

Séries, intégrales et probabilités

Thierry MEYRE

Préparation à l'agrégation interne. Année 2017-2018.

Université Paris Diderot. IREM.

<https://www.lpsm.paris/dw/doku.php?id=users:meyre:index>

Table des matières

1	Intégrale de Riemann.	11
1.1	Applications réglées.	11
1.2	Construction de l'intégrale de Riemann	16
1.2.1	Rappels sur les applications linéaires continues	16
1.2.2	Intégrale d'une application en escalier	17
1.2.3	Intégrale d'une application réglée	18
1.2.4	Propriétés élémentaires	20
1.2.5	Intégrale d'une application à valeurs réelles	22
1.3	Outils pratiques de calcul d'une intégrale	26
1.3.1	Utilisation d'une primitive	26
1.3.2	Intégration par parties	30
1.3.3	Changement de variable	32
1.3.4	Exercices suggérés.	33
1.3.5	Applications suggérées.	35
1.4	Sommes de Riemann	37
1.5	Convergences de suites d'applications	40
1.6	Intégrale d'une fonction dépendant d'un paramètre	44
1.7	Arcs paramétrés	46
1.7.1	Rectification d'un arc paramétré.	46
1.7.2	Intégrale curviligne d'une forme différentielle.	46
2	Intégrales impropres.	47
2.1	Définition des intégrales impropres.	47
2.2	Étude de la convergence : cas positif	49
2.2.1	Comparaison avec une intégrale de référence.	50
2.2.2	Intégration des relations de comparaison	52
2.2.3	Comparaison avec une série	55
2.3	Cas général : critère de Cauchy et autres méthodes.	56
2.3.1	Critère de Cauchy et conséquences.	56
2.3.2	Méthodes directes	59
2.4	Calcul des intégrales impropres	60

2.4.1	Intégration par parties généralisée	60
2.4.2	Changement de variable généralisé	62
2.5	Application à l'étude de séries.	62
2.6	Transformée de Laplace à variable réelle	67
2.6.1	Quelques propriétés générales	67
2.6.2	Transformée de Laplace et calcul des probabilités	69
3	Intégration sur un intervalle quelconque	71
3.1	Fonctions intégrables	71
3.1.1	Le cas positif	71
3.1.2	Le cas réel	72
3.1.3	Le cas complexe	73
3.2	Théorèmes de convergence	73
3.3	Intégrale dépendant d'un paramètre	80
3.3.1	Étude de la continuité	80
3.3.2	Étude de la dérivabilité	84
3.4	La fonction Γ d'Euler	86
3.5	Transformation de Laplace à variable complexe	90
4	Intégrales multiples	93
4.1	Intégrale double sur un domaine simple borné	93
4.2	Intégrale double sur un produit d'intervalles quelconques . . .	94
4.2.1	Le cas positif	95
4.2.2	Le cas réel ou complexe	96
4.3	Changement de variable	98
5	Méthodes de calcul approché d'une intégrale	101
5.1	Intégrale d'une fonction polynomiale.	102
5.1.1	Rang et ordre d'une méthode de quadrature	102
5.1.2	Exemples	104
5.1.3	Méthodes de Newton-Cotes (subdivision régulière) . . .	105
5.2	Interpolation polynomiale.	108
5.2.1	Interpolation de Lagrange et Hermite	108
5.2.2	Évaluation de l'erreur	109
5.3	Intégration numérique	111
5.3.1	Avec un polynôme de Lagrange	111
5.3.2	Avec un polynôme d'Hermite	112
5.3.3	Méthodes composées.	115
5.3.4	Lien avec la méthode de Romberg-Richardson	116
5.4	Méthodes de quadrature de Gauss.	116
5.5	Une méthode probabiliste	118

6	Applications de l'analyse au calcul des grandeurs	119
6.1	Longueur d'un arc de classe C^1	119
6.1.1	Rectification d'un arc paramétré.	119
6.1.2	Calcul pratique	121
6.2	Produit mixte et produit vectoriel dans \mathbb{R}^n	124
6.2.1	Orientation de \mathbb{R}^n	124
6.2.2	Produit mixte dans \mathbb{R}^n euclidien orienté	125
6.2.3	Produit vectoriel dans \mathbb{R}^n euclidien orienté	126
6.2.4	Interprétations géométriques en dimension 3	127
6.3	Théorème du changement de variables	128
6.3.1	Enoncé général	128
6.3.2	Quelques cas particuliers	130
6.4	Formule de Green-Riemann	133
6.5	Aire d'une nappe géométrique	134
6.5.1	Nappe paramétrée, nappe géométrique	134
6.5.2	Aire d'une nappe géométrique	135
7	Espaces de Banach, suites et séries	137
7.1	Définition et premiers exemples	137
7.2	Séries à termes réels positifs	138
7.2.1	Définitions et théorèmes de comparaison	138
7.2.2	Comparaison à une série géométrique	141
7.2.3	Série et intégrale impropre positives.	143
7.3	Séries à valeurs dans un espace de Banach	144
7.4	Espaces de Banach usuels de suites et de fonctions	146
7.5	Suites d'applications à valeurs dans un espace de Banach	149
7.6	Séries d'applications à valeurs dans un espace de Banach	152
8	Exponentielle dans une algèbre de Banach	155
8.1	Définitions et premières propriétés	155
8.2	Exponentielle de matrice	158
8.2.1	Calcul explicite	159
8.2.2	Quelques applications topologiques	163
8.2.3	Application aux systèmes différentiels linéaires	164
9	Espaces préhilbertiens	167
9.1	Produit scalaire sur un \mathbb{K} -espace vectoriel.	167
9.2	Norme euclidienne, norme hermitienne	169
9.3	Orthogonalité et procédé de Gram-Schmidt	171
9.4	Projection orthogonale, meilleure approximation	173
9.5	Inégalité de Bessel et égalité de Parseval	176

10 Séries de Fourier	179
10.1 Polynôme et série trigonométriques	179
10.2 Coefficients et série de Fourier	182
10.3 Le théorème de convergence de Dirichlet	183
10.4 Le théorème de Fejér	190
10.5 Approximation en moyenne quadratique	193
11 Séries entières	197
11.1 L'algèbre des séries entières	197
11.2 Rayon de convergence	198
11.3 Continuité et dérivabilité	202
11.4 Principe des zéros isolés	205
11.5 Formule de Cauchy, égalité de Parseval	207
11.6 Fonctions développables en série entière	209
12 Espaces mesurés. Espaces probabilisés.	213
12.1 Clans et tribus.	213
12.1.1 Définitions.	213
12.1.2 Clans et tribus engendrés.	215
12.2 Mesures positives.	216
12.2.1 Définitions.	216
12.2.2 Propriétés d'une mesure.	217
12.3 Application à la modélisation du hasard.	220
12.3.1 Un modèle simple : le modèle additif.	220
12.3.2 Modèle définitif : les espaces probabilisés.	225
12.4 Probabilités conditionnelles. Indépendance d'événements.	227
12.4.1 Définition des probabilités conditionnelles.	227
12.4.2 Formule de Bayes.	230
12.4.3 Indépendance d'événements.	231
12.5 Exercices.	234
12.5.1 Clans.	234
12.5.2 Formule de Poincaré.	235
12.5.3 Équiprobabilité.	236
12.5.4 Probabilités conditionnelles.	237
12.5.5 Événements indépendants.	238
13 Variables aléatoires réelles	243
13.1 La tribu borélienne réelle	243
13.2 Les variables aléatoires réelles et leurs lois	245
13.3 Fonction de répartition d'une v.a.r.	248
13.4 Variables discrètes	250

14 Variables à densité	255
14.1 Définitions	255
14.2 Exemples classiques de lois à densité	256
14.3 Fonction de répartition	258
14.4 Espérance. Théorème de transfert	260
14.5 Moments d'une variable aléatoire	261
14.6 Exercices sur les fonctions de répartition	262
14.6.1 Minimum de variables exponentielles.	262
14.6.2 Variables amnésiques.	262
14.6.3 Loi du χ^2 à un degré de liberté.	262
14.6.4 Loi gaussienne dans \mathbb{R}	263
14.6.5 Avec une loi de Cauchy.	263
15 Vecteurs aléatoires et indépendance	265
15.1 Les vecteurs aléatoires et leurs lois	266
15.2 Vecteurs aléatoires discrets	268
15.3 Vecteurs aléatoires à densité	268
15.4 Indépendance de p v.a.r.	272
15.5 Produit de convolution de densités	274
15.6 Exercices	277
15.6.1 Calcul de densité marginale.	277
15.6.2 Densité gaussienne en dimension 2.	277
15.6.3 Loi uniforme sur le disque	277
15.6.4 Gauss et Cauchy.	278
15.6.5 Avec des lois exponentielles.	278
15.6.6 Méthode de Box-Muller	278
16 Loi des grands nombres.	281
16.1 Suite indépendante équadistribuée.	282
16.2 Convergences de suites de v.a.r.	285
16.2.1 Événements négligeables, égalité presque sûre.	285
16.2.2 Convergences en probabilité et presque sûre.	286
16.3 Loi des grands nombres	288
17 Loi normale	291
17.1 La loi normale centrée réduite	291
17.2 La loi normale générale	293
17.3 Approximation normale de la loi binomiale	295

18 Le théorème-limite central	299
18.1 Énoncé classique	300
18.2 TLC et fonctions de répartition	301
18.3 Un énoncé sans équidistribution	304
18.4 Test du χ^2 d'ajustement.	305

Séries et intégrales : bibliographie

Ouvrages de référence

CMP Analyse MP. Cours (avec exercices-types, méthodes, exercices et problèmes corrigés). Monier. Éditions Dunod. 5ème édition. 2007

CMPSI Analyse MPSI. Cours, méthodes et exercices corrigés. Monier. Éditions Dunod. 5ème édition. 2006

DAN Mathématiques pour l'agrégation interne. Analyse et Probabilités. Cours et exercices corrigés. Dantzer. Éditions Vuibert. 2007.

EMP Analyse MP. 200 exercices développés. 800 exercices d'entraînement. Rappels de cours. Monier. Éditions Dunod. 2ème édition. 2004

EMPSI Les méthodes et exercices de Mathématiques MPSI. Monier. Éditions Dunod. 2008.

POM Agrégation de Mathématiques. Cours d'analyse. Pommelet. Éditions Ellipses. 1994

RUD Principes d'analyse mathématique. Rudin. Éditions Dunod. 2006

SKA Analyse - Résumés et exercices. Skandalis.

Ouvrages plus difficiles

BP Théorie de l'intégration. Briane, Pagès. Éditions Vuibert. 4ème édition. 2006

GRA Intégration. Gramain. Éditions Hermann. 1998

MAD Leçons d'analyse. Madère. Éditions Ellipses. 1997

Chapitre 1

Intégrale de Riemann.

Il existe différentes théories de l'intégration qui s'appliquent à des classes de fonctions plus ou moins vastes. La théorie que nous allons examiner dans ce cours a été formulée par Bernhard Riemann, professeur à l'université de Göttingen (Allemagne), en 1854. Elle a le mérite de reposer sur une construction relativement simple et s'applique bien aux fonctions continues par morceaux, qui sont les seules au programme officiel.

Dans ce chapitre, nous considérons des applications définies sur un intervalle $[a, b]$ et à valeurs dans un espace de Banach E sur le corps $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

Dans la pratique, nous intégrerons très souvent des applications à valeurs dans \mathbb{R} ou \mathbb{C} mais la construction de l'intégrale de Riemann pour des applications à valeurs dans un espace de Banach général n'apporte aucune difficulté supplémentaire.

De même, nous allons définir l'intégrale de Riemann sur une classe plus large que celle des applications continues par morceaux, à savoir la classe des applications *réglées*, car cela ne nous demandera aucun effort de construction supplémentaire.

1.1 Applications réglées.

Définition 1.1.1 On appelle *subdivision* d'un intervalle réel $[a, b]$ toute famille finie $\sigma = (a_i)_{0 \leq i \leq n}$ d'éléments du segment $[a, b]$ telle que :

$$a = a_0 < a_1 < \cdots < a_n = b$$

On appelle pas de la subdivision σ le réel positif

$$\|\sigma\| = \max_{1 \leq i \leq n} (a_i - a_{i-1})$$

On dit qu'une subdivision σ' est plus fine qu'une subdivision σ si tous les éléments de σ appartiennent à σ' .

Nous allons maintenant définir deux classes de fonctions de $[a, b]$ dans E qui nous seront utiles dans la construction de l'intégrale de Riemann.

Définition 1.1.2 On dit qu'une application $f : [a, b] \rightarrow E$ est en escalier s'il existe une subdivision $\sigma = (a_i)_{0 \leq i \leq n}$ de $[a, b]$ et des constantes $(c_i)_{0 \leq i \leq n-1} \in E^n$ telles que, pour tout $0 \leq i \leq n-1$, $f|_{]a_i, a_{i+1}[} \equiv c_i$.

On dit que $f : [a, b] \rightarrow E$ est une application continue par morceaux s'il existe une subdivision $\sigma = (a_i)_{0 \leq i \leq n}$ de $[a, b]$ et des applications f_0, \dots, f_{n-1} telles que, pour tout $0 \leq i \leq n-1$, f_i est définie et continue sur $[a_i, a_{i+1}]$ avec

$$\forall x \in]a_i, a_{i+1}[\quad f(x) = f_i(x)$$

Dans les deux cas, on dit que la subdivision σ est adaptée à l'application en escalier (resp. continue par morceaux) f .

Remarques.

1. Cette dernière condition équivaut au fait que, pour tout $0 \leq i \leq n-1$, l'application $f|_{]a_i, a_{i+1}[}$ est continue **et** prolongeable par continuité aux extrémités de l'intervalle. Ainsi, l'application suivante $f : [-1, 1] \rightarrow \mathbb{R}$ est continue sur $] -1, 0[$ et $] 0, 1[$ mais n'est pas continue par morceaux : $f(x) = \frac{1}{x}$ si $x \neq 0$ et $f(0) = 0$.
2. On peut avoir $f(a_i) \neq f_i(a_i)$. L'hypothèse de continuité sur les applications f_i nous donne les relations :

$$\forall i \in \llbracket 0; n-1 \rrbracket, \quad f_i(a_i) = f(a_i+) \quad \text{et} \quad f_i(a_{i+1}) = f(a_{i+1}-).$$

3. Si σ est adaptée à f , alors toute subdivision σ' plus fine que σ est encore adaptée à f .
4. On en déduit que si f et g sont deux applications en escalier (resp. continues par morceaux), avec des subdivisions adaptées respectives σ et σ' , alors $\sigma \cup \sigma'$ est une subdivision adaptée aux deux applications f et g à la fois.
5. Dans le cas particulier où $E = \mathbb{R}$ ou \mathbb{C} , le lecteur pourra montrer à titre d'exercice que le produit de deux applications en escalier (resp. continues par morceaux) est encore une application en escalier (resp. continue par morceaux).
6. Pour tout $k \in \mathbb{N}^* \cup \{+\infty\}$, on définit de façon analogue la notion d'application de classe C^k par morceaux.

Notations et rappels : Nous noterons $C([a, b], E)$, resp. $C_M([a, b], E)$, $\mathcal{E}([a, b], E)$, $\mathcal{B}([a, b], E)$ l'ensemble des applications $f : [a, b] \rightarrow E$ continues, resp. continues par morceaux, en escalier, bornées. Nous rappelons que $(\mathcal{B}([a, b], E), +, \cdot)$ est un espace vectoriel que nous munirons de la norme suivante :

$$\|f\|_\infty = \sup_{x \in [a, b]} \|f(x)\|_E$$

Cette norme est encore appelée norme *de la convergence uniforme* pour une raison que nous allons rappeler maintenant.

Définitions 1.1.3 Soit $(f_n)_{n \in \mathbb{N}}$ et f des applications définies sur l'intervalle $[a, b]$ et à valeurs dans l'espace de Banach E .

On dit que la suite d'applications $(f_n)_{n \in \mathbb{N}}$ converge simplement vers l'application f sur l'intervalle $[a, b]$ si

$$\forall x \in [a, b], \quad \forall \epsilon > 0, \quad \exists N \in \mathbb{N}, \quad \forall n \geq N, \quad \|f_n(x) - f(x)\| \leq \epsilon.$$

On dit que la suite d'applications $(f_n)_{n \in \mathbb{N}}$ converge uniformément vers l'application f sur l'intervalle $[a, b]$ si

$$\forall \epsilon > 0, \quad \exists N \in \mathbb{N}, \quad \forall n \geq N, \quad \forall x \in [a, b], \quad \|f_n(x) - f(x)\| \leq \epsilon. \quad (1.1)$$

Remarques.

1. La convergence simple de (f_n) vers f sur $[a, b]$ équivaut donc à dire que pour tout $x \in [a, b]$, la suite $(f_n(x))_{n \in \mathbb{N}}$ converge vers $f(x)$ dans E .
2. La convergence uniforme est plus forte que la convergence simple, la différence venant de ce que N ne dépend pas de x dans le cas de la convergence uniforme, contrairement à ce qui se passe pour la convergence simple. Ainsi, si l'on prend $f_n(x) = 1$ si $x \in]0, 1/n[$ et $f_n(x) = 0$ sinon, la suite $(f_n)_{n \in \mathbb{N}^*}$ converge simplement vers $f \equiv 0$ sur \mathbb{R} mais ne converge pas uniformément vers f sur \mathbb{R} .

