathcal{S} \text{ subdivision de } [a, b]\}.$$

Théorème 6.6.1 *Soit f bornée sur $[a, b]$ alors f est intégrable si et seulement si $s_a^b(f) = S_a^b(f)$. Cette valeur commune est alors l'intégrale de f sur $[a, b]$.*

Démonstration On a toujours

$$s_a^b(f) \leq i_a^b(f) \leq I_a^b(f) \leq S_a^b(f),$$

de manière évidente. Donc si $s_a^b(f) = S_a^b(f)$ on a bien que f est intégrable et que son intégrale est cette valeur commune.

Réciproquement, si f est intégrable, alors $i_a^b(f) = \int_a^b f(x) dx$. On sait aussi qu'il existe φ fonction en escalier, telle que $\varphi \leq f$ et

$$i_a^b(f) - \varepsilon \leq \int_a^b \varphi(x) dx \leq i_a^b(f).$$

Pour une fonction en escalier φ , montrons que

$$s_a^b(\varphi) = \int_a^b \varphi(x) dx.$$

Soit M tel que $|\varphi| \leq M$. Sur chaque intervalle $[x_k, x_{k+1}]$ d'une subdivision adaptée à φ on découpe l'intervalle en $[x_k, x_k + \delta] \cup [x_k + \delta, x_{k+1} - \delta] \cup [x_{k+1} - \delta, x_{k+1}]$. Sur cet intervalle et pour cette nouvelle subdivision, l'erreur commise entre $s(\mathcal{S}, \varphi)$ et $\int_{x_k}^{x_{k+1}} \varphi(x) dx$ est inférieure à $4M\delta$. Donc on peut rendre cette différence aussi petite que voulue. D'où le résultat annoncé.

Enfin, notons que si $f \geq g$ sur $[a, b]$ alors $s_a^b(f) \geq s_a^b(g)$ trivialement.

On en déduit que $i_a^b(f) - \varepsilon \leq s_a^b(\varphi) \leq s_a^b(f)$. Ca prouve que $i_a^b(f) \leq s_a^b(f)$ et que donc ils sont égaux. On fait de même avec $I_a^b(f)$ et $S_a^b(f)$ et on a le résultat annoncé. \square

6.6.2 Sommes de Riemann

L'idée des sommes de Riemann est un peu différente de celle des sommes de Darboux et constitue une approximation assez utile des intégrales.

Soit f une fonction bornée sur $[a, b]$, soit $\mathcal{S} = \{a = x_0 < x_1 < \dots < x_n = b\}$ une subdivision de f , soit $\mathcal{E} = \{\xi_0, \dots, \xi_{n-1}\}$ une famille de réels tels que $\xi_k \in [x_k, x_{k+1}]$ pour tout k . On pose

$$R(\mathcal{S}, \mathcal{E}, f) = \sum_{k=0}^{n-1} f(\xi_k)(x_{k+1} - x_k).$$

Théorème 6.6.2 *Soit f une fonction bornée et intégrable sur $[a, b]$, alors*

$$\int_a^b f(x) dx = \lim_{\delta(\mathcal{S}) \rightarrow 0} R(\mathcal{S}, \mathcal{E}, f).$$

Démonstration Par définition du fait que f est intégrable, on sait que pour tout $\varepsilon > 0$ il existe deux fonctions étagées φ et ψ telles que $\varphi \leq f \leq \psi$ et

$$\int_a^b \psi(x) - \varphi(x) dx \leq \frac{\varepsilon}{2}.$$

Donnons-nous une subdivision $\mathcal{S} = \{x_0, \dots, x_p\}$ de $[a, b]$, de diamètre δ , adaptée à φ . Soit $\mathcal{S}' = \{y_0, \dots, y_n\}$, une subdivision quelconque, de diamètre $\delta' < \delta$. Dans chaque intervalle $]y_k, y_{k+1}[$, on choisit un point ξ_k . On pose

$$S = \sum_{k=0}^{n-1} f(\xi_k)(y_{k+1} - y_k).$$

Cette somme se divise en deux : ceux des k pour lesquels $]y_k, y_{k+1}[$ est inclus dans un des intervalles $]x_j, x_{j+1}[$, et les autres. Ceux de la deuxième espèce sont au plus au nombre de $p+1$ (car ils chevauchent des intervalles $]x_j, x_{j+1}[$).

On note σ la somme des termes $f(\xi_k)(y_{k+1} - y_k)$ de la première espèce et σ' ceux de la deuxième. On a donc $S = \sigma + \sigma'$. On a aussi

$$|\sigma'| \leq \delta' M(p+1),$$

où M est un majorant de φ . Si on choisit $\delta' \leq \varepsilon/(2M(p+1))$ alors $|S - \sigma| \leq \varepsilon/2$.

D'autre part, si k est un point de la première espèce alors φ est constante sur $]y_k, y_{k+1}[$, de valeurs $\varphi(\xi_k)$. Si on rajoute à \mathcal{S}' les $p+1$ points de \mathcal{S} on obtient une subdivision \mathcal{S}'' qui est adaptée à φ . La somme σ est alors égale à $\int_a^b \varphi(x) dx$ sauf pour les points de $\mathcal{S}'' \setminus \mathcal{S}'$. Ces points sont au plus $p+1$ et l'erreur commise est au plus $\delta' M$. Donc au total

$$\left| \int_a^b \varphi(x) dx - \sigma \right| \leq \frac{\varepsilon}{2}.$$

On a montré que

$$\left| \int_a^b \varphi(x) dx - S \right| \leq \varepsilon$$

dès que la subdivision \mathcal{S}' est de diamètre plus petit qu'un seuil fixé.

Appelons $S(f)$ la somme de Riemann apparaissant dans l'énoncé. Si on note $S(\varphi)$ et $S(\psi)$ les mêmes sommes mais avec φ et ψ qui remplacent f , on a alors

$$S(\varphi) \leq S(f) \leq S(\psi).$$

On a aussi montré plus haut que, pour un pas de la subdivision suffisamment petit

$$\int_a^b \varphi(x) dx - \frac{\varepsilon}{2} \leq S(\varphi)$$

et

$$S(\psi) \leq \int_a^b \psi(x) dx + \frac{\varepsilon}{2}.$$

Ce qui donne en tout

$$\int_a^b \varphi(x) - \psi(x) dx - \frac{\varepsilon}{2} \leq S(f) - \int_a^b f(x) dx \leq \int_a^b \psi(x) - \varphi(x) dx + \frac{\varepsilon}{2}.$$

Soit finalement

$$\left| \int_a^b f(x) dx - S(f) \right| \leq \varepsilon.$$

□

Corollaire 6.6.3 *Si f est une fonction bornée et intégrable sur $[a, b]$ alors*

$$\int_a^b f(t) dt = \lim_{n \rightarrow +\infty} \frac{(b-a)}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{(b-a)}{n}\right).$$

Cette dernière identité permet parfois de calculer efficacement certaines limites de sommes. Par exemple, posons

$$u_n = \frac{1}{n+1} + \dots + \frac{1}{2n}.$$

Alors

$$u_n = \frac{1}{n} \sum_{k=1}^n \frac{1}{1+k/n}.$$

Posons $f(t) = 1/(1+t)$, alors

$$u_n = \frac{1}{n} \sum_{k=1}^n f(k/n).$$

On reconnaît la somme de Riemann de f sur $[0, 1]$ et donc

$$\lim_{n \rightarrow +\infty} u_n = \int_0^1 f(t) dt = \ln(2).$$

6.6.3 Estimation d'erreurs

Proposition 6.6.4 *Si f est C^1 sur $[0, 1]$ et si on pose*

$$R_n = \frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right)$$

on a

$$\left| \int_0^1 f(x) dx - R_n \right| \leq \frac{M_1}{2n}$$

où $M_1 = \sup\{|f'(x)| ; x \in [0, 1]\}$.

Démonstration Sur l'intervalle $[x_k, x_{k+1}]$ l'erreur est, en utilisant l'inégalité

des accroissements finis

$$\begin{aligned}
 \left| \int_{x_k}^{x_{k+1}} f(x) - f(x_k) dx \right| &\leq \int_{x_k}^{x_{k+1}} |f(x) - f(x_k)| dx \\
 &\leq M_1 \int_{x_k}^{x_{k+1}} |x - x_k| dx \\
 &= M_1 \left[\frac{(x - x_k)^2}{2} \right]_{x_k}^{x_{k+1}} \\
 &= \frac{M_1}{2n^2}.
 \end{aligned}$$

Cette erreur étant obtenue sur n intervalles, on a le résultat annoncé. \square

On peut améliorer cette approximation en approchant l'intégrale par des trapèzes plutôt que des rectangles. Sur chaque intervalle $[x_k, x_{k+1}]$ on pose

$$f_n(x) = f(x_k) + \frac{(f(x_{k+1}) - f(x_k))}{x_{k+1} - x_k} (x - x_k).$$

Ainsi

$$\begin{aligned}
 \int_{x_k}^{x_{k+1}} f_n(x) dx &= f(x_k)(x_{k+1} - x_k) + \frac{(f(x_{k+1}) - f(x_k))}{x_{k+1} - x_k} \frac{(x_{k+1} - x_k)^2}{2} \\
 &= f(x_k) \frac{1}{n} + (f(x_{k+1}) - f(x_k)) \frac{1}{2n} \\
 &= \frac{f(x_k) + f(x_{k+1})}{2n}.
 \end{aligned}$$

Posons $T_n = \int_0^1 f_n(x) dx$.