Proposition 1.1.4 Soit (f_n) une suite d'applications bornées de $[a, b]$ dans E qui converge uniformément vers une application $f : [a, b] \rightarrow E$.

Alors f est une application bornée et $\|f_n - f\|_\infty \rightarrow 0$.

Démonstration: En prenant $\epsilon = 1$ dans la définition de la convergence uniforme, nous avons en particulier :

$$\exists N \in \mathbb{N}, \quad \forall n \geq N, \quad \forall x \in [a, b], \quad \|f_n(x) - f(x)\| \leq 1.$$

En particulier, nous avons la majoration :

$$\forall x \in [a, b], \quad \|f_N(x) - f(x)\| \leq 1.$$

Or, par hypothèse, l'application f_N est bornée, disons par M . L'inégalité triangulaire nous donne alors :

$$\forall x \in [a, b], \quad \|f(x)\| \leq \|f(x) - f_N(x)\| + \|f_N(x)\| \leq 1 + M,$$

si bien que l'application f est bornée.

Ainsi, toutes les applications en jeu sont bornées et nous pouvons raisonner dans l'espace vectoriel normé $(\mathcal{B}([a, b], E), +, \cdot, \|\cdot\|_\infty)$.

La formule (1.1) se réécrit :

$$\forall \epsilon > 0, \quad \exists N \in \mathbb{N}, \quad \forall n \geq N, \quad \|f_n - f\|_\infty \leq \epsilon,$$

ce qui nous donne bien la convergence annoncée. □

Remarque. L'espace vectoriel normé $(\mathcal{B}([a, b], E), \|\cdot\|_\infty)$ est un espace de Banach car il "hérite" de la complétude de E . Preuve : [DAN 43].

Il est facile de vérifier qu'une application continue par morceaux est bornée, en utilisant le fait que chacune des applications f_0, \dots, f_{n-1} de la définition est continue sur un intervalle compact donc bornée.

Nous constatons alors que $\mathcal{E}([a, b], E)$ est un sous-espace vectoriel de $C_M([a, b], E)$, qui est lui-même un sous-espace vectoriel de $\mathcal{B}([a, b], E)$.

Nous allons maintenant prouver l'égalité suivante :

$$C_M([a, b], E) = C([a, b], E) + \mathcal{E}([a, b], E)$$

Commençons en introduisant une définition simple mais bien utile dans la pratique.

Définition 1.1.5 Soit F un ensemble quelconque et A une partie de F . On appelle fonction indicatrice de A l'application notée $\mathbf{1}_A : F \rightarrow \{0, 1\}$ définie par $\mathbf{1}_A(x) = 1$ si $x \in A$ et $\mathbf{1}_A(x) = 0$ si $x \notin A$.

Remarque. Cette application est appelée dans certains ouvrages d'analyse "fonction caractéristique" de A mais nous n'utiliserons pas ce vocable dans ce cours car la notion de fonction caractéristique en calcul des probabilités est tout autre.

Proposition 1.1.6 Pour toute application $f \in C_M([a, b], E)$, il existe une application $f_c \in C([a, b], E)$ et une application $f_e \in \mathcal{E}([a, b], E)$ telles que :

$$f = f_c + f_e$$

Démonstration: Nous allons prouver par récurrence sur $n \in \mathbb{N}$ la proposition suivante :

\mathcal{P}_n : ‘Toute application $f \in C_M([a, b], E)$ admettant au plus n points de discontinuité peut s’écrire sous la forme annoncée’

La proposition \mathcal{P}_0 est trivialement vraie. Montrons maintenant que \mathcal{P}_{n-1} implique \mathcal{P}_n .

Soit f admettant au plus n points de discontinuités et x_0 l’un de ces points. Nous définissons l’application $g \in \mathcal{E}([a, b], E)$ par :

$$\forall x \in [a, b], \quad g(x) = f(x_0-) \mathbf{1}_{[a, x_0[}(x) + f(x_0) \mathbf{1}_{\{x_0\}}(x) + f(x_0+) \mathbf{1}_{]x_0, b]}(x)$$

On vérifie facilement que l’application $h = f - g$ admet au plus $n - 1$ points de discontinuité donc par hypothèse de récurrence, on peut écrire $h = h_c + h_e$. Nous avons donc $f = h_c + h_e + g$ et nous concluons en posant :

$$f_c = h_c \quad , \quad f_e = h_e + g$$

□

À titre d’exercice, le lecteur pourra montrer que les applications f_c et f_e sont uniques à constante additive près.

Définition 1.1.7 On dit qu’une application $f : [a, b] \rightarrow E$ est réglée si elle est limite uniforme d’une suite d’applications en escalier.

Autrement dit, l’ensemble $\mathcal{R}([a, b], E)$ des applications réglées est le sous-espace vectoriel de $\mathcal{B}([a, b], E)$ défini comme l’adhérence de $\mathcal{E}([a, b], E)$ pour la norme de la convergence uniforme :

$$\mathcal{R}([a, b], E) = \overline{\mathcal{E}([a, b], E)}^{\|\cdot\|_\infty}$$

Proposition 1.1.8 Toute application $f : [a, b] \rightarrow E$ continue par morceaux est réglée. Autrement dit, on a l’inclusion : $C_M([a, b], E) \subset \mathcal{R}([a, b], E)$.

Démonstration: Nous avons évidemment $\mathcal{E}([a, b], E) \subset \mathcal{R}([a, b], E)$ donc, d’après la proposition précédente, il suffit de prouver l’inclusion :

$$C([a, b], E) \subset \mathcal{R}([a, b], E)$$

Considérons donc $f : [a, b] \rightarrow E$ continue et, pour tout $n \in \mathbb{N}^*$, définissons la fonction en escalier $f_n : [a, b] \rightarrow E$ par

$$\forall x \in [a, b], \quad f_n(x) = \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) \mathbf{1}_{[a+k \frac{b-a}{n}, a+(k+1) \frac{b-a}{n}]}(x) + f(b) \mathbf{1}_{\{b\}}(x)$$

Puisque f est continue sur l'intervalle compact $[a, b]$, elle est uniformément continue d'après le théorème de Heine. Ainsi, pour $\epsilon > 0$ arbitraire, on peut trouver $\alpha > 0$ tel que

$$\forall (x, y) \in [a, b]^2 \quad |x - y| \leq \alpha \implies \|f(x) - f(y)\| \leq \epsilon$$

Posons $N = \left\lceil \frac{b-a}{\alpha} \right\rceil + 1$ de sorte que $n \geq N \implies \frac{b-a}{n} \leq \alpha$. Nous avons alors :

$$\forall n \geq N, \quad \|f - f_n\|_\infty \leq \epsilon$$

Comme $\epsilon > 0$ était arbitraire, ceci prouve que f est limite uniforme de la suite (f_n) . □

On peut caractériser les fonctions réglées par la propriété suivante. Nous renvoyons le lecteur à [SK 57] pour la démonstration.

Proposition 1.1.9 *Une application $f : [a, b] \rightarrow E$ est réglée si et seulement si elle admet une limite à droite en tout point de $[a, b[$ et une limite à gauche en tout point de $]a, b]$.*

Remarque. Nous déduisons de cette caractérisation des applications réglées que toute application monotone $f : [a, b] \rightarrow \mathbb{R}$ est réglée et qu'un produit de deux applications réglées de $[a, b]$ dans \mathbb{R} (ou \mathbb{C}) est encore une application réglée.

Exemple. L'application $f : [0, 1] \rightarrow \mathbb{R}$ définie par $f(x) = (E[x^{-1}])^{-1}$ pour $0 < x \leq 1$ et $f(0) = 0$ est réglée : on peut le prouver directement d'après la définition ou bien constater qu'elle est croissante.

En revanche, elle n'est pas continue par morceaux. En effet, l'ensemble de ses discontinuités est infini dénombrable : c'est $\{1/n, n \geq 2\}$.

Nous constatons donc que l'inclusion énoncée dans la proposition 1.1.8 est stricte : $C_M([a, b], E) \subsetneq \mathcal{R}([a, b], E)$.

1.2 Construction de l'intégrale de Riemann

1.2.1 Rappels sur les applications linéaires continues

Soient $(F, \|\cdot\|_F)$ et $(G, \|\cdot\|_G)$ deux \mathbb{K} -e.v.n. Nous noterons $\mathcal{L}(F, G)$, resp. $\mathcal{L}_c(F, G)$ le \mathbb{K} -espace vectoriel des applications linéaires, resp. linéaires continues de F dans G . Rappelons le théorème suivant dont le lecteur pourra retrouver la démonstration dans [DAN 167].

Théorème 1.2.1 Soit $f \in \mathcal{L}(F, G)$. Il y a équivalence entre :

1. f est lipschitzienne sur F ;
2. f est uniformément continue sur F ;
3. f est continue sur F ;
4. f est continue en 0 ;
5. f est bornée sur la boule unité fermée de F ;
6. f est bornée sur la sphère unité de F ;
7. Il existe une constante $M \geq 0$ telle que $\forall x \in F, \quad \|f(x)\|_G \leq M\|x\|_F$.

Définition et proposition 1.2.2 Soient $(F, \|\cdot\|_F)$ et $(G, \|\cdot\|_G)$ deux \mathbb{K} -espaces vectoriels normés. On définit une norme sur $\mathcal{L}_c(F, G)$ en posant :

$$\|u\| = \sup_{x \neq 0} \frac{\|u(x)\|_G}{\|x\|_F} = \sup_{\|x\|_F=1} \|u(x)\|_G.$$

On l'appelle norme subordonnée aux normes $\|\cdot\|_F$ et $\|\cdot\|_G$.

1.2.2 Intégrale d'une application en escalier

Proposition et définition 1.2.3 Soit $f : [a, b] \rightarrow E$ une application en escalier et $\sigma = (a_i)_{0 \leq i \leq n}$ une subdivision adaptée à f :

$$\forall 0 \leq i \leq n-1, \quad f|_{]a_i, a_{i+1}[} \equiv c_i.$$

Alors la somme $\sum_{i=0}^{n-1} (a_{i+1} - a_i) c_i$ est un élément de E indépendant de la subdivision σ choisie. On l'appelle intégrale sur $[a, b]$ de l'application f et l'on note :

$$\int_a^b f(x) \, dx = \sum_{i=0}^{n-1} (a_{i+1} - a_i) c_i$$

Démonstration: Pour prouver que les sommes associées à deux subdivisions adaptées σ et σ' sont identiques, on compare chacune d'entre elles à la somme associée à la subdivision plus fine $\sigma \cup \sigma'$, ce qui permet de conclure facilement. \square

Proposition 1.2.4 Pour toute application $f \in \mathcal{E}([a, b], E)$, on a $\|f\| \in \mathcal{E}([a, b], \mathbb{R})$ et

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

Démonstration: Si $\sigma = (a_i)_{0 \leq i \leq n}$ est une subdivision adaptée à f , nous constatons immédiatement qu'elle est aussi adaptée à l'application $\|f\|$ qui est encore en escalier. L'inégalité que nous cherchons à prouver s'écrit donc

$$\left\| \sum_{i=0}^{n-1} (a_{i+1} - a_i) c_i \right\| \leq \sum_{i=0}^{n-1} (a_{i+1} - a_i) \|c_i\|$$

et résulte de l'inégalité triangulaire. \square

Proposition 1.2.5 *L'application $I : \mathcal{E}([a, b], E) \rightarrow E$ définie par*

$$\forall f \in \mathcal{E}([a, b], E) \quad I(f) = \int_a^b f(x) \, dx$$

est linéaire et continue, de norme subordonnée $\|I\|_{\mathcal{L}} = b - a$.

Démonstration: On vérifie immédiatement que

$$\forall (\lambda, \mu) \in \mathbb{R}^2 \text{ (ou } \mathbb{C}^2 \text{)} \quad \forall (f, g) \in \mathcal{E}([a, b], E)^2 \quad I(\lambda f + \mu g) = \lambda I(f) + \mu I(g)$$

en effectuant le calcul avec une subdivision σ adaptée à la fois à f et g . La proposition précédente nous donne

$$\left\| \int_a^b f(x) \, dx \right\| \leq \sum_{i=0}^{n-1} (a_{i+1} - a_i) \|c_i\| \leq \sum_{i=0}^{n-1} (a_{i+1} - a_i) \|f\|_{\infty} = (b - a) \|f\|_{\infty}$$

donc I est une application linéaire continue de norme au plus égale à $(b - a)$. Il nous reste à remarquer que l'égalité est obtenue pour une fonction constante non nulle, ce qui nous permet de conclure. \square

1.2.3 Intégrale d'une application réglée

Pour prolonger l'application linéaire continue I de $\mathcal{E}([a, b], E)$ à $\mathcal{R}([a, b], E)$, nous allons utiliser un procédé général énoncé dans la proposition suivante.

Proposition 1.2.6 *Soit F un \mathbb{K} -espace vectoriel normé, soient F_1 un sous-espace dense de F et E un \mathbb{K} -espace de Banach. Alors, toute application linéaire continue $u : F_1 \rightarrow E$ se prolonge de manière unique en une application linéaire continue $\hat{u} : F \rightarrow E$. En outre, \hat{u} et u ont même norme subordonnée : $\|\hat{u}\| = \|u\|$.*

Démonstration: Soit $x \in F$; il existe une suite $(x_n) \in F_1^{\mathbb{N}}$ qui converge vers x . La suite (x_n) étant convergente dans F , elle est de Cauchy et l'inégalité

$$\|u(x_n) - u(x_m)\| \leq \|u\| \|x_n - x_m\|$$

prouve que la suite $(u(x_n))$ est de Cauchy dans l'espace de Banach E donc convergente.

Nous pouvons poser $\hat{u}(x) = \lim u(x_n)$ sans ambiguïté car la valeur de cette limite ne dépend pas du choix de la suite (x_n) qui converge vers x ; en effet, si (x_n) et (y_n) convergent toutes deux vers x , alors $\|x_n - y_n\| \rightarrow 0$ et l'inégalité $\|u(x_n) - u(y_n)\| \leq \|u\| \|x_n - y_n\|$ implique $\lim u(x_n) = \lim u(y_n)$.

L'application $\hat{u} : F \rightarrow E$ ainsi définie est linéaire puisque, si (x_n) converge vers x et (y_n) converge vers y , alors la suite $(\lambda x_n + \mu y_n)$ converge vers $\lambda x + \mu y$ d'où :

$$\hat{u}(\lambda x + \mu y) = \lim u(\lambda x_n + \mu y_n) = \lim (\lambda u(x_n) + \mu u(y_n)) = \lambda \hat{u}(x) + \mu \hat{u}(y)$$

Si (x_n) converge vers x , en passant à la limite dans l'inégalité

$$\|u(x_n)\| \leq \|u\| \|x_n\|,$$

nous obtenons $\|\hat{u}(x)\| \leq \|u\| \|x\|$. Ceci étant vrai pour $x \in F$ arbitraire, nous en déduisons la continuité de \hat{u} , puis l'inégalité entre normes subordonnées :

$$\|\hat{u}\| \leq \|u\|.$$

Par ailleurs, en notant S (resp. S_1) la sphère unité de F (resp. F_1), on a :

$$\|\hat{u}\| = \sup_{x \in S} \|\hat{u}(x)\| \geq \sup_{x \in S_1} \|\hat{u}(x)\| = \sup_{x \in S_1} \|u(x)\| = \|u\|,$$

d'où finalement $\|\hat{u}\| = \|u\|$.

Il nous reste à démontrer l'unicité de ce prolongement. Soit donc $v : F \rightarrow E$ une application linéaire continue telle que $\forall x \in F_1, v(x) = u(x)$.

Soit $x \in F$ quelconque; alors il existe $(x_n) \in F_1^{\mathbb{N}}$ qui converge vers x .

A fortiori, (x_n) est une suite d'éléments de F qui converge vers x et donc, par continuité de v sur F , $v(x) = \lim v(x_n)$.

D'autre part, pour tout $n \in \mathbb{N}$, $x_n \in F_1$ et donc $v(x_n) = u(x_n)$.

En définitive, nous obtenons pour $x \in F$ quelconque :

$$v(x) = \lim v(x_n) = \lim u(x_n) = \hat{u}(x)$$

□

Nous en déduisons immédiatement la proposition suivante, qui nous permet de définir l'intégrale de Riemann d'une application réglée.

Proposition 1.2.7 *L'application $I : \mathcal{E}([a, b], E) \rightarrow E$ définie par*

$$\forall f \in \mathcal{E}([a, b], E), \quad I(f) = \int_a^b f(x) \, dx,$$

se prolonge de façon unique en une application linéaire continue $\hat{I} : \mathcal{R}([a, b], E) \rightarrow E$ de norme subordonnée $\|\hat{I}\| = b - a$.

Pour toute application $f : [a, b] \rightarrow E$ réglée, nous appellerons intégrale de Riemann de f sur $[a, b]$ la valeur $\hat{I}(f)$ et nous noterons

$$\hat{I}(f) = \int_a^b f(x) \, dx.$$

Corollaire 1.2.8 (Inégalité de la moyenne)

$$\forall f \in \mathcal{R}([a, b], E) \quad \left\| \int_a^b f(x) \, dx \right\| \leq (b - a) \|f\|_\infty$$

Remarques.