Proposition 6.6.5 *Si f est C^2 sur $[0, 1]$ alors*

$$\left| \int_0^1 f(x) dx - T_n \right| \leq \frac{M_2}{12n^2},$$

où $M_2 = \sup\{|f''(x)| ; x \in [0, 1]\}$.

Sans preuve.

Enfin, parlons de la méthode de Simpson : on approche f par un arc de parabole sur chaque intervalle $[x_k, x_{k+1}]$. En fait, on calcule quel polynôme de degré 2 coïncide avec f aux points x_k, x_{k+1} et $(x_k + x_{k+1})/2$.

Un calcul un peu long donne alors une erreur en $1/n^4$ si f est C^4 .

6.6.4 Lien avec les surfaces

Discutons le lien entre intégrale et surface. Il n'y a maintenant presque plus rien à démontrer, si on se base sur une définition intuitive de la notion de surface. Le problème c'est que, étant donné un domaine borné du plan, il n'est pas forcément évident que l'on puisse toujours lui attribuer une surface, voire même on peut être bien embêté pour définir la surface d'un tel ensemble.

L'idée intuitive que l'on peut avoir est celle qui a toujours prévalu depuis les grecs. On sait bien définir et calculer l'aire des rectangles et des unions de rectangles. On peut facilement en déduire, par des arguments géométriques, l'aire des polygones. Par contre pour des surfaces plus complexes, délimitées par des courbes quelconques, l'idée naturelle est que si on peut encadrer le domaine par deux unions de rectangles de façon à ce que la différence des surfaces soit aussi petite que voulue, alors la surface du domaine existe et elle est égale à cette valeur commune de surface.

Théorème 6.6.6 *Soit f une fonction intégrable et positive sur $[a, b]$, alors $\int_a^b f(x) dx$ est l'aire du domaine*

$$D = \{(x, y) \in \mathbb{R}^2; a \leq x \leq b, 0 \leq y \leq f(x)\}.$$

Pour f de signe quelconque alors $\int_a^b f(x) dx = \int_a^b f_+(x) dx - \int_a^b f_-(x) dx$ est la différence des surfaces des domaines D_+ et D_- associés à f_+ et à f_- .

Démonstration Les sommes de Darboux sont des sommes de surfaces de rectangles qui encadrent le domaine D comme précisé ci-dessus. Si f est intégrable, on a vu que ces sommes convergent vers l'intégrale de f . \square

6.7 Intégrales de suites de fonctions

6.7.1 Ce qui ne marche pas

Dans cette section on va s'intéresser à une famille de problèmes extrêmement importants, mais assez difficiles à comprendre. Voici la problématique générale. On a une suite de fonctions (f_n) , c'est à dire une famille de fonctions définies sur un même intervalle $[a, b]$ et qui dépendent d'un paramètre $n \in \mathbb{N}$. Par exemple :

$$\begin{aligned} f_n(x) &= x^n && \text{sur } [0, 1] \\ g_n(x) &= \frac{\sin(nx)}{\sqrt{n}} && \text{sur } \mathbb{R} \\ h_n(x) &= n^2 x(1-x^2)^n && \text{sur } [0, 1]. \end{aligned}$$

On s'intéresse à la limite *ponctuelle* de ces fonctions, c'est à dire une fonction f sur $[a, b]$ telle que

$$f(x) = \lim_{n \rightarrow +\infty} f_n(x).$$

Dans nos exemples ça donne :

$$f(x) = \begin{cases} 0 & \text{si } x \in [0, 1[, \\ 1 & \text{si } x = 1. \end{cases}$$

$$g(x) = 0$$

$$h(x) = 0.$$

Les principales questions que l'on se pose en général sont : quelles sont les propriétés communes aux f_n qui passent à f ? Si les f_n sont continues, est-ce que f est continue ? Si les f_n sont dérivables, est-ce que f aussi ? Dans ce cas a t-on $f'(x) = \lim_n f'_n(x)$? Même question avec l'intégrabilité ? Est-il vrai que

$$\int_a^b f(x) dx = \lim_{n \rightarrow +\infty} \int_a^b f_n(x) dx ?$$

Nous allons voir avec les exemples ci-dessus que la réponse est non à toutes les questions posées ci-dessus.