1. On appelle *moyenne* de l'application f sur l'intervalle $[a, b]$ l'élément de E suivant :

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

Sa norme (dans E) est donc majorée par $\|f\|_\infty$.

2. Par convention, nous poserons

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

1.2.4 Propriétés élémentaires

Proposition 1.2.9 (Relation de Chasles) *Soit $f \in \mathcal{R}([a, b], E)$ et soit $c \in]a, b[$; alors, on a l'égalité*

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

Démonstration: Si $f \in \mathcal{E}([a, b], E)$, il suffit de choisir une subdivision σ adaptée à f et contenant le point c (ce qui est toujours possible quitte à rajouter ce point à la subdivision) pour obtenir l'égalité.

Dans le cas général, il existe une suite (f_n) d'éléments de $\mathcal{E}([a, b], E)$ qui converge uniformément vers f et nous concluons en passant à la limite dans l'égalité

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

□

Proposition 1.2.10

$$\forall f \in \mathcal{R}([a, b], E), \quad \left\| \int_a^b f(x) dx \right\| \leq \int_a^b \|f(x)\| dx.$$

Démonstration: Nous prenons une suite (f_n) d'éléments de $\mathcal{E}([a, b], E)$ qui converge uniformément vers f et nous écrivons l'inégalité donnée par la proposition 1.2.4 :

$$\left\| \int_a^b f_n(x) dx \right\| \leq \int_a^b \|f_n(x)\| dx. \quad (1.2)$$

Remarquons maintenant que $(\|f_n\|)$ est une suite d'éléments de $\mathcal{E}([a, b], \mathbb{R})$ qui converge uniformément vers $\|f\|$ puisque, par inégalité triangulaire,

$$\forall n \in \mathbb{N}, \quad \forall x \in [a, b], \quad \left| \|f_n(x)\| - \|f(x)\| \right| \leq \|f_n(x) - f(x)\| \leq \|f_n - f\|_\infty.$$

Nous en déduisons que $\|f\| \in \mathcal{R}([a, b], \mathbb{R})$ et que, par définition même de l'intégrale de Riemann, nous avons l'égalité :

$$\int_a^b \|f(x)\| dx = \lim \int_a^b \|f_n(x)\| dx.$$

Il nous reste alors à passer à la limite dans l'égalité (1.2) pour conclure.

□

Exercice. Soit $f \in \mathcal{R}([a, b], \mathbb{C})$. Nous savons qu'il existe une suite (f_n) d'éléments de $\mathcal{E}([a, b], \mathbb{C})$ qui converge uniformément vers f . Montrer que $(\Re f_n)$ est une suite dans $\mathcal{E}([a, b], \mathbb{R})$ qui converge uniformément vers $\Re f$. Procéder à un raisonnement similaire avec $(\Im f_n)$. En déduire l'égalité :

$$\int_a^b g(t) dt = \int_a^b \Re g(t) dt + i \int_a^b \Im g(t) dt.$$

1.2.5 Intégrale d'une application à valeurs réelles

Nous allons maintenant énoncer des propriétés de l'intégrale de Riemann spécifiques au cas où f est une application à valeurs dans $E = \mathbb{R}$.

Proposition 1.2.11 *Soient $f, g : [a, b] \rightarrow \mathbb{R}$ deux applications réglées. Nous avons les propriétés suivantes.*

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

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

Démonstration: La première propriété résulte de la proposition 1.2.10 puisque nous avons :

$$\int_a^b f(x) \, dx = \int_a^b |f(x)| \, dx \geq \left| \int_a^b f(x) \, dx \right| \geq 0.$$

La seconde propriété est une conséquence de la première par linéarité. \square

Proposition 1.2.12 *Soit $f \in C([a, b], \mathbb{R})$ telle que $f \geq 0$. Alors, nous avons l'implication suivante.*

$$\int_a^b f(x) \, dx = 0 \Rightarrow f \equiv 0.$$

Démonstration: Nous raisonnons par l'absurde en supposant l'existence de $x_0 \in [a, b]$ tel que $f(x_0) > 0$. Nous remarquons alors que, par continuité de f en x_0 , il existe un segment non réduit à un point $[c, d]$ inclus dans $[a, b]$ et tel que, pour tout $x \in [c, d]$, $f(x) \geq f(x_0)/2$.

En utilisant la relation de Chasles et la croissance de l'intégrale de Riemann, nous en déduisons :

$$\int_a^b f(x) \, dx \geq \int_c^d f(x) \, dx \geq \frac{(d-c)f(x_0)}{2} > 0,$$

d'où une contradiction. \square

Remarque. La proposition précédente ne se généralise pas au cas d'une application f continue par morceaux. Il suffit de considérer $f = \mathbf{1}_{\{a\}}$ pour le constater.

Proposition 1.2.13 (Inégalité de Cauchy-Schwarz) *Soient $f, g \in \mathcal{R}([a, b], \mathbb{R})$; alors nous avons l'inégalité suivante :*

$$\left| \int_a^b f(x)g(x) \, dx \right| \leq \left(\int_a^b f^2(x) \, dx \right)^{\frac{1}{2}} \left(\int_a^b g^2(x) \, dx \right)^{\frac{1}{2}}.$$

Si nous supposons en outre f et g continues sur $[a, b]$, alors il y a égalité si et seulement si f et g sont linéairement liées.

Démonstration: Considérons l'application $P : \mathbb{R} \rightarrow \mathbb{R}$ définie par

$$\forall \lambda \in \mathbb{R} \quad P(\lambda) = \int_a^b (\lambda f + g)^2(x) \, dx.$$

D'après la positivité de l'intégrale de Riemann établie dans la proposition 1.2.11, nous avons $P(\lambda) \geq 0$ pour tout $\lambda \in \mathbb{R}$. En outre, nous constatons que P est un polynôme de degré au plus 2 puisque

$$\forall \lambda \in \mathbb{R} \quad P(\lambda) = \lambda^2 \int_a^b f^2(x) \, dx + 2\lambda \int_a^b f(x)g(x) \, dx + \int_a^b g^2(x) \, dx.$$

Si $\int_a^b f^2(x) \, dx = 0$, l'inégalité $P(\lambda) \geq 0, \forall \lambda \in \mathbb{R}$ implique $\int_a^b f(x)g(x) \, dx = 0$ (faire tendre λ vers $-\infty$ ou $+\infty$) et donc l'inégalité de Cauchy-Schwarz est immédiatement vérifiée.

Sinon, le polynôme P est de degré 2 et ne prend jamais de valeur strictement négative donc son discriminant est négatif ou nul, ce qui nous donne l'inégalité voulue.

Passons maintenant au cas d'égalité en supposant que

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

Si $f = 0$, il y a évidemment une liaison linéaire entre f et g . Sinon, d'après la proposition 1.2.12, nous avons $\int_a^b f^2(x) \, dx > 0$ et P est un polynôme de degré 2 dont le discriminant est nul. Il admet alors une unique racine double que nous notons λ_0 . Nous avons donc

$$\int_a^b (\lambda_0 f + g)^2(x) \, dx = 0$$

et la proposition 1.2.12 nous permet d'en déduire, puisque $(\lambda_0 f + g)^2 \in C([a, b], \mathbb{R})$, que $\lambda_0 f + g = 0$.

Il est facile de vérifier réciproquement que s'il existe $(\alpha, \beta) \in \mathbb{R}^2 - \{(0, 0)\}$ tel que $\alpha f + \beta g = 0$, alors nous nous trouvons dans le cas d'égalité de Cauchy-Schwarz. \square

Remarques.

1. On démontre par une preuve très similaire l'inégalité de Cauchy-Schwarz dans le cadre beaucoup plus général des espaces préhilbertiens réels. Le lecteur pourra trouver cette preuve dans [DAN 9-10].
2. Dans le cas où $f, g \in \mathcal{R}([a, b], \mathbb{C})$, l'inégalité de Cauchy-Schwarz s'écrit :

$$\left| \int_a^b f(x) \bar{g}(x) \, dx \right| \leq \left(\int_a^b |f(x)|^2 \, dx \right)^{\frac{1}{2}} \left(\int_a^b |g(x)|^2 \, dx \right)^{\frac{1}{2}}$$

et le cas d'égalité pour des applications continues reste celui où f et g sont linéairement liées. Le lecteur pourra adapter la démonstration précédente à titre d'exercice en choisissant, pour (f, g) fixé, un couple $(\rho, \theta) \in \mathbb{R}_+ \times \mathbb{R}$ tel que $\int_a^b f(x) \bar{g}(x) \, dx = \rho e^{i\theta}$ puis en considérant :

$$P(\lambda) := \int_a^b |f(x) + \lambda e^{i\theta} g(x)|^2 \, dx.$$

Exercices Référence :EMPSI pages 81-83

1. Soit $f \in C([a, b], \mathbb{R})$ telle qu'il existe $x_1 \in [a, b]$ tel que $f(x_1) > 0$, et $\int_a^b f = 0$. Montrer qu'il existe $x_2 \in [a, b]$ tel que $f(x_2) < 0$.
2. Déterminer les limites respectives des suites :

$$a_n = \int_0^1 \frac{x^n}{1+x} dx; \quad b_n = \int_0^\pi \frac{\sin x}{x+n} dx; \quad c_n = \int_0^\pi \frac{n \sin x}{x+n} dx; \quad d_n = \int_0^1 \sqrt{1+x^n} dx.$$

3. Soient $f, g \in C([0, 1], \mathbb{R})$ telles que $f \geq 0$, $g \geq 0$, $fg \geq 1$. Prouver l'inégalité

$$\int_0^1 f(x) \, dx \int_0^1 g(x) \, dx \geq 1.$$

4. Soit $f \in C([0, 1], \mathbb{R})$ telle que

$$\int_0^1 f^2 = \int_0^1 f^3 = \int_0^1 f^4.$$

Montrer que $f \equiv 0$ ou $f \equiv 1$.

Nous terminons ce paragraphe en énonçant deux résultats classiques (mais hors programme) faisant intervenir la notion de moyenne qui a été définie dans la remarque suivant le corollaire 1.2.8.

Proposition 1.2.14 (Première formule de la moyenne) Soient $f \in C^0([a, b], \mathbb{R})$ et $g \in \mathcal{R}([a, b], \mathbb{R})$ telle que $g \geq 0$. Alors :

$$\exists c \in [a, b] \quad \int_a^b f(x)g(x) \, dx = f(c) \int_a^b g(x) \, dx$$

Démonstration: Considérons l'application $h : [a, b] \rightarrow \mathbb{R}$ définie par :

$$\forall x \in [a, b] \quad h(x) = f(x) \int_a^b g(x) \, dx$$

L'application f étant continue sur le compact, connexe $[a, b]$ et l'intégrale de g sur $[a, b]$ étant positive, nous avons :

$$h([a, b]) = \left[\inf_{[a, b]} f \int_a^b g(x) \, dx, \sup_{[a, b]} f \int_a^b g(x) \, dx \right]$$

Puisque $g \geq 0$, nous avons les inégalités :

$$\forall x \in [a, b] \quad \inf_{[a, b]} f g(x) \leq f(x)g(x) \leq \sup_{[a, b]} f g(x)$$

Par intégration sur l'intervalle $[a, b]$, nous en déduisons :

$$\inf_{[a, b]} f \int_a^b g(x) \, dx \leq \int_a^b f(x)g(x) \, dx \leq \sup_{[a, b]} f \int_a^b g(x) \, dx$$

Autrement dit, le réel $\int_a^b f(x)g(x) \, dx$ appartient à $h([a, b])$, d'où l'existence de $c \in [a, b]$ tel que

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

□

Proposition 1.2.15 (Seconde formule de la moyenne) Soient $f : [a, b] \rightarrow \mathbb{R}$, supposée positive et décroissante et $g \in \mathcal{R}([a, b], \mathbb{R})$. Alors,

$$\exists c \in [a, b] \quad \int_a^b f(x)g(x) \, dx = f(a+) \int_a^c g(x) \, dx$$

Démonstration: Voir [DAN 201-202] qui prouve ce résultat dans le cas où f et g sont continues par morceaux. Dans le cas général, remarquons que $\int_a^b f(x)g(x) dx$ est bien définie à cause de la remarque faite à la suite de la proposition 1.1.9 : f est réglée car décroissante et fg est réglée en tant que produit d'applications réglées. Le reste de la démonstration s'adapte sans difficulté.

□

1.3 Outils pratiques de calcul d'une intégrale

Tous les outils que nous présenterons dans ce paragraphe reposent en fait sur le lien entre intégrale et primitive, que nous allons donc traiter en premier. Pour cette partie très classique, le lecteur pourra se reporter de préférence à [DAN 190-195] ou encore à [CMPSI 230-234].

Dans tout ce paragraphe, I désigne un intervalle réel d'intérieur non vide (*i.e.* I ni vide ni réduit à un point). Rappelons d'abord le résultat suivant.

Proposition 1.3.1 (Inégalité des accroissements finis) Soit E un espace vectoriel normé et soit $f : [a, b] \rightarrow E$ une application continue sur $[a, b]$ et dérivable sur $]a, b[$. S'il existe $M \in \mathbb{R}_+$ tel que $\|f'(x)\| \leq M$, pour tout $x \in]a, b[$, alors nous avons l'inégalité :

$$\|f(b) - f(a)\| \leq M(b - a).$$

1.3.1 Utilisation d'une primitive

Définition 1.3.2 Soient f et F deux applications de I dans E . On dit que F est une primitive de f sur I si F est dérivable sur I et telle que $F' = f$.

Proposition 1.3.3 Soient F une primitive de f sur I et $G : I \rightarrow E$. Alors, G est une primitive de f sur I si, et seulement si,

$$\exists C \in E, \quad \forall x \in I, \quad G(x) = F(x) + C.$$

Démonstration: Il est immédiat de vérifier qu'une application G de la forme précédente est bien une primitive de f sur I .

Réciproquement, si nous supposons que G est une primitive de f sur I , nous constatons que l'application $G - F$ est dérivable sur I , de dérivée nulle.

L'inégalité des accroissements finis entraîne alors (en majorant la norme de la dérivée par la constante nulle) que l'application $G - F$ est constante sur l'intervalle I , d'où la conclusion. □

Remarque. Dans la définition et les deux propositions précédentes, nous avons simplement supposé que E est un espace vectoriel normé quelconque. Désormais, nous aurons besoin de supposer que E est un espace de Banach afin de pouvoir introduire des intégrales d'applications à valeurs dans E .

Théorème 1.3.4 (fondamental de l'analyse) Soit $f : I \rightarrow E$ une application continue et soit $a \in I$. Alors, l'application $F : I \rightarrow E$ définie par

$$\forall x \in I, \quad F(x) = \int_a^x f(t) \, dt, \quad (1.3)$$

est une primitive de f sur I .

Démonstration: Soit $x_0 \in I$; nous allons prouver que F est dérivable en x_0 et que $F'(x_0) = f(x_0)$.

Soit $\epsilon > 0$ arbitraire. Puisque l'application f est continue au point x_0 , nous pouvons choisir $\alpha > 0$ tel que

$$\forall t \in I, \quad |t - x_0| \leq \alpha \implies \|f(t) - f(x_0)\| \leq \epsilon.$$

Pour tout réel h non nul tel que $x_0 + h \in I$ et $|h| \leq \alpha$, nous avons, en utilisant la relation de Chasles,

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

L'inégalité triangulaire dans une intégrale nous donne alors :

$$\left\| \frac{1}{h} (F(x_0 + h) - F(x_0)) - f(x_0) \right\| \leq \frac{1}{|h|} \left| \int_{x_0}^{x_0+h} \epsilon \, dt \right| \leq \epsilon.$$

Nous en déduisons que F est dérivable en x_0 et que $F'(x_0) = f(x_0)$. □

L'importance de l'hypothèse de continuité de l'application f est manifeste dans cette preuve. Nous conservons néanmoins un résultat intéressant en affaiblissant un peu cette hypothèse.

Définition 1.3.5 Une application $f : I \rightarrow E$ est dite continue par morceaux (resp. C^k par morceaux) si elle l'est sur tout segment $[a, b] \subseteq I$.

Proposition 1.3.6 Soit $f : I \rightarrow E$ une application continue par morceaux et soit $a \in I$. Alors, l'application $F : I \rightarrow E$ définie par

$$\forall x \in I, \quad F(x) = \int_a^x f(t) \, dt$$

est continue, de classe C^1 par morceaux, et vérifie $F'(x) = f(x)$ en tout point $x \in I$ où f est continue.

Démonstration: La continuité étant une propriété locale, il suffit de montrer que l'application F est continue sur tout segment $[b, c] \subseteq I$ pour en déduire que $F \in C^0(I, E)$.

Or, sur un tel segment, l'application f est continue par morceaux, donc bornée, disons par M . Si nous considérons $b \leq x \leq y \leq c$, la relation de Chasles et l'inégalité triangulaire impliquent :

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

Ainsi, l'application F est M -lipschitzienne sur $[b, c]$, donc en particulier continue sur ce segment.