- Les f_n sont continues sur $[0, 1]$, la fonction f ne l'est pas.
- Les g_n sont dérivables, de dérivée

$$g'_n(x) = \sqrt{n} \cos(nx).$$

En particulier $g'_n(0) = \sqrt{n}$ qui tend vers $+\infty$, alors que $g(x) = 0$ pour tout x et que donc $g'(0) = 0$.

On a

$$\int_0^1 n^2 x (1 - x^2)^n = \left[\frac{n^2}{-2(n+1)} (1 - x^2)^{n+1} \right]_0^1 = \frac{n^2}{2(n+1)}.$$

Donc sa limite est $+\infty$ alors que l'intégrale de h est nulle.

On voit à travers ces contre-exemples qu'il faut faire très attention quand on passe à la limite sur des fonctions. Essayons de voir dans le détail ce qui ne marche pas.

Le fait que $f(x) = \lim_n f_n(x)$ veut dire que

$$\forall x \in [a, b], \forall \varepsilon > 0, \exists N \in \mathbb{N}; n \geq N \Rightarrow |f_n(x) - f(x)| \leq \varepsilon.$$

Dire que f_n est continue c'est

$$\forall x \in [a, b], \forall \varepsilon > 0, \exists \delta > 0; |x - y| \leq \delta \Rightarrow |f_n(x) - f_n(y)| \leq \varepsilon.$$

Enfin, dire que f est continue c'est

$$\forall x \in [a, b], \forall \varepsilon > 0, \exists \delta > 0; |x - y| \leq \delta \Rightarrow |f(x) - f(y)| \leq \varepsilon.$$

Si on voulait montrer que f est continue, l'idée serait de dire

$$|f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)|.$$

Chacun des trois termes peut être choisi aussi petit qu'on veut, donc on n'aura pas de mal à montrer que $|f(x) - f(y)| \leq \varepsilon$. Mais regardons de plus près.

Soit x fixé. Soit $\varepsilon > 0$ fixé. On veut construire $\delta > 0$ comme dans la définition de la continuité de f . Ce x étant fixé, on sait qu'il existe N tel que $|f_N(x) - f(x)| \leq \varepsilon/3$. Ce N étant fixé on choisit $\delta > 0$ tel que $|f_N(x) - f_N(y)| \leq \varepsilon/3$ pour tout $|x - y| \leq \delta$. Le problème c'est que ce N dépend de x . Le même N ne marchera pas pour les autres points y , on n'a pas de contrôle sur le terme $|f_N(y) - f(y)|$.

On voit bien ce qui ne marche pas. Et même on voit ce qu'il faudrait changer pour que ça marche ! Si dans l'écriture de la limite des f_n ci-dessus, le même N marchait pour tous les x on aurait démontré la continuité de la limite.

6.7.2 Limite uniforme

On dit qu'une suite (f_n) de fonction sur $[a, b]$ converge *uniformément* vers f sur $[a, b]$ si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}; n \geq N \Rightarrow |f_n(x) - f(x)| \leq \varepsilon, \forall x \in [a, b].$$

On note

$$M_n([a, b]) = \sup\{|f_n(x) - f(x)|; x \in [a, b]\}.$$

Théorème 6.7.1 *La suite (f_n) converge uniformément vers f si et seulement si*

$$\lim_{n \rightarrow +\infty} M_n([a, b]) = 0.$$

Démonstration Si (f_n) converge uniformément vers f alors

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}; n \geq N \Rightarrow \sup\{|f_n(x) - f(x)|, x \in [a, b]\} \leq \varepsilon.$$

C'est bien la définition de $\lim_{n \rightarrow +\infty} M_n([a, b]) = 0$.

La réciproque est facile. \square

De manière pratique, comment montre-t-on que (f_n) tend uniformément vers f ? Faisons-le avec deux exemples : $f_n(x) = x^n$ sur $[0, 1]$ et $g_n(x) = \sin(nx)/\sqrt{n}$ sur \mathbb{R} . On calcule d'abord la limite ponctuelle de la suite : $\lim_n f_n(x) = 0$ sur $[0, 1[$ et 1 en 1 alors que $\lim_n g_n(x) = 0$. On regarde $|f_n(x) - f(x)|$ et on regarde si le sup pour tous les x tend vers 0 ou non. Ce qui revient au même que d'essayer de trouver une suite (u_n) qui ne dépend pas de x , qui tend vers 0 et qui vérifie $|f_n(x) - f(x)| \leq u_n$. Ici on trouve

$$\sup\{|f_n(x) - f(x)| ; x \in [0, 1]\} = \sup\{x^n ; x \in [0, 1]\} = 1.$$

On n'a pas convergence uniforme dans ce cas. Pour l'autre on trouve

$$|g_n(x) - g(x)| \leq \frac{1}{\sqrt{n}}.$$

Ce qui prouve la convergence uniforme dans ce cas.