Montrons maintenant que l'application F est de classe C^1 par morceaux sur tout segment $[b, c] \subseteq I$, ce qui suffira. Puisque l'application f est continue par morceaux sur le segment $[b, c]$, il existe une subdivision $\sigma = (a_i)_{0 \leq i \leq n}$ de $[b, c]$ et des applications f_0, \dots, f_{n-1} telles que :

$$\forall i \in \llbracket 0; n-1 \rrbracket, \quad f_i \in C^0([a_i, a_{i+1}], E) \quad \text{et} \quad \forall x \in]a_i, a_{i+1}[, \quad f(x) = f_i(x).$$

Soit $x_0 \in [b, c]$ fixé. Si x_0 est un point de continuité de f , la démonstration du théorème précédent s'applique et prouve que F est dérivable au point x_0 , de dérivée $F'(x_0) = f(x_0)$. Sinon, x_0 est un élément de la subdivision σ . Dans un premier temps, supposons que $x_0 = a_i$, avec $i \in \llbracket 0; n-1 \rrbracket$, et montrons que l'application F est dérivable à droite au point a_i , avec :

$$F'_d(a_i) = f_i(a_i) = \lim_{\substack{x \rightarrow a_i \\ x > a_i}} f(x).$$

Soit $\epsilon > 0$ arbitraire. Puisque l'application f_i est continue au point a_i , nous pouvons choisir $\alpha \in]0, a_{i+1} - a_i[$ tel que $\forall t \in [a_i, a_i + \alpha]$, $|f_i(t) - f_i(a_i)| \leq \epsilon$. Pour tout $h \in]0, \alpha[$, nous avons, en utilisant la relation de Chasles,

$$\begin{aligned} \left\| \frac{1}{h} (F(a_i + h) - F(a_i)) - f_i(a_i) \right\| &= \left\| \frac{1}{h} \int_{a_i}^{a_i+h} f(t) \, dt - f_i(a_i) \right\| \\ &= \left\| \frac{1}{h} \int_{a_i}^{a_i+h} (f_i(t) - f_i(a_i)) \, dt \right\|. \end{aligned}$$

L'inégalité triangulaire dans une intégrale nous donne alors :

$$\left\| \frac{1}{h} (F(a_i + h) - F(a_i)) - f_i(a_i) \right\| \leq \frac{1}{h} \int_{a_i}^{a_i+h} \epsilon \, dt \leq \epsilon,$$

ce qui nous donne le résultat voulu.

Un raisonnement analogue montre que, pour tout $i \in \llbracket 1; n \rrbracket$, l'application F est dérivable à gauche au point a_i , avec :

$$F'_g(a_i) = f_{i-1}(a_i) = \lim_{\substack{x \rightarrow a_i \\ x < a_i}} f(x).$$

Finalement, nous avons prouvé que, pour tout $i \in \llbracket 0; n-1 \rrbracket$, F est dérivable sur le segment $[a_i, a_{i+1}]$, de dérivée f_i , ce qui nous permet de conclure. \square

Le théorème fondamental de l'analyse admet le corollaire immédiat suivant.

Corollaire 1.3.7 *Toute application continue $f : I \rightarrow E$ admet au moins une primitive sur I .*

En choisissant $f : \mathbb{R}_+^* \rightarrow \mathbb{R}$ définie par $f(x) = 1/x$, nous pouvons définir le *logarithme népérien* comme l'unique primitive de f s'annulant au point 1.

Grâce au théorème fondamental de l'analyse, nous obtenons une méthode qui est à la base du calcul pratique des intégrales : l'utilisation d'une primitive. C'est de cette méthode de calcul que découlent l'*intégration par parties* et le *changement de variable* que nous énoncerons ci-dessous.

Proposition 1.3.8 *Soit $f : I \rightarrow E$ une application continue, $[a, b]$ un segment inclus dans I et F une primitive de f sur $[a, b]$. Nous avons alors l'égalité*

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

Démonstration: Nous définissons l'application $G : I \rightarrow E$ par

$$\forall x \in I, \quad G(x) = \int_a^x f(t) \, dt.$$

En appliquant la Proposition 1.3.3, nous constatons que l'application $G - F$ est constante sur $[a, b]$. En particulier, $G(b) - F(b) = G(a) - F(a)$, ce qui s'écrit encore $G(b) - G(a) = F(b) - F(a)$, d'où la conclusion. \square

Remarque. Nous utiliserons la notation suivante :

$$[F(x)]_a^b = F(b) - F(a).$$

1.3.2 Intégration par parties

Proposition 1.3.9 Soient $f, g \in C^1([a, b], \mathbb{K})$. Alors on a l'égalité :

$$\int_a^b f(x)g'(x) \, dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x) \, dx.$$

Démonstration: Nous utilisons l'égalité $(fg)' = f'g + fg'$ et, puisque toutes les applications en jeu sont continues sur $[a, b]$, nous l'intégrons entre a et b :

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

D'après la proposition 1.3.8, le membre de gauche est égal à $[f(x)g(x)]_a^b$, d'où la conclusion. □

Remarque. On peut facilement généraliser la proposition précédente au cas où $f \in C^1([a, b], \mathbb{R})$ et $g \in C^1([a, b], E)$. Voir [DAN 192].

Exemple. Une application classique de l'intégration par parties est le calcul des *intégrales de Wallis* ([CMPSI 234] ou [DAN 206-208] plus détaillé) :

$$\forall n \in \mathbb{N}, \quad I_n := \int_0^{\pi/2} \sin^n x \, dx.$$

On peut en déduire la *formule de Wallis* [CMP 261] qui nous donne l'équivalent suivant lorsque n tend vers l'infini :

$$I_n \sim \sqrt{\frac{\pi}{2n}}$$

Cet équivalent est utile notamment pour déterminer la constante intervenant dans la *formule de Stirling* : $n! \sim \sqrt{2\pi} n^{n+1/2} e^{-n}$.

Exemple. [EMP 221-223] Pour toute application $\varphi \in C^1([a, b], \mathbb{C})$, on établit facilement la convergence suivante par intégration par parties.

$$\int_a^b \varphi(x) \sin(nx) \, dx \xrightarrow{n \rightarrow +\infty} 0.$$

Si l'on définit l'application $\varphi : [0, \pi/2] \rightarrow \mathbb{R}$ par $\varphi(0) = 0$ et

$$\forall x \in]0, \pi/2], \quad \varphi(x) = \frac{1}{x} - \frac{1}{\sin x},$$

on vérifie que $\varphi(0+) = 0$ et $\varphi'(0+) = -1/6$ grâce à des développements limités ; ceci entraîne que $\varphi \in C^1([0, \pi/2], \mathbb{R})$ et nous déduisons donc de ce qui précède, par extraction de sous-suite,

$$\int_0^{\pi/2} \left(\frac{\sin((2n+1)x)}{x} - \frac{\sin((2n+1)x)}{\sin x} \right) dx \xrightarrow{n \rightarrow +\infty} 0.$$

Or, on établit par des calculs élémentaires que pour tout $x \in [0, \pi/2]$,

$$\frac{\sin((2n+1)x)}{\sin x} = 1 + 2 \sum_{k=1}^n \cos(2kx), \text{ d'où } \int_0^{\pi/2} \frac{\sin((2n+1)x)}{\sin x} dx = \frac{\pi}{2}.$$

Nous en déduisons la convergence

$$\int_0^{\pi/2} \frac{\sin((2n+1)x)}{x} dx \xrightarrow{n \rightarrow +\infty} \frac{\pi}{2}.$$

En anticipant sur la suite, avec le changement de variable $t = (2n+1)x$, nous en déduisons la valeur de *l'intégrale de Dirichlet* :

$$\int_0^{+\infty} \frac{\sin t}{t} dt = \frac{\pi}{2}.$$

On établit la proposition suivante par récurrence en utilisant une intégration par parties.

Proposition 1.3.10 (formule de Taylor avec reste intégral) Soient I un intervalle non réduit à un point, $n \in \mathbb{N}$ et soit $f \in C^{n+1}(I, \mathbb{K})$. Pour tout $(a, b) \in I^2$ avec $a \neq b$, nous avons l'égalité :

$$f(b) = \sum_{k=0}^n f^{(k)}(a) \frac{(b-a)^k}{k!} + \int_a^b f^{(n+1)}(t) \frac{(b-t)^n}{n!} dt.$$

Exercices

1. (EMPSI 83) Soit $f \in C^2([0, 2\pi], \mathbb{R})$ une application convexe. Prouver l'inégalité

$$\int_0^{2\pi} f(x) \cos x \, dx \geq 0.$$

2. (CMPSI 235) Prouver la convergence suivante lorsque $x \rightarrow +\infty$:

$$\frac{\int_1^x e^t \ln t \, dt}{e^x \ln x} \longrightarrow 1.$$

1.3.3 Changement de variable

Proposition 1.3.11 Soient $f \in C^0(I, E)$ et $u \in C^1([a, b], I)$. Alors, on a l'égalité :

$$\int_a^b f(u(t)) u'(t) \, dt = \int_{u(a)}^{u(b)} f(x) \, dx.$$

Démonstration: Considérons l'application $G : I \rightarrow E$ définie par

$$\forall x \in I, \quad G(x) = \int_{u(a)}^x f(t) \, dt,$$

et l'application $F : [a, b] \rightarrow E$ définie par $F = G \circ u$.

D'après le théorème 1.3.4, l'application G est une primitive de f sur I ; en particulier, $G \in C^1(I, E)$. Comme en outre $u \in C^1([a, b], I)$, nous en déduisons par composition que $F \in C^1([a, b], E)$ et

$$\forall t \in [a, b], \quad F'(t) = f(u(t))u'(t).$$

Ainsi, F est une primitive de l'application continue $t \mapsto f(u(t))u'(t)$ sur $[a, b]$ et nous concluons par la proposition 1.3.8. \square

Remarque. On dit couramment que l'on a posé le changement de variable $x = u(t)$ et donc que (formellement) $dx = u'(t) \, dt$.

Proposition 1.3.12 Soient $f \in C_M([\alpha, \beta], E)$ et $u \in C^1([a, b], [\alpha, \beta])$ supposée strictement monotone. Alors, on a l'égalité :

$$\int_a^b f(u(t)) u'(t) \, dt = \int_{u(a)}^{u(b)} f(x) \, dx.$$

Démonstration: Voir [DAN 194] qui montre aussi qu'en l'absence de l'hypothèse u strictement monotone, on peut avoir $f \circ u \notin C_M([a, b], E)$, si bien que l'intégrale du membre de gauche n'est pas définie. \square

Voici une application du changement de variable dans une intégrale au cas d'une fonction T -périodique, où $T > 0$.

Proposition 1.3.13 *Soit $f \in C(\mathbb{R}, E)$ une application T -périodique. Alors,*

$$\forall (t_1, t_2) \in \mathbb{R}^2, \quad \int_{t_1}^{t_1+T} f(x) \, dx = \int_{t_2}^{t_2+T} f(x) \, dx.$$

Démonstration: D'après la relation de Chasles,

$$\int_{t_1}^{t_1+T} f(x) \, dx = \int_{t_1}^{t_2} f(x) \, dx + \int_{t_2}^{t_2+T} f(x) \, dx + \int_{t_2+T}^{t_1+T} f(x) \, dx.$$

Il nous suffit donc pour conclure de prouver l'égalité :

$$\int_{t_1}^{t_2} f(x) \, dx = \int_{t_1+T}^{t_2+T} f(x) \, dx,$$

qui s'écrit encore, en vertu de la T -périodicité de l'application f ,

$$\int_{t_1}^{t_2} f(x+T) \, dx = \int_{t_1+T}^{t_2+T} f(y) \, dy$$

et cette dernière égalité résulte du changement de variable $y = x + T$. \square

Remarque. En vertu de la remarque précédente, cette proposition se généralise au cas où f est continue par morceaux. En effet, le changement de variable que nous avons utilisé dans la démonstration est strictement croissant puisqu'il s'agit simplement d'une translation.

1.3.4 Exercices suggérés.

Référence : EMPSI pages 81-83

1. Soit $f \in C([0, 1], \mathbb{R})$. Prouver l'inégalité

$$\left| \int_0^1 (f(x) + xf(1-x)) \, dx \right| \leq \frac{3}{2} \|f\|_\infty.$$

2. Calculer la valeur de l'intégrale

$$\int_0^{2\pi} \sqrt{\frac{1 + \cos x}{2}} dx.$$

3. Calculer la valeur de l'intégrale

$$\int_0^{\pi/4} \ln(1 + \tan x) dx.$$

4. En utilisant une comparaison avec une intégrale, prouver la majoration

$$\forall n \in \mathbb{N}^* \quad \sum_{k=1}^n \frac{1}{k} \leq 1 + \ln n$$

et en déduire :

$$\forall n \in \mathbb{N}^* \quad \sum_{d|n} d \leq n(1 + \ln n).$$

5. Déterminer les limites suivantes :

$$\lim_{u \rightarrow 0^+} \int_0^{\pi/2} e^{-u \sin x} dx \quad ; \quad \lim_{u \rightarrow 0^+} \int_u^{3u} \frac{\cos x}{x} dx.$$

6. Déterminer l'ensemble des applications $f \in C(\mathbb{R}_+, \mathbb{R})$ telles que $f \geq 0$ et

$$\forall x \in \mathbb{R}_+ \quad f(x) \leq \int_0^x f(t) dt.$$

Indication : Introduire $g(x) = \exp(-x) \int_0^x f(t) dt$.

7. Etudier et représenter graphiquement l'application $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par l'intégrale dépendant d'un paramètre aux bornes suivante :

$$\forall x \in \mathbb{R} \quad f(x) = \int_x^{2x} e^{-t^2} dt.$$

Indication : Utiliser les valeurs numériques qui vous sont données ci-dessous :

$$\alpha = \sqrt{\frac{\ln 2}{3}} \approx 0,481 \quad ; \quad f(\alpha) \approx 0,286.$$

8. Soit $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ une application k -lipschitzienne ($k \geq 0$). On définit l'application $F : \mathbb{R}_+ \rightarrow \mathbb{R}$ par $F(0) = f(0)$ et

$$\forall x \in \mathbb{R}_+^* \quad F(x) = \frac{1}{x} \int_0^x f(t) dt.$$

Montrer que l'application F est $k/2$ -lipschitzienne.

Indication : Effectuer un changement de variable pour fixer les bornes de l'intégrale.

9. Soit $f \in C([0, 1], \mathbb{R})$ telle que

$$\forall (x, y) \in [0, 1]^2 \quad xf(y) + yf(x) \leq 1.$$

Prouver l'inégalité :

$$\int_0^1 f(x) dx \leq \frac{\pi}{4}.$$

Indication : Ecrire de deux façons différentes cette intégrale en utilisant les changements de variable $x = \sin u$ et $x = \cos v$.

Référence :CMP 143

1. Calculer les intégrales suivantes :

$$I = \int_0^{\pi/2} \frac{\cos x}{\sqrt{1 + \sin x \cos x}} dx \quad ; \quad J = \int_0^{\pi/2} \frac{\sin x}{\sqrt{1 + \sin x \cos x}} dx.$$

2. Soit $\theta \in [0, \pi/2[$ fixé. Calculer l'intégrale

$$\int_0^{2\theta} \frac{x}{\cos(x - \theta)} dx.$$

3. Calculer l'intégrale suivante :

$$\int_{-1}^1 \frac{1}{\sqrt{1-x} + \sqrt{1+x} + 2} dx.$$

Indication : Effectuer les changements de variables successifs $x = \cos \theta$ puis $\theta = 2(\varphi + \pi/4)$.

4. Soient $a > 0$ fixé et $f \in C([0, a], \mathbb{R}_+^*)$. Calculer l'intégrale

$$\int_0^a \frac{f(t)}{f(t) + f(a-t)} dt.$$

En déduire la valeur de l'intégrale suivante :

$$\int_0^{\pi/2} \frac{(\cos t)^{\sin t}}{(\cos t)^{\sin t} + (\sin t)^{\cos t}} dt.$$

1.3.5 Applications suggérées.

Pour un exposé systématique sur le calcul de primitives, le lecteur pourra par exemple consulter le chapitre 9 de [CMPSI 309].

Le lemme de (Riemann-)Lebesgue, qui concerne les séries de Fourier, est traité dans le cas C^1 par intégration par parties dans [CMPSI 234-235] et dans le cas plus général d'une application continue par morceaux dans [DAN 209-211].

Les applications de \mathbb{R} dans \mathbb{R} qui sont additives et continues sont exactement les applications linéaires : une démonstration à l'aide d'intégrales de ce résultat classique est donnée dans [EMPSI 83].

Le lecteur trouvera dans [CMPSI 232] la démonstration du lemme classique suivant en utilisant le lien entre primitive et intégrale :

Lemme 1.3.14 (Gronwall) Soient $T \in \bar{\mathbb{R}}_+^*$ et $f, g : [0, T[\rightarrow \mathbb{R}$ continues, $f \geq 0$, $g \geq 0$, telles que

$$\exists C > 0 \quad \forall x \in [0, T[\quad f(x) \leq C + \int_0^x fg.$$

Alors on a l'inégalité suivante :

$$\forall x \in [0, T[\quad f(x) \leq C \exp \left(\int_0^x g \right).$$

Le lemme de Gronwall est très utile en théorie des équations différentielles. Prouvons par exemple que l'équation

$$\frac{dy}{dt} = b(y) \quad ; \quad y(0) = a$$

avec $a \in \mathbb{R}$ fixé et $b : \mathbb{R} \rightarrow \mathbb{R}$ supposée k -lipschitzienne, admet au plus une solution.

Supposons donc que x et y soient deux solutions de cette équation différentielle si bien que, pour $T > 0$ arbitraire,

$$\forall t \in [0, T[\quad x_t = a + \int_0^t b(x_s) ds \text{ et } y_t = a + \int_0^t b(y_s) ds.$$

Nous en déduisons, en utilisant l'inégalité de Cauchy-Schwarz, que pour tout $t \in [0, T[$

$$(x_t - y_t)^2 = \left(\int_0^t [b(x_s) - b(y_s)] ds \right)^2 \leq t \int_0^t [b(x_s) - b(y_s)]^2 ds.$$

Puisque l'application b est supposée k -lipschitzienne, nous en déduisons :

$$\forall t \in [0, T[\quad (x_t - y_t)^2 \leq Tk^2 \int_0^t (x_s - y_s)^2 ds.$$

En appliquant le lemme de Gronwall avec $g \equiv Tk^2$ et $C > 0$ arbitraire, nous en déduisons :

$$\forall t \in [0, T[\quad (x_t - y_t)^2 \leq C \exp(Tk^2 t).$$

Nous pouvons alors conclure en faisant tendre C vers 0 puis en utilisant $T > 0$ arbitraire.

1.4 Sommes de Riemann

La notion de *sommes de Riemann* est introduite dans [CMPSI 222-226] et illustrée par plusieurs exercices.

Elle ne figure pas au programme officiel mais sera néanmoins bien utile dans plusieurs leçons pour fournir des illustrations ou des exercices.

Le résultat essentiel, que nous allons énoncer immédiatement après une première définition, est assez intuitif (faire un dessin).

Définition 1.4.1 On appelle *subdivision pointée d'un segment* $[a, b]$ un couple (σ, Θ) où $\sigma = (a_i)_{0 \leq i \leq n}$ est une subdivision de $[a, b]$ et $\Theta = (\theta_i)_{1 \leq i \leq n}$ est une famille de points de $[a, b]$ telle que :

$$\forall 1 \leq i \leq n \quad \theta_i \in [a_{i-1}, a_i].$$

Définition et proposition 1.4.2 Soit $f \in \mathcal{R}([a, b], E)$ et (σ, Θ) une subdivision pointée de $[a, b]$. On appelle *somme de Riemann associée à l'application f et à la subdivision pointée (σ, Θ)* la quantité :

$$S(f, \sigma, \Theta) = \sum_{i=1}^n (a_i - a_{i-1}) f(\theta_i).$$

Considérons maintenant une suite $(\sigma_N, \Theta_N)_{N \in \mathbb{N}^*}$ de subdivisions pointées de $[a, b]$ telle que la suite des pas $(\|\sigma_N\|)_{n \in \mathbb{N}^*}$ tende vers 0. Alors nous avons la convergence suivante lorsque N tend vers l'infini :

$$S(f, \sigma_N, \Theta_N) \longrightarrow \int_a^b f(x) dx.$$

Démonstration: Le lecteur trouvera la preuve dans [DAN 196-197] qui se restreint au cas f continue par morceaux mais la démonstration est exactement la même pour f réglée.

L'idée est d'examiner d'abord le cas où f est de la forme $c \mathbf{1}_{[\alpha, \beta]}$, $c \in E$, puis de passer au cas d'une application en escalier par linéarité pour terminer la démonstration par densité de $\mathcal{E}([a, b], E)$ dans $(\mathcal{R}([a, b], E), \|\cdot\|_\infty)$.

□

Remarque. Dans le cas où $f \in C([a, b], E)$, l'énoncé précédent résulte de la construction de l'intégrale de Riemann. En effet, en constatant d'une part qu'une somme de Riemann peut s'écrire comme intégrale d'une fonction en escalier bien choisie et d'autre part que le théorème de Heine nous donne l'uniforme continuité de f , nous pouvons écrire le résultat précédent sous la forme

$$\int_a^b f = \lim_{N \rightarrow +\infty} \int_a^b f_N,$$

où (f_N) est une suite de fonctions en escalier qui converge uniformément vers f sur $[a, b]$.

Un cas particulier souvent utilisé dans la pratique est obtenu en prenant des subdivisions régulières de $[a, b]$. Plus précisément, pour tout $N \in \mathbb{N}^*$, nous définissons la subdivision σ_N comme la famille $(a_i)_{0 \leq i \leq N}$ telle que :

$$\forall 0 \leq i \leq N, \quad a_i = a + i \frac{b-a}{N}.$$

En outre, nous définissons la subdivision pointée (σ_N, Θ_N) en choisissant simplement $\theta_i = a_{i-1}$ pour tout $1 \leq i \leq N$.

La proposition précédente nous donne alors :

Corollaire 1.4.3 *Pour toute application $f \in C_M([a, b])$, nous avons la convergence suivante lorsque N tend vers l'infini :*

$$\frac{b-a}{N} \sum_{i=0}^{N-1} f\left(a + i \frac{b-a}{N}\right) \longrightarrow \int_a^b f(x) \, dx.$$

Ce résultat nous fournit une méthode pour calculer une approximation d'une intégrale dite *méthode des rectangles*.

Dans le cas où f est suffisamment régulière, on peut utiliser l'inégalité de Taylor-Lagrange pour mesurer la qualité de cette approximation, ce qui est essentiel dans la pratique. Nous écrivons la proposition suivante dans le cas $a = 0$ et $b = 1$, ce qui ne fait pas perdre de généralité puisqu'on peut toujours se ramener à ce cas par un simple changement de variable (transformation affine).

Proposition 1.4.4 *Supposons $f \in C^3([0, 1], \mathbb{R})$. Alors on a le développement asymptotique :*

$$\frac{1}{N} \sum_{i=0}^{N-1} f\left(\frac{i}{N}\right) = \int_0^1 f(x) \, dx - \frac{f(1) - f(0)}{2N} + \frac{f'(1) - f'(0)}{12N^2} + \mathcal{O}\left(\frac{1}{N^3}\right).$$

Démonstration: Voir [DAN 198-199].

□

Nous reviendrons beaucoup plus précisément sur les différentes méthodes de calcul de la valeur approchée d'une intégrale dans un chapitre ultérieur.

Exemple. (Intégrale de Poisson) Pour tout $a \in \mathbb{R} \setminus \{-1; 1\}$, on définit :

$$I(a) = \int_0^\pi \ln(1 - 2a \cos x + a^2) dx.$$

D'après le Corollaire 1.4.3, nous avons donc :

$$I(a) = \lim_{n \rightarrow +\infty} \frac{\pi}{n} \sum_{k=1}^{n-1} \ln(a^2 - 2a \cos \frac{k\pi}{n} + 1) = \lim_{n \rightarrow +\infty} \frac{\pi}{n} \ln P_n(a), \quad (1.4)$$

où le polynôme P_n est donné par :

$$P_n(X) = \prod_{k=1}^{n-1} \left(X - \exp\left(\frac{2ik\pi}{2n}\right) \right) \left(X - \exp\left(\frac{-2ik\pi}{2n}\right) \right) = \frac{X^{2n} - 1}{(X - 1)(X + 1)}.$$

Revenant à la formule (1.4), nous en déduisons :

$$|a| < 1 \Rightarrow I(a) = 0 \quad ; \quad |a| > 1 \Rightarrow I(a) = 2\pi \ln |a|.$$

Exercices. Référence : EMPSI pages 82-83

1. Montrer que la suite de terme général

$$u_n = \sum_{k=1}^n \frac{1}{\sqrt{n^2 + 2kn}}$$

est convergente et calculer sa limite.

2. Montrer que la suite de terme général

$$u_n = \prod_{k=1}^n \left(1 + \frac{k^2}{n^2} \right)^{1/n}$$

est convergente et calculer sa limite.

3. Montrer que la suite de terme général

$$u_n = \sum_{k=0}^n (n+k)^{-\alpha} (n+k+1)^{-\beta},$$

où $\alpha > 0, \beta > 0$ et $\alpha + \beta = 1$, est convergente et calculer sa limite.

Indication : Introduire $v_n = \sum_{k=0}^n (n+k)^{-1}$ et prouver que

$$v_n - \frac{1}{n} + \frac{1}{2n+1} \leq u_n \leq v_n.$$

4. Montrer que la suite de terme général

$$u_n = \sum_{k=1}^n \sin \frac{k}{n^2} \sin \frac{k}{n}$$

est convergente et calculer sa limite.

Indication : Introduire $v_n = \sum_{k=1}^n \frac{k}{n^2} \sin \frac{k}{n}$ et montrer l'inégalité

$$\forall x \in \mathbb{R}_+ \quad x - \frac{x^3}{6} \leq \sin x \leq x.$$

5. Soit $f \in C([0, 1], \mathbb{R})$ et $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ convexe. Montrer l'inégalité de Jensen :

$$\varphi \left(\int_0^1 f \right) \leq \int_0^1 \varphi(f).$$

1.5 Convergences de suites d'applications

Les définitions des convergences simple et uniforme ont été données page 13. Rappelons simplement ici que la continuité en un point x_0 (ou sur l'intervalle $[a, b]$) est conservée par passage à une limite uniforme [DAN 271], ce qui n'est pas le cas pour une limite simple.

On peut en déduire que $(C([a, b], E), \|\cdot\|_\infty)$ est un espace de Banach en tant que sous-espace fermé de $(\mathcal{B}([a, b], E), \|\cdot\|_\infty)$ qui est lui-même un espace de Banach.

Proposition 1.5.1 *Soit $(f_n)_{n \in \mathbb{N}}$ une suite d'applications réglées de $[a, b]$ dans E qui converge uniformément sur $[a, b]$ vers une application f .*

Alors $f \in \mathcal{R}([a, b], E)$ et nous avons l'égalité :

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

Démonstration: Dans l'espace vectoriel normé $(\mathcal{B}([a, b], E), \|\cdot\|_\infty)$, nous avons $f = \lim f_n$ avec $(f_n) \in (\mathcal{R}([a, b], E))^{\mathbb{N}}$. Or $\mathcal{R}([a, b], E)$ est un sous-espace fermé de $(\mathcal{B}([a, b], E), \|\cdot\|_\infty)$ puisque, par définition,

$$\mathcal{R}([a, b], E) = \overline{\mathcal{E}([a, b], E)}^{\|\cdot\|_\infty}$$

Nous en déduisons que $f \in \mathcal{R}([a, b], E)$ et donc que l'intégrale du membre de gauche est bien définie.

Par construction, l'intégrale de Riemann est une application linéaire continue $\hat{I} : (\mathcal{R}([a, b], E), \|\cdot\|_\infty) \rightarrow E$ de norme $(b - a)$, comme nous l'avons établi dans la proposition 1.2.7. Nous déduisons de cette propriété de continuité que $\hat{I}(f) = \lim \hat{I}(f_n)$, d'où la conclusion.

Une autre façon d'aboutir à cette conclusion est d'écrire les inégalités :

$$\left\| \int_a^b f(x) dx - \int_a^b f_n(x) dx \right\| \leq \int_a^b \|f(x) - f_n(x)\| dx \leq (b - a) \|f - f_n\|_\infty.$$

□

Remarque. Cette interversion entre limite et intégrale n'est pas valable lorsqu'il y a seulement convergence simple, même si toutes les applications en jeu sont continues. Un contre-exemple est obtenu en prenant $a = 0$, $b = 1$ et en définissant comme suit f_n , $n \in \mathbb{N}^*$: $f_n(x) = n^2 x$ si $0 \leq x \leq 1/n$, $f_n(x) = 2n - n^2 x$ si $1/n \leq x \leq 2/n$ et $f_n(x) = 0$ sinon. □

Dans toute la suite de ce paragraphe, nous allons considérer des applications numériques définies sur $[a, b]$ et nous allons définir de nouveaux types de convergences pour des suites de telles applications.

Proposition et définition 1.5.2 *L'application $\|\cdot\|_1$ définie sur $C([a, b], \mathbb{R})$ par*

$$\forall f \in C([a, b], \mathbb{R}) \quad \|f\|_1 = \int_a^b |f(x)| dx$$

est une norme sur $C([a, b], \mathbb{R})$ appelée norme 1.

Soient $(f_n)_{n \in \mathbb{N}}$ et f des applications numériques continues sur $[a, b]$. Si la suite (f_n) converge vers f dans $(C([a, b], \mathbb{R}), \|\cdot\|_1)$, on dit que (f_n) converge en moyenne vers f .

Démonstration: La seule vérification qui ne soit pas immédiate est celle de l'implication suivante : $\|f\|_1 = 0 \Rightarrow f = 0$. Elle résulte de la proposition 1.2.12. □

Remarque. L'énoncé précédent se généralise sans difficulté à $C([a, b], E)$, avec E espace de Banach quelconque. En revanche, le produit scalaire que nous allons introduire maintenant pourrait être généralisé à $C([a, b], \mathbb{C})$ mais pas au-delà.

Proposition et définition 1.5.3 *L'application $\langle \cdot, \cdot \rangle$ définie sur $(C([a, b], \mathbb{R}))^2$ par*

$$\forall (f, g) \in (C([a, b], \mathbb{R}))^2 \quad \langle f, g \rangle = \int_a^b f(x)g(x) \, dx$$

est un produit scalaire sur $C([a, b], \mathbb{R})$. La norme induite par ce produit scalaire est appelée norme 2 et donc définie par :

$$\forall f \in C([a, b], \mathbb{R}) \quad \|f\|_2 = \left(\int_a^b f^2(x) \, dx \right)^{\frac{1}{2}}.$$

Soient $(f_n)_{n \in \mathbb{N}}$ et f des applications numériques continues sur $[a, b]$. Si la suite (f_n) converge vers f dans $(C([a, b], \mathbb{R}), \|\cdot\|_2)$, on dit que (f_n) converge en moyenne quadratique vers f .

Démonstration: Le lecteur vérifiera sans difficulté que l'application $\langle \cdot, \cdot \rangle$ est bilinéaire symétrique positive. Le fait qu'elle soit définie résulte de la proposition 1.2.12. \square

Remarque. L'inégalité de Cauchy-Schwarz, énoncée dans la proposition 1.2.13, se réécrit donc dans notre cadre

$$\forall (f, g) \in (C([a, b], \mathbb{R}))^2 \quad |\langle f, g \rangle| \leq \|f\|_2 \|g\|_2.$$

Sous cette forme, elle est d'ailleurs vraie dans le cadre beaucoup plus général des espaces préhilbertiens, ainsi que son cas d'égalité. Le lecteur pourra se reporter à [DAN 9-10].

C'est précisément l'inégalité de Cauchy-Schwarz que nous allons utiliser pour comparer les trois normes que nous avons précédemment définies sur $C([a, b], \mathbb{R})$.

Proposition 1.5.4 *Pour toute application $f \in C([a, b], \mathbb{R})$, nous avons les inégalités*

$$\|f\|_1 \leq \sqrt{b-a} \|f\|_2 \leq (b-a) \|f\|_\infty.$$

Démonstration: Grâce à l'inégalité de Cauchy-Schwarz, nous pouvons écrire (en notant $\mathbf{1}$ l'application constante et égale à 1 sur $[a, b]$)

$$\|f\|_1 = \int_a^b |f(x)| \, dx = \langle |f|, \mathbf{1} \rangle \leq \|f\|_2 \|\mathbf{1}\|_2 = \sqrt{b-a} \|f\|_2,$$

d'où la première inégalité. Nous obtenons alors la seconde inégalité en écrivant

$$\|f\|_2 = \left(\int_a^b f^2(x) \, dx \right)^{\frac{1}{2}} \leq \left(\int_a^b \|f\|_\infty^2 \, dx \right)^{\frac{1}{2}} = \sqrt{b-a} \|f\|_\infty.$$

□

Corollaire 1.5.5 Soient $(f_n)_{n \in \mathbb{N}}$ et f des applications numériques continues sur $[a, b]$.

Si (f_n) converge uniformément vers f sur $[a, b]$, alors (f_n) converge en moyenne quadratique vers f sur $[a, b]$.

Si (f_n) converge en moyenne quadratique vers f sur $[a, b]$, alors (f_n) converge en moyenne vers f sur $[a, b]$

Dans les deux cas, la réciproque est fausse comme nous allons le voir maintenant.

Contre-exemples :

1. Prenons $a = 0$, $b = 1$, $f \equiv 0$ et définissons f_n , $n \in \mathbb{N}^*$ par :
 $f_n(x) = n^4 x$ si $0 \leq x \leq 1/n^3$, $f_n(x) = 2n - n^4 x$ si $1/n^3 \leq x \leq 2/n^3$ et $f_n(x) = 0$ sinon.
 Alors la suite (f_n) converge en moyenne quadratique vers f mais ne converge pas uniformément vers f .
2. Prenons $a = 0$, $b = 1$, $f \equiv 0$ et définissons f_n , $n \in \mathbb{N}^*$ par :
 $f_n(x) = n^3 x$ si $0 \leq x \leq 1/n^2$, $f_n(x) = 2n - n^3 x$ si $1/n^2 \leq x \leq 2/n^2$ et $f_n(x) = 0$ sinon.
 Alors la suite (f_n) converge en moyenne vers f mais ne converge pas en moyenne quadratique vers f : on pourra remarquer que $f_n(x) \geq n/2$ dès que $1/(2n^2) \leq x \leq 3/(2n^2)$.