Théorème 6.7.2 *Si (f_n) converge uniformément vers f sur $[a, b]$ et si chacune des fonctions f_n est continue sur $[a, b]$, alors f est continue sur $[a, b]$.*

Démonstration Reprenons ce qu'on avait entamé. Le fait que f est limite uniforme de (f_n) veut dire que

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}; n \geq N \Rightarrow |f_n(x) - f(x)| \leq \varepsilon, \forall x \in [a, b].$$

Dire que f_n est continue c'est

$$\forall x \in [a, b], \forall \varepsilon > 0, \exists \delta > 0; |x - y| \leq \delta \Rightarrow |f_n(x) - f_n(y)| \leq \varepsilon.$$

Enfin, dire que f est continue c'est

$$\forall x \in [a, b], \forall \varepsilon > 0, \exists \delta > 0; |x - y| \leq \delta \Rightarrow |f(x) - f(y)| \leq \varepsilon.$$

Soit $\varepsilon > 0$ fixé et soit n tel que $|f_n(x) - f(x)| \leq \varepsilon/3, \forall x \in [a, b]$. On a

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &\leq \frac{2\varepsilon}{3} + |f_n(x) - f_n(y)|. \end{aligned}$$

Soit x fixé, maintenant on choisit $\delta > 0; |x - y| \leq \delta \Rightarrow |f_n(x) - f_n(y)| \leq \varepsilon/3$. On a $|f(x) - f(y)| \leq \varepsilon$ et on a prouvé la continuité de f au point x . \square

Hélas, la convergence uniforme n'est pas suffisante pour assurer la convergence des dérivées, ni même la dérivabilité de la fonction limite. Par exemple, prenons $g_n(x) = \sin(nx)/\sqrt{n}$ sur $[0, \pi]$. Cette fonction est dérivable sur \mathbb{R} , elle converge uniformément vers la fonction nulle. Pourtant sa dérivée est $\sqrt{n} \cos(nx)$ qui n'est pas identiquement nulle.

En fait pour la dérivabilité, il faut demander la convergence uniforme des dérivées f'_n . Nous donnons ici le résultat à titre informatif seulement.

Théorème 6.7.3 *Soit (f_n) une suite de fonctions dérivables sur $[a, b]$. Si (f_n) converge ponctuellement sur $[a, b]$ vers une fonction f et si (f'_n) converge uniformément sur $[a, b]$ vers une fonction g , alors f est dérivable sur $[a, b]$, de dérivée $f' = g$.*

Avec l'intégration ça marche mieux.

Théorème 6.7.4 *Si (f_n) est une suite de fonctions bornées et intégrables sur $[a, b]$ qui convergent uniformément vers f sur $[a, b]$, alors f est intégrable sur $[a, b]$ et*

$$\int_a^b f(x) dx = \lim_{n \rightarrow +\infty} \int_a^b f_n(x) dx.$$

Démonstration Soit $\varepsilon > 0$ fixé et $N \in \mathbb{N}$ associé à la suite (f_n) dans la définition de la convergence uniforme vers f . En particulier, sur chaque intervalle I_k d'une partition \mathcal{S} donnée on a, pour $n \geq N$

$$\sup_{I_k} f - \inf_{I_k} f \leq \sup_{I_k} f_n - \inf_{I_k} f_n + 2\varepsilon.$$

En particulier

$$S_a^b(f) - s_a^b(f) \leq S_a^b(f_n) - s_a^b(f_n) + 2\varepsilon(b - a).$$

Si on choisit une subdivision telle que $S_a^b(f_n) - s_a^b(f_n) \leq \varepsilon$, on obtient

$$S_a^b(f) - s_a^b(f) \leq (1 + 2(b - a))\varepsilon.$$

Cette quantité peut-être rendue arbitrairement petite. Cela prouve l'intégrabilité de f .

Maintenant, on a

$$\begin{aligned} \left| \int_a^b f(x) dx - \int_a^b f_n(x) dx \right| &\leq \int_a^b |f(x) - f_n(x)| dx \\ &\leq \int_a^b \varepsilon dx \\ &= \varepsilon(b - a). \end{aligned}$$

Ce qui prouve la limite annoncée. □

Chapitre 7

Equations différentielles