Remarque et exercice suggéré : On peut montrer que pour tout $p \in \mathbb{N}^*$, l'application $\|\cdot\|_p$ définie sur $C([a, b], \mathbb{R})$ par

$$\forall f \in C([a, b], \mathbb{R}) \quad \|f\|_p = \left(\int_a^b |f(x)|^p \, dx \right)^{\frac{1}{p}}$$

est une norme sur $C([a, b], \mathbb{R})$: on l'appelle norme p . Ce résultat qui repose en fait sur une inégalité de convexité n'est pas trivial. Le lecteur intéressé par la démonstration pourra se reporter aux inégalités dites de Hölder et Minkowski dans tout livre de théorie de la mesure et de l'intégration.

La notation que nous avons utilisée pour la norme de la convergence uniforme est alors justifiée par le résultat suivant :

$$\forall f \in C([a, b], \mathbb{R}) \quad \lim_{p \rightarrow +\infty} \|f\|_p = \|f\|_\infty.$$

Le lecteur trouvera ce résultat sous la forme d'un exercice corrigé dans [DAN 205-206].

1.6 Intégrale d'une fonction dépendant d'un paramètre

Nous considérons une application $f : E \times [a, b] \rightarrow F$, où (E, d) est un espace métrique et F un espace de Banach.

Rappelons ici que $E \times [a, b]$ est alors lui-même un espace métrique en tant que produit de deux espaces métriques.

Théorème 1.6.1 (Continuité) *Si $f : E \times [a, b] \rightarrow F$ est continue, alors nous pouvons définir l'application $\Phi : E \rightarrow F$ par :*

$$\forall x \in E, \quad \Phi(x) = \int_a^b f(x, t) dt,$$

et cette application Φ est continue.

Démonstration: Pour $x \in E$ arbitrairement fixé, l'application $t \mapsto f(x, t)$ est continue du segment $[a, b]$ dans l'espace de Banach F , ce qui permet de définir l'intégrale ci-dessus.

Nous allons prouver la continuité de Φ par caractérisation séquentielle : si $x \in E$ et si $(x_n)_{n \in \mathbb{N}^*} \in E^{\mathbb{N}^*}$ est une suite qui converge vers x , alors $(\Phi(x_n))_{n \in \mathbb{N}^*}$ converge vers $\Phi(x)$ dans F .

Sous ces hypothèses, il est bien connu que $\{x_n, n \in \mathbb{N}^*\} \cup \{x\}$ est un sous-ensemble compact de E (cf. [SKA 23]). Ainsi, il suffit de démontrer la continuité de la restriction de Φ à un sous-ensemble compact quelconque K de l'espace métrique E .

Puisque $K \times [a, b]$, muni de la distance $d_1((x, s), (y, t)) := d(x, y) + |s - t|$, est alors un compact, le théorème de Heine implique que la restriction de

1.6. INTÉGRALE D'UNE FONCTION DÉPENDANT D'UN PARAMÈTRE 45

l'application f à $K \times [a, b]$ est uniformément continue. Soit $x_0 \in E$ arbitraire fixé. Pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que, pour tout $x \in E$ et tout $t \in [a, b]$,

$$d(x, x_0) < \delta \Rightarrow d_1((x, t), (x_0, t)) < \delta \Rightarrow |f(x, t) - f(x_0, t)| < \frac{\epsilon}{b - a}.$$

Grâce à l'inégalité triangulaire dans une intégrale, nous en déduisons :

$$d(x, x_0) < \delta \Rightarrow |\Phi(x) - \Phi(x_0)| \leq \int_a^b |f(x, t) - f(x_0, t)| dt < \epsilon,$$

ce qui nous permet de conclure. □

Exemple. [CMP 190-191] Soit $T > 0$ et soient $f, g : \mathbb{R} \rightarrow \mathbb{C}$ deux applications continues et T -périodiques. Le *produit de convolution* de f et g est l'application $f * g : \mathbb{R} \rightarrow \mathbb{C}$ définie par :

$$\forall x \in \mathbb{R}, \quad f * g(x) = \int_0^T f(t)g(x - t) dt.$$

Grâce au théorème précédent, on vérifie sans peine que l'application $f * g$ est continue et T -périodique.

Nous supposons désormais que $E = I$, intervalle réel.

Soit $(x_0, t_0) \in I \times [a, b]$. Rappelons que la dérivée partielle $f'_x(x_0, t_0)$ désigne, sous réserve de son existence, la dérivée au point $x_0 \in I$ de l'application $x \mapsto f(x, t_0)$ de I dans F .

À l'aide de l'inégalité des accroissements finis, on peut établir le résultat suivant (cf. Ramis-Deschamps-Odoux tome 3, paragraphe 8.1.8, dont on peut adapter la preuve en se référant à la démonstration précédente).

Théorème 1.6.2 (Dérivabilité) *Si $f : I \times [a, b] \rightarrow F$ est continue, telle que la dérivée partielle f'_x existe et soit continue sur $I \times [a, b]$, alors l'application $\Phi : I \rightarrow F$ définie par :*

$$\forall x \in I, \quad \Phi(x) = \int_a^b f(x, t) dt$$

est de classe C^1 et nous avons l'égalité :

$$\forall x \in I, \quad \Phi'(x) = \int_a^b f'_x(x, t) dt.$$

Ces résultats seront considérablement généralisés dans la section 3.3 grâce au *théorème de convergence dominée*.

1.7 Arcs paramétrés

1.7.1 Rectification d'un arc paramétré.

Le lecteur pourra se reporter à [DAN 203-205] à propos de la rectification des arcs paramétrés, que nous présentons très succinctement ici.

Définition 1.7.1 Soit $d \in \mathbb{N}^*$ fixé et $f \in C([a, b], \mathbb{R}^d)$. On dit que le couple $\Gamma = ([a, b], f)$ est un arc paramétré continu de \mathbb{R}^d . Cet arc est dit rectifiable si

$$\sup \left\{ \sum_{i=0}^{n-1} \|f(a_{i+1}) - f(a_i)\| \right\} < +\infty,$$

cette borne supérieure portant sur toutes les subdivisions $(a_i)_{0 \leq i \leq n}$ ($n \in \mathbb{N}^*$) de l'intervalle $[a, b]$ et $\|\cdot\|$ désignant la norme euclidienne dans \mathbb{R}^d .

Dans ce cas, cette borne supérieure est appelée longueur de l'arc paramétré continu Γ et notée $L(\Gamma)$.

Un exemple d'arc continu non rectifiable nous est donné par l'application $f : [0, 1] \rightarrow \mathbb{R}^2$ définie par

$$f(0) = (0, 0) \text{ et } \forall t \in]0, 1] \quad f(t) = (t, t \sin \frac{1}{t}).$$

Le lecteur pourra démontrer que la borne supérieure ci-dessus est infinie en choisissant judicieusement des subdivisions et en utilisant la divergence de la série harmonique (i.e. $\sum_n \frac{1}{n} = +\infty$).

Proposition 1.7.2 Soit $f \in C^1([a, b], \mathbb{R}^d)$ et $\Gamma = ([a, b], f)$ l'arc paramétré associé. Alors cet arc est rectifiable, de longueur

$$L(\Gamma) = \int_a^b \|f'(x)\| dx.$$

Exercice suggéré : Référence : “Topologie, Analyse” par G. FLORY, exercices avec solutions, tome 3, éditions Vuibert.

Calculer la longueur totale de l'astroïde définie sous forme paramétrée par :

$$f(t) = (a \cos^3 t, a \sin^3 t) \quad , \quad 0 \leq t \leq 2\pi.$$

1.7.2 Intégrale curviligne d'une forme différentielle.

Une autre application aux arcs paramétrés concerne la notion d'intégrale curviligne d'une forme différentielle le long d'une courbe orientée. Cette notion permet en particulier de calculer le travail d'une force le long d'une courbe en Physique. Référence : [CMPSI 396-403].

Chapitre 2

Intégrales impropres.

Dans tout ce chapitre, nous considérons des applications à valeurs dans le corps $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

Nous pourrions en fait définir les intégrales impropres dans le cadre plus général des applications à valeurs dans un espace de Banach E mais dans la pratique ce sont surtout des intégrales impropres réelles ou complexes qui apparaissent. Le lecteur intéressé par les énoncés généraux dans E pourra les trouver dans les anciens livres de classes préparatoires (Ramis-Deschamps-Odoux par exemple).

2.1 Définition des intégrales impropres.

Définition 2.1.1 *Soit I un intervalle réel et soit $f : I \rightarrow \mathbb{K}$. Nous dirons que l'application f est continue par morceaux sur I si la restriction de f à tout segment inclus dans I est continue par morceaux. Dans ce cas, nous noterons $f \in C_M(I, \mathbb{K})$.*

Exemple. Soit $I = [a, b[$, avec $-\infty < a < b \leq +\infty$. Alors $f \in C_M([a, b[, \mathbb{K})$ si et seulement si, pour tout $x \in [a, b[$, la restriction de f à l'intervalle compact $[a, x]$ est continue par morceaux.

Remarque. Si nous définissons une application en escalier sur $[a, b[$ par l'existence d'une subdivision $\sigma : a = a_0 < \dots < a_n = b$, telle que

$$\exists (c_i)_{0 \leq i \leq n-1} \in \mathbb{K}^n, \quad \forall i = 0, \dots, n-1, \quad f|_{[a_i, a_{i+1}[} \equiv c_i,$$

remarquons qu'une application continue par morceaux sur $[a, b[$ n'est pas forcément limite uniforme sur $[a, b[$ de fonctions en escalier car cela impliquerait

qu'elle soit bornée. L'application $f : [0, 1[\rightarrow \mathbb{R}$ définie par $f(x) = (1 - x)^{-1}$ fournit alors un contre-exemple.

Définition 2.1.2 Soit $f \in C_M([a, b[, \mathbb{K})$ avec $-\infty < a < b \leq +\infty$. Nous définissons l'application $F : [a, b[\rightarrow \mathbb{K}$ par :

$$\forall x \in [a, b[, \quad F(x) = \int_a^x f(t) dt.$$

Si l'application F admet une limite $l \in \mathbb{K}$ lorsque x tend vers b , nous dirons que l'intégrale impropre $\int_a^b f(t) dt$ est convergente et que sa valeur est :

$$\int_a^b f(t) dt = l.$$

Dans le cas contraire (F n'admet pas de limite ou bien tend vers l'infini), nous dirons que l'intégrale impropre est divergente.

Remarque. Si $c \in [a, b[$, la relation de Chasles implique immédiatement que les intégrales impropres $\int_a^b f(t) dt$ et $\int_c^b f(t) dt$ ont même nature, convergentes ou divergentes. Dans le cas où elles convergent, nous avons l'égalité :

$$\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$$

La notion d'intégrale impropre généralise l'intégrale de Riemann dans deux directions : intervalle non borné (cas $b = +\infty$) ou bien application non bornée (par exemple, $f(x)$ tend vers $+\infty$ lorsque $x \rightarrow b-$). Nous allons rencontrer ces deux situations dans l'exemple très classique suivant.

Exemple de Riemann L'intégrale impropre $\int_1^{+\infty} \frac{dt}{t^\alpha}$ converge si et seulement si $\alpha > 1$.

L'intégrale impropre $\int_a^b \frac{dt}{(b-t)^\alpha}$ converge si et seulement si $\alpha < 1$.

Malgré la simplicité de la preuve (prendre une primitive), cet exemple est important car il sert souvent de référence lorsqu'on emploie des relations de comparaison pour étudier la convergence d'une intégrale impropre (nous y reviendrons ci-dessous).

Même si, pour fixer les idées, nous avons travaillé jusqu'à présent avec des intégrales impropres sur un intervalle de la forme $[a, b[$, il est évident que tout ce qui précède s'adapte immédiatement pour définir des intégrales impropres sur un intervalle de la forme $]a, b]$. Ainsi, l'exemple de Riemann se réécrit comme suit sur $]a, b]$:

Exemple. L'intégrale impropre $\int_a^b \frac{dt}{(t-a)^\alpha}$ converge si et seulement si $\alpha < 1$.

Nous pouvons alors combiner les deux cas précédents pour définir comme suit une intégrale impropre sur un intervalle ouvert.

Proposition et définition 2.1.3 Soit $I =]a, b[$ avec $-\infty \leq a < b \leq +\infty$ et $f \in C_M(]a, b[, \mathbb{K})$. Supposons qu'il existe $c_0 \in]a, b[$ tel que les intégrales $\int_a^{c_0} f(t) dt$ et $\int_{c_0}^b f(t) dt$ soient convergentes. Alors, pour tout $c \in]a, b[$, les intégrales $\int_a^c f(t) dt$ et $\int_c^b f(t) dt$ convergent et leur somme est indépendante du choix de c .

On dit que l'intégrale impropre $\int_a^b f(t) dt$ est convergente et que sa valeur est :

$$\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$$

Dans le cas contraire, on dit que l'intégrale impropre $\int_a^b f(t) dt$ est divergente.

Nous pourrions même considérer le cas d'une application $f : [a, b] \rightarrow \mathbb{K}$, avec $-\infty < a < b < +\infty$, qui n'est pas bornée au voisinage d'un certain nombre de points d'un ensemble fini $F \subset [a, b]$. On découpe alors le segment $[a, b]$ en segments dont l'intérieur ne contient aucun point de F et l'on étudie la convergence de l'intégrale de f , à gauche et à droite, aux points de F comme ci-dessus. Notons néanmoins que, dans ce cas, l'application f n'est pas forcément continue par morceaux sur $[a, b]$.

Dans la suite de ce cours, nous allons mettre en évidence diverses méthodes d'étude de la convergence d'une intégrale impropre ; nous écrirons les propositions dans le cas d'une intégrale impropre définie sur un intervalle de la forme $I = [a, b[$, laissant au lecteur le soin de les adapter aux autres cas.

2.2 Étude de la convergence : cas positif

Avant d'examiner plusieurs critères de convergence ou divergence, nous commençons par un lemme, propre au cas positif.

Soit $f \in C_M([a, b[, \mathbb{R}_+)$; nous définissons l'application $F : [a, b[\rightarrow \mathbb{R}$ par :

$$\forall x \in [a, b[, \quad F(x) = \int_a^x f(t) dt. \quad (2.1)$$

Alors, l'hypothèse $f \geq 0$ implique la croissance de F , comme le montre la relation de Chasles. Nous en déduisons le résultat suivant.

Lemme 2.2.1 Soit $f \in C_M([a, b[, \mathbb{R}_+)$ et soit l'application $F : [a, b[\rightarrow \mathbb{R}$ définie par la formule (2.1). Alors, nous avons l'égalité :

$$\lim_{x \rightarrow b-} F(x) = \sup_{[a, b[} F \leq +\infty.$$

Preuve : Si $\sup_{[a, b[} F = +\infty$, alors $\forall A > 0$, $\exists c \in [a, b[, F(c) > A$.

Grâce à la croissance de l'application F , nous pouvons donc écrire :

$$\forall A > 0, \quad \exists c \in [a, b[, \quad \forall x \in [c, b[, \quad F(x) > A,$$

si bien que $\lim_{x \rightarrow b-} F(x) = +\infty$.

Si $\sup_{[a, b[} F = M \in \mathbb{R}_+$ alors, la borne supérieure étant le plus petit des majorants, $\forall \epsilon > 0$, $\exists c \in [a, b[, F(c) > M - \epsilon$.

En utilisant la croissance de F , nous obtenons donc :

$$\forall \epsilon > 0, \quad \exists c \in [a, b[, \quad \forall x \in [c, b[, \quad M - \epsilon < F(x) \leq M,$$

d'où $\lim_{x \rightarrow b-} F(x) = M$. □

Il résulte de ce lemme que, si $\sup_{[a, b[} F < +\infty$, l'intégrale impropre $\int_a^b f(t)$ converge et l'on a :

$$\int_a^b f(t) dt = \sup_{[a, b[} F.$$

Dans le cas contraire, l'intégrale impropre $\int_a^b f(t)$ est divergente et nous adopterons la convention suivante.

Convention 2.2.2 Soit $f \in C_M([a, b[, \mathbb{R}_+)$ et soit l'application $F : [a, b[\rightarrow \mathbb{R}$ définie par la formule (2.1). Si $\sup_{[a, b[} F = +\infty$, alors nous noterons :

$$\int_a^b f(t) dt = +\infty.$$

Insistons sur le fait que cette convention est propre au cas où l'intégrande f est une application positive.

Nous allons maintenant décrire plusieurs méthodes permettant de déterminer la nature d'une intégrale impropre (convergente ou divergente) dans ce cas.

2.2.1 Comparaison avec une intégrale de référence.

Proposition 2.2.3 Soient $f, g \in C_M([a, b[, \mathbb{R})$ telles que $0 \leq f \leq g$. Alors, on a les implications suivantes :

1. Si $\int_a^b g(t) dt$ converge, alors $\int_a^b f(t) dt$ converge.
2. Si $\int_a^b f(t) dt$ diverge, alors $\int_a^b g(t) dt$ diverge.

Démonstration: Définissons l'application $F : [a, b[\rightarrow \mathbb{R}$ comme précédemment et l'application G de façon similaire. Puisque $0 \leq f \leq g$, nous avons

$$\forall x \in [a, b[, \quad F(x) \leq G(x).$$

1. Les hypothèses impliquent que

$$\sup_{[a, b[} F \leq \sup_{[a, b[} G = \int_a^b g(t) dt < +\infty,$$

d'où la convergence de $F(x)$ vers $\sup_{[a, b[} F$ lorsque x tend vers b par valeurs inférieures.

2. C'est la contraposée de la première implication.

□

En utilisant l'exemple de Riemann, nous en déduisons les critères suivants, souvent utiles dans la pratique.

Corollaire 2.2.4 Soient $-\infty < a < b < +\infty$ et soit $f \in C_M([a, b[, \mathbb{R}_+)$.

1. S'il existe $\alpha < 1$ tel que $\lim_{x \rightarrow b} (b - x)^\alpha f(x) = 0$, alors l'intégrale impropre $\int_a^b f(x) dx$ est convergente.
2. S'il existe $\alpha \geq 1$ tel que $\lim_{x \rightarrow b} (b - x)^\alpha f(x) = +\infty$, alors l'intégrale impropre $\int_a^b f(x) dx$ est divergente.

Démonstration: Nous traitons le premier cas, laissant le second à titre d'exercice au lecteur.

L'hypothèse implique l'existence de $c \in [a, b[$ tel que :

$$\forall x \in [c, b[, \quad (b - x)^\alpha f(x) \leq 1.$$

Nous concluons en appliquant la proposition précédente aux intégrales

$$\int_c^b f(t) dt \quad \text{et} \quad \int_c^b \frac{dt}{(b - t)^\alpha},$$

cette dernière intégrale étant convergente d'après l'exemple de Riemann.

□

Corollaire 2.2.5 Soient $a \in \mathbb{R}$ et $f \in C_M([a, +\infty[, \mathbb{R}_+)$.

1. S'il existe $\alpha > 1$ tel que $\lim_{x \rightarrow +\infty} x^\alpha f(x) = 0$, alors l'intégrale impropre $\int_a^{+\infty} f(x) dx$ est convergente.
2. S'il existe $\alpha \leq 1$ tel que $\lim_{x \rightarrow +\infty} x^\alpha f(x) = +\infty$, alors l'intégrale impropre $\int_a^{+\infty} f(x) dx$ est divergente.

Ce corollaire implique immédiatement la convergence de l'intégrale de Gauss :

$$\int_{-\infty}^{+\infty} e^{-t^2} dt.$$

2.2.2 Intégration des relations de comparaison

Dans toute cette sous-section, nous considérons deux applications continues par morceaux $f, g : [a, b[\rightarrow \mathbb{R}_+$, avec $-\infty < a < b \leq +\infty$.

Le lecteur pourra réviser les notions de comparaison locale ou asymptotique de fonctions (domination, prépondérance et équivalence) dans [DAN 119ss].

Proposition 2.2.6 (Intégration de la domination) *Supposons qu'au voisinage de b , l'application f est dominée par g , ce que nous notons $f = O_b(g)$. Alors, nous avons les implications suivantes :*

1. Si $\int_a^b g(t) dt$ converge, alors $\int_a^b f(t) dt$ converge et

$$\int_x^b f(t) dt = O_b \left(\int_x^b g(t) dt \right).$$

2. Si $\int_a^b f(t) dt$ diverge, alors $\int_a^b g(t) dt$ diverge et

$$\int_a^x f(t) dt = O_b \left(\int_a^x g(t) dt \right).$$

Remarque. Lorsque $x \rightarrow b$, les intégrales qui apparaissent dans les relations de domination établies par ce théorème ont pour limite 0 dans le premier cas (on parle de *reste intégral* d'une intégrale impropre convergente) et pour limite $+\infty$ dans le second cas.

Démonstration: Voir [DAN 220]. Nous détaillons plutôt la preuve de la proposition suivante qui est légèrement plus difficile.

Proposition 2.2.7 (Intégration de la prépondérance) *Supposons qu'au voisinage de b , l'application f est négligeable devant g (on dit encore que g est prépondérante devant f), ce que nous notons $f = o_b(g)$. Alors, nous avons les implications suivantes :*

1. Si $\int_a^b g(t) dt$ converge, alors $\int_a^b f(t) dt$ converge et

$$\int_x^b f(t) dt = o_b \left(\int_x^b g(t) dt \right).$$

2. Si $\int_a^b f(t) dt$ diverge, alors $\int_a^b g(t) dt$ diverge et

$$\int_a^x f(t) dt = o_b \left(\int_a^x g(t) dt \right).$$

Démonstration: Soit $\epsilon > 0$ arbitraire ; notre hypothèse nous donne :

$$\exists c \in [a, b[, \quad \forall t \in [c, b[, \quad 0 \leq f(t) \leq \epsilon g(t). \quad (2.2)$$

1. Puisque $\int_a^b g(t) dt$ converge, il en est de même de $\int_c^b g(t) dt$. L'inégalité précédente implique alors la convergence de $\int_c^b f(t) dt$, d'où celle de $\int_a^b f(t) dt$. En outre, si $c \leq x \leq y < b$, nous avons :

$$\int_x^y f(t) dt \leq \epsilon \int_x^y g(t) dt.$$

En passant à la limite lorsque $y \rightarrow b-$, nous obtenons :

$$\forall x \in [c, b[, \quad \int_x^b f(t) dt \leq \epsilon \int_x^b g(t) dt,$$

d'où la relation de prépondérance annoncée lorsque $x \rightarrow b-$ entre les *restes intégraux*.

2. Puisque $\int_a^b f(t) dt$ diverge, il en est de même de $\int_c^b f(t) dt$. L'inégalité (2.2) implique alors la divergence de $\int_c^b g(t) dt$, d'où celle de $\int_a^b g(t) dt$. Cette divergence (vers $+\infty$) implique l'existence de $c' \in [c, b[$ tel que :

$$\forall x \in [c', b[, \quad \int_a^x g(t) dt \geq \frac{1}{\epsilon} \int_a^c f(t) dt.$$

Nous en déduisons, pour tout $x \in [c', b[$,

$$\int_a^x f(t) dt = \int_a^c f(t) dt + \int_c^x f(t) dt \leq \epsilon \int_a^c g(t) dt + \epsilon \int_c^x g(t) dt$$

En majorant cette dernière intégrale par $\int_a^x g(t) dt$, nous en déduisons :

$$\forall x \in [c', b[, \quad \int_a^x f(t) dt \leq 2\epsilon \int_a^x g(t) dt,$$

ce qui nous permet de conclure puisque ϵ était arbitraire.

□

Proposition 2.2.8 (Intégration de l'équivalence) *Nous supposons qu'au voisinage de b , les applications f et g sont équivalentes, ce que nous notons $f \sim_b g$. Alors, nous avons les implications suivantes :*

1. Les intégrales $\int_a^b f(t) dt$ et $\int_a^b g(t) dt$ sont de même nature.

2. (a) Si ces intégrales convergent, alors :

$$\int_x^b f(t) dt \underset{b}{\sim} \int_x^b g(t) dt.$$

(b) Si ces intégrales divergent, alors :

$$\int_a^x f(t) dt \underset{b}{\sim} \int_a^x g(t) dt.$$

Démonstration:

1. Puisque $f = O(g)$ et $g = O(f)$, la proposition 2.2.6 nous donne le résultat.
2. (a) Puisque $f - g = o_b(g)$, ce qui équivaut à $|f - g| = o_b(g)$, l'inégalité triangulaire et la proposition 2.2.7 impliquent :

$$\left| \int_x^b f(t) dt - \int_x^b g(t) dt \right| \leq \int_x^b |f(t) - g(t)| dt = o_b\left(\int_x^b g(t) dt\right),$$

ce qui nous permet de conclure.

(b) On raisonne d'une façon analogue en écrivant :

$$\left| \int_a^x f(t) dt - \int_a^x g(t) dt \right| \leq \int_a^x |f(t) - g(t)| dt = o_b\left(\int_a^x g(t) dt\right).$$

Exemple. On définit $f \in C^1(\mathbb{R}_+, \mathbb{R})$ en posant :

$$\forall t \in \mathbb{R}_+, \quad f(t) = \frac{e^{-t^2}}{2t}.$$

On montre facilement que $-f'(t) \sim \exp(-t^2)$ lorsque $t \rightarrow +\infty$ et l'on en déduit, grâce au corollaire précédent,

$$\int_x^{+\infty} e^{-t^2} dt \sim \frac{e^{-x^2}}{2x} \quad \text{lorsque } x \rightarrow +\infty.$$

En calcul des probabilités, on évalue ainsi la *queue de distribution* gaussienne, c'est-à-dire le comportement asymptotique de $x \mapsto P(Z > x)$, où Z est une variable gaussienne centrée réduite.

Exercices suggérés L'exercice 14.4 de [DAN 222-223], l'exercice 3.1.1 de [CMP 164], les exercices 3.3.1 à 3.3.3 de [CMP 184], l'exercice 3.1 de [EMP 195] permettent de mettre en oeuvre les méthodes précédentes sur des exemples assez simples.

Remarque. Les résultats précédents tombent en défaut lorsque nous considérons des fonctions de signe quelconque.

Ainsi, il est facile de constater que nous avons l'équivalence suivante :

$$\frac{\sin t}{\sqrt{t}} + \frac{|\sin t|}{t} \sim \frac{\sin t}{\sqrt{t}} \text{ lorsque } t \rightarrow +\infty.$$

En effet, nous avons bien sûr :

$$\frac{|\sin t|}{t} = \frac{|\sin t|}{\sqrt{t}} \times \frac{1}{\sqrt{t}} = o\left(\frac{\sin t}{\sqrt{t}}\right) \text{ lorsque } t \rightarrow +\infty.$$

Néanmoins, à l'aide de méthodes que nous verrons ci-dessous, nous pourrions établir que :

$$\int_1^{+\infty} \left(\frac{\sin t}{\sqrt{t}} + \frac{|\sin t|}{t} \right) dt \text{ diverge et } \int_1^{+\infty} \frac{\sin t}{\sqrt{t}} dt \text{ converge.}$$

2.2.3 Comparaison avec une série

En utilisant la fonction F définie par l'égalité (2.1), le lecteur montrera facilement la proposition suivante.

Proposition 2.2.9 *Soit $a \in \mathbb{R}$ et soit $f \in C_M([a, +\infty[, \mathbb{R}_+)$.*

Si $\int_a^{+\infty} f$ converge, alors pour toute suite $(x_n) \in [a, +\infty[^{\mathbb{N}}$ de limite $+\infty$, la série $\sum \int_{x_n}^{x_{n+1}} f$ converge.

S'il existe une suite croissante $(x_n) \in [a, +\infty[^{\mathbb{N}}$ de limite $+\infty$ telle que la série $\sum \int_{x_n}^{x_{n+1}} f$ converge, alors l'intégrale impropre $\int_a^{+\infty} f$ converge.

Exemple. La fonction $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ définie par $f(t) = \frac{|\sin t|}{t}$ si $t \neq 0$ et $f(0) = 1$, est continue sur \mathbb{R}_+ . Nous allons prouver la divergence suivante.

$$\int_0^{+\infty} f := \int_0^{+\infty} \frac{|\sin t|}{t} dt = +\infty.$$

En contraposant la première phrase de la proposition ci-dessus, nous constatons que l'existence d'une suite $(x_n) \in [0, +\infty[^{\mathbb{N}}$ de limite $+\infty$ telle que la série $\sum \int_{x_n}^{x_{n+1}} f$ diverge implique la divergence de l'intégrale impropre $\int_0^{+\infty} f$. Prenons $x_n = n\pi$ et étudions donc la suite $(u_n)_{n \in \mathbb{N}}$ définie par :

$$\forall n \in \mathbb{N}, \quad u_n = \int_{n\pi}^{(n+1)\pi} \frac{|\sin t|}{t} dt.$$

Par décroissance de l'application $t \rightarrow \frac{1}{t}$ sur \mathbb{R}_+^* , nous avons la minoration :

$$\forall n \in \mathbb{N}, \quad u_n \geq \frac{1}{(n+1)\pi} \int_{n\pi}^{(n+1)\pi} |\sin t| dt,$$

qui s'écrit encore, en calculant l'intégrale par π -périodicité et à l'aide d'une primitive :

$$\forall n \in \mathbb{N}, \quad u_n \geq \frac{2}{(n+1)\pi}.$$

Nous en déduisons la divergence de la série $\sum u_n$, ce qui nous permet de conclure.

Remarque. Contrairement à ce qui se passe pour les séries positives, la convergence d'une intégrale impropre positive sur $[a, +\infty[$ n'implique pas que son intégrande tende vers 0 : on trouvera un contre-exemple avec un intégrande continu et non borné au voisinage de l'infini dans [DAN 234]. En revanche, si l'intégrande est uniformément continu, il est nécessaire qu'il tende vers 0 en $+\infty$ pour que l'intégrale impropre converge : voir par exemple « Oraux X-ENS, Analyse 3 » par Francinou, Gianella et Nicolas page 183.

2.3 Cas général : critère de Cauchy et autres méthodes.

Dans toute cette section, sauf indication explicite, nous considérons une application $f \in C_M([a, b[, \mathbb{K})$ avec $-\infty < a < b \leq +\infty$.

2.3.1 Critère de Cauchy et conséquences.

Puisque $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} est complet, nous avons la condition nécessaire et suffisante de convergence suivante :

Proposition 2.3.1 (Critère de Cauchy pour les intégrales) *L'intégrale impropre $\int_a^b f(t) dt$ converge si et seulement si :*

$$\forall \epsilon > 0, \exists c \in [a, b[, \forall (x, y) \in [c, b]^2 \quad \left| \int_x^y f(t) dt \right| < \epsilon.$$

Démonstration: Soit $F : [a, b[\rightarrow \mathbb{R}$ définie par

$$\forall x \in [a, b[\quad F(x) = \int_a^x f(t) \, dt.$$

Il s'agit de démontrer l'équivalence entre les deux propositions suivantes :

1. L'application F admet une limite dans \mathbb{K} lorsque $x \rightarrow b$.
2. Le critère de Cauchy suivant est satisfait :

$$\forall \epsilon > 0, \exists c \in [a, b[, \forall (x, y) \in [c, b[\quad |F(y) - F(x)| < \epsilon$$

L'implication $(1 \Rightarrow 2)$ résulte facilement de l'inégalité triangulaire.

Pour prouver l'implication réciproque, on utilise une caractérisation séquentielle et le fait que si $(x_n) \in [a, b[^\mathbb{N}$ tend vers b , alors $(F(x_n))$ est une suite de Cauchy dans \mathbb{K} qui est complet.

Le lecteur trouvera les détails dans [DAN 223-224]. □

Nous allons maintenant énoncer trois conséquences de cette proposition, en commençant par un corollaire presque immédiat.

Corollaire 2.3.2 *Si $-\infty < a < b < +\infty$ et si $f \in C_M([a, b[, \mathbb{K})$ est bornée sur $[a, b[$, alors l'intégrale impropre $\int_a^b f(t) \, dt$ converge.*

Démonstration: En notant $M = \sup_{[a, b[} |f| < +\infty$ et en prenant $\epsilon > 0$ arbitraire, nous constatons que le critère précédent est satisfait en prenant

$$c = \max \left(a, b - \frac{\epsilon}{M} \right)$$

puisque nous avons les inégalités

$$\left| \int_x^y f(t) \, dt \right| \leq \int_x^y |f(t)| \, dt \leq M(y - x).$$

□

Par exemple, l'intégrale impropre $\int_0^1 \sin(1/t) \, dt$ est convergente.

Une importante conséquence du critère de Cauchy pour les intégrales concerne la notion d'absolue convergence.

Proposition et définition 2.3.3 *Soit $f \in C_M([a, b[, \mathbb{K})$. Alors nous avons l'implication :*

$$\int_a^b |f(t)| \, dt \text{ converge} \implies \int_a^b f(t) \, dt \text{ converge.}$$

On dit dans ce cas que l'intégrale impropre de f est absolument convergente sur $[a, b[$. Si l'intégrale impropre de f sur $[a, b[$ est convergente mais non absolument convergente, on dit qu'elle est semi-convergente.

Démonstration: Puisque l'intégrale impropre $\int_a^b |f(t)| dt$ est convergente, le critère de Cauchy suivant est satisfait :

$$\forall \epsilon > 0, \exists c \in [a, b[, \forall (x, y) \in [c, b]^2 \quad \int_x^y |f(t)| dt < \epsilon.$$

Il suffit alors d'appliquer l'inégalité, valable si $c \leq x \leq y < b$,

$$\left| \int_x^y f(t) dt \right| \leq \int_x^y |f(t)| dt$$

pour conclure en utilisant de nouveau le critère de Cauchy pour les intégrales. \square

Exemples. Nous montrerons ci-dessous que l'intégrale $\int_0^{+\infty} \frac{\sin t}{t} dt$ est semi-convergente.

Pour $\alpha > 1$, l'intégrale impropre $\int_1^{+\infty} \frac{\sin t}{t^\alpha} dt$ est absolument convergente. \square

En utilisant la seconde formule de la moyenne, nous pouvons énoncer une autre conséquence du critère de Cauchy pour les intégrales.

Proposition 2.3.4 (Règle d'Abel) *Considérons deux applications $f \in C_M([a, b[, \mathbb{R}_+)$ et $g \in C_M([a, b[, \mathbb{K})$. Supposons que les deux conditions suivantes sont satisfaites :*

1. *f est décroissante avec $\lim_{x \rightarrow b} f(x) = 0$,*
2. *il existe $M > 0$ tel que*

$$\forall x \in [a, b[\quad \left| \int_a^x g(t) dt \right| \leq M.$$

Alors l'intégrale impropre $\int_a^b f(t)g(t) dt$ converge.

Démonstration: Traitons dans un premier temps le cas $\mathbb{K} = \mathbb{R}$.

Puisque $f \geq 0$ et $\lim_{x \rightarrow b} f(x) = 0$, pour $\epsilon > 0$ arbitraire, nous pouvons trouver $c \in [a, b[$ tel que

$$\forall x \in [c, b[\quad 0 \leq f(x) \leq \epsilon.$$

Soit $(x, y) \in [c, b]^2$ tel que $x < y$. La seconde formule de la moyenne (Proposition 1.2.15 page 25) implique l'existence de $z \in [x, y]$ tel que

$$\int_x^y f(t)g(t) dt = f(x+) \int_x^z g(t) dt$$

Nous en déduisons

$$\left| \int_x^y f(t)g(t) dt \right| \leq 2M\epsilon,$$

ce qui permet de conclure puisque le critère de Cauchy pour les intégrales est ainsi satisfait.

Passons maintenant au cas $\mathbb{K} = \mathbb{C}$ en écrivant $g = \Re g + i\Im g$. Les hypothèses sur g impliquent alors que, pour tout $x \in [a, b[$,

$$\left| \int_a^x \Re g(t) dt \right| = \left| \Re \left(\int_a^x g(t) dt \right) \right| \leq \left| \int_a^x g(t) dt \right| \leq M$$

(on a utilisé l'inégalité, valable pour tout $z \in \mathbb{C}$, $|\Re z| \leq |z|$).

Nous constatons que les applications f et $\Re g$ satisfont les hypothèses de l'énoncé dans le cas $\mathbb{K} = \mathbb{R}$. Nous en déduisons la convergence de l'intégrale impropre $\int_a^b f(t)\Re g(t) dt$. Par un raisonnement similaire, nous obtenons la convergence de l'intégrale impropre $\int_a^b f(t)\Im g(t) dt$, d'où finalement celle de l'intégrale impropre complexe $\int_a^b f(t)g(t) dt$.

□

Exemple. Pour tout $\alpha > 0$, l'intégrale impropre suivante est convergente :

$$\int_1^{+\infty} \frac{\sin t}{t^\alpha} dt$$

Il suffit en effet d'appliquer la règle d'Abel avec $f(x) = x^{-\alpha}$ et $g(x) = \sin x$ sur $[1, +\infty[$, en constatant que

$$\forall x \in [1, +\infty[\quad \left| \int_1^x \sin t dt \right| = |\cos x - \cos 1| \leq 2.$$

On peut obtenir ce résultat sans la règle d'Abel, en procédant par intégration par parties, mais c'est un peu plus long : voir [EMP 216-217].

2.3.2 Méthodes directes

On peut étudier directement la convergence d'une intégrale impropre en appliquant les outils classiques de calcul d'une intégrale, comme l'intégration par parties ou le changement de variable. Une application classique de ce dernier est l'exemple de Bertrand, qui généralise l'exemple de Riemann.

Exemple de Bertrand Pour $(\alpha, \beta) \in \mathbb{R}^2$, nous considérons l'intégrale impropre

$$\int_2^{+\infty} \frac{1}{x^\alpha (\ln x)^\beta} dx.$$

Cette intégrale converge si et seulement si $\alpha > 1$ ou $(\alpha = 1 \text{ et } \beta > 1)$. Le lecteur pourra consulter la démonstration dans [CMP 162].

On peut avoir recours à un développement asymptotique pour appliquer ensuite l'intégration des relations de domination, prépondérance ou équivalence. Par exemple, dans son exercice 3.4.1 d), [CMP 189] établit ainsi la convergence de l'intégrale impropre

$$\int_0^{+\infty} \frac{e^{i(x+\frac{1}{x})}}{x} dx.$$

Intégrale de Dirichlet Par comparaison avec des séries, [GRA 34] prouve la semi-convergence de l'intégrale impropre $\int_0^{+\infty} \frac{\sin t}{t} dt$.

2.4 Calcul des intégrales impropres

Puisqu'une intégrale impropre est par définition la limite d'une intégrale de Riemann (fonction de la borne d'en haut par exemple), nous allons retrouver dans cette section les méthodes classiques de calcul d'une intégrale de Riemann : utilisation d'une primitive, intégration par parties, changement de variables. Ces deux dernières méthodes admettent des généralisations qui nous fournissent des formules directement applicables aux intégrales impropres comme nous allons le voir maintenant.

2.4.1 Intégration par parties généralisée

Proposition 2.4.1 *Considérons deux applications $f, g \in C^1([a, b[, \mathbb{K})$. Alors nous avons les propriétés suivantes :*

1. *Si $\lim_{x \rightarrow b} f(x)g(x)$ existe, alors les deux intégrales impropres suivantes sont de même nature :*

$$\int_a^b f(t)g'(t) dt \text{ et } \int_a^b f'(t)g(t) dt.$$

2. *Dans le cas où ces deux intégrales convergent, nous avons l'égalité*

$$\int_a^b f(t)g'(t) dt = \lim_{x \rightarrow b} f(x)g(x) - f(a)g(a) - \int_a^b f'(t)g(t) dt.$$

Démonstration: Pour tout $x \in [a, b[$, la formule d'intégration par parties classique nous donne :

$$\int_a^x f(t)g'(t) dt = f(x)g(x) - f(a)g(a) - \int_a^x f'(t)g(t) dt.$$

Si $\lim_{x \rightarrow b} f(x)g(x)$ existe, la convergence d'une des intégrales impropres équivaut donc à celle de l'autre, d'où le premier point.

Nous obtenons alors le second point par simple passage à la limite lorsque $x \rightarrow b$ dans l'égalité précédente. \square

Remarque. Dans le cas où les deux intégrales impropres convergent, nous emploierons encore la notation

$$\int_a^b f(t)g'(t) dt = [f(x)g(x)]_a^b - \int_a^b f'(t)g(t) dt.$$

Exemple. Existence et calcul de l'intégrale :

$$\int_0^{+\infty} \left(\int_x^{+\infty} \frac{\sin t}{t} dt \right) dx. \quad (2.3)$$

Posons, pour tout $x \geq 0$, $F(x) = \int_x^{+\infty} \frac{\sin t}{t} dt$ et calculons, pour tout $y \geq 0$,

$$I_y := \int_0^y F(x) dx = [xF(x)]_0^y - \int_0^y xF'(x) dx = yF(y) + \int_0^y \sin x dx,$$

d'où l'égalité :

$$\forall y \in \mathbb{R}_+, \quad I_y = yF(y) - \cos y + 1. \quad (2.4)$$

Or, une IPP généralisée nous donne, pour tout $y > 0$,

$$F(y) = \int_y^{+\infty} \frac{\sin t}{t} dt = \left[\frac{-\cos t}{t} \right]_y^{+\infty} - \int_y^{+\infty} \frac{\cos t}{t^2} dt = \frac{\cos y}{y} - \int_y^{+\infty} \frac{\cos t}{t^2} dt.$$

En reportant cette égalité dans (2.4), nous obtenons :

$$\forall y \in \mathbb{R}_+^*, \quad I_y = 1 - y \int_y^{+\infty} \frac{\cos t}{t^2} dt. \quad (2.5)$$

Une nouvelle IPP généralisée nous donne :

$$\int_y^{+\infty} \frac{\cos t}{t^2} dt = \left[\frac{\sin t}{t^2} \right]_y^{+\infty} + \int_y^{+\infty} \frac{2 \sin t}{t^3} dt = -\frac{\sin y}{y^2} + \int_y^{+\infty} \frac{2 \sin t}{t^3} dt.$$

Par inégalité triangulaire dans une intégrale, on montre facilement que ce dernier terme est un $O(1/y^2)$ lorsque $y \rightarrow +\infty$ et finalement, revenant à l'égalité (2.5), nous obtenons $\lim_{y \rightarrow +\infty} I_y = 1$, ce qui nous donne l'existence et la valeur de l'intégrale (2.3).

2.4.2 Changement de variable généralisé

Proposition 2.4.2 Soit u une application de classe C^1 bijective strictement croissante de l'intervalle $[a, b[$ sur l'intervalle $[\alpha, \beta[$ et $f \in C_M([\alpha, \beta[, \mathbb{K})$. Alors les deux intégrales impropres

$$\int_a^b f(u(t))u'(t) dt \text{ et } \int_\alpha^\beta f(x) dx$$

sont de même nature et égales en cas de convergence.

Démonstration: Voir [DAN 230-231]. Notons qu'en cas de convergence des intégrales impropres, l'égalité énoncée par la proposition s'écrit encore :

$$\int_a^b f(u(t))u'(t) dt = \int_{u(a)}^{u(b-)} f(x) dx.$$

□

Remarque. La proposition reste valable avec $u \in C^1([a, b[,]\alpha, \beta])$ strictement décroissante et $f \in C_M(] \alpha, \beta], \mathbb{K})$.

Exercices suggérés Quand il s'agit de calculer une intégrale impropre, on commence toujours par étudier la convergence de celle-ci. Le lecteur pourra procéder ainsi pour l'exercice 3.2.6 de [CMP 180] et pour les exercices 3.2, 3.5, 3.6 de [EMP 195-205]. L'exercice 1.3 de [EMP 201] utilise la formule de Wallis (qui nous fournit un équivalent de l'intégrale de Wallis) pour calculer l'intégrale impropre suivante, dite *intégrale de Gauss*, utile par exemple en calcul des probabilités pour définir la densité gaussienne.

$$\int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

2.5 Application à l'étude de séries.

Proposition 2.5.1 Soit $f \in C_M([1, +\infty[, \mathbb{R}_+)$ une application décroissante. Alors la série $\sum_{n \geq 1} f(n)$ et l'intégrale impropre $\int_1^{+\infty} f(t) dt$ sont de même nature. En outre, nous avons :

1. Si $\int_1^{+\infty} f(t) dt$ diverge, alors $S_n := \sum_{k=1}^n f(k) \sim \int_1^n f(t) dt$.

2. Si $\int_1^{+\infty} f(t) dt$ converge et si la condition suivante est satisfaite

$$\lim_{n \rightarrow +\infty} \frac{\int_{n+1}^{+\infty} f(t) dt}{\int_n^{+\infty} f(t) dt} = 1$$

alors $R_n := \sum_{k=n+1}^{+\infty} f(k) \sim \int_n^{+\infty} f(t) dt$.

Remarque. Si $\int_1^{+\infty} f(t) dt$ converge, nous savons que la suite $\left(\int_n^{+\infty} f(t) dt\right)$ converge vers 0 en décroissant. La condition énoncée dans le deuxième point exprime que cette convergence est *lente* (le lecteur pourra se reporter aux définitions générales dans [DAN 147]). Si cette condition n'est pas respectée, la conclusion tombe en défaut comme le montre le contre-exemple suivant : Si $f(t) = e^{-t}$ pour tout $t \in [1, +\infty[$, alors on trouve

$$\forall n \in \mathbb{N} \quad R_n = \frac{e^{-n}}{e-1} \text{ et } \int_n^{+\infty} f(t) dt = e^{-n}$$

donc les suites (R_n) et $\left(\int_n^{+\infty} f(t) dt\right)$ ne sont pas équivalentes.

La démonstration de la proposition précédente passe par un lemme que nous allons maintenant prouver, laissant le lecteur se reporter à [DAN 242-244] pour le reste de la démonstration.

Lemme 2.5.2 Soit $f \in C_M([1, +\infty[, \mathbb{R}_+)$ une application décroissante. Nous définissons la suite $(u_n)_{n \in \mathbb{N}^*}$ par

$$\forall n \in \mathbb{N}^* \quad u_n = \sum_{k=1}^n f(k) - \int_1^n f(t) dt.$$

Alors la suite $(u_n)_{n \in \mathbb{N}^*}$ est positive et décroissante donc convergente.

Démonstration: Par décroissance de l'application f , nous avons, pour tout $k \in \mathbb{N}^*$,

$$\forall t \in [k, k+1] \quad f(k+1) \leq f(t) \leq f(k).$$

Par croissance de l'intégrale de Riemann, nous en déduisons

$$\forall k \in \mathbb{N}^* \quad f(k+1) \leq \int_k^{k+1} f(t) dt \leq f(k). \quad (2.6)$$

Nous remarquons maintenant que, pour tout $n \in \mathbb{N}^*$,

$$u_{n+1} - u_n = f(n+1) - \int_n^{n+1} f(t) dt \leq 0,$$

cette dernière inégalité résultant de (2.6). Ceci prouve la décroissance de la suite $(u_n)_{n \in \mathbb{N}^*}$.

En outre, il est évident que $u_1 \geq 0$ et, en utilisant encore (2.6), nous obtenons

$$\forall n \geq 2 \quad u_n = \sum_{k=1}^{n-1} \left(f(k) - \int_k^{k+1} f(t) dt \right) + f(n) \geq 0,$$

ce qui nous permet de conclure. □

Corollaire 2.5.3 *La suite $(u_n)_{n \in \mathbb{N}^*}$ définie par*

$$\forall n \in \mathbb{N}^* \quad u_n = \sum_{k=1}^n \frac{1}{k} - \ln n$$

est convergente. Sa limite est notée γ et appelée constante d'Euler.

Démonstration: Il suffit d'appliquer la proposition précédente avec f définie par

$$\forall x \in [1, +\infty[\quad f(x) = \frac{1}{x}$$

qui est une application continue décroissante et positive, admettant pour primitive le logarithme népérien. □

Remarque. Une valeur approchée de la constante d'Euler est $\gamma \approx 0,577$. On ne sait pas à l'heure actuelle si γ est rationnel ou non.

Une application classique de la proposition 2.5.1 est donnée dans l'exemple suivant. Le lecteur pourra démontrer les résultats à titre d'exercice ou se référer à [DAN 244-245].

Exemple : Séries de Riemann Soit $\alpha \in \mathbb{R}$. La série de Riemann

$$\sum_{n=1}^{+\infty} \frac{1}{n^\alpha}$$

converge si et seulement si $\alpha > 1$ et, dans ce cas,

$$R_n := \sum_{k=n+1}^{+\infty} \frac{1}{k^\alpha} \sim \frac{1}{(\alpha-1)n^{\alpha-1}}.$$

Pour $\alpha \in]0, 1[$ (cas de divergence) on a

$$S_n := \sum_{k=1}^n \frac{1}{k^\alpha} \sim \frac{n^{1-\alpha}}{1-\alpha}$$

Enfin, le cas $\alpha = 1$ (cas de divergence) a déjà été traité dans le corollaire 2.5.3 qui nous donne un résultat plus précis que la simple équivalence $S_n \sim \ln n$. \square

En utilisant l'intégrale de Bertrand, le lecteur pourra démontrer le résultat suivant, qui est traité dans [CMP 229].

Exemple : Séries de Bertrand Pour tout $(\alpha, \beta) \in \mathbb{R}^2$ fixé, la série

$$\sum_{n=2}^{+\infty} \frac{1}{n^\alpha (\ln n)^\beta}$$

converge si et seulement si $\alpha > 1$ ou $(\alpha = 1 \text{ et } \beta > 1)$.

En utilisant le même genre de technique que pour prouver le lemme 2.5.2, on obtient [GOU 155] :

Proposition 2.5.4 *Soit $f \in C_M(\mathbb{R}, \mathbb{R})$ une application décroissante telle que l'intégrale $\int_0^{+\infty} f$ converge et est non nulle. Alors, pour tout $t > 0$, la série $\sum_{n \in \mathbb{N}^*} f(nt)$ converge et nous avons l'équivalent suivant lorsque $t \rightarrow 0+$:*

$$\sum_{n=1}^{+\infty} f(nt) \sim \frac{1}{t} \int_0^{+\infty} f$$

En appliquant cette proposition avec $f(t) = \exp(-t^2)$, on peut établir l'équivalent suivant lorsque $x \rightarrow 1-$:

$$\sum_{n=1}^{+\infty} x^{n^2} \sim \frac{\sqrt{\pi}}{2\sqrt{1-x}}$$

En utilisant la formule de Taylor avec reste intégral, on peut démontrer la proposition suivante [GOU 217] :

Proposition 2.5.5 *Soit $f \in C^1(\mathbb{R}_+, \mathbb{C})$ telle que :*

$$\int_1^{+\infty} |f'(t)| dt < +\infty$$

Alors la série $\sum_{n \in \mathbb{N}^} f(n)$ a la même nature que la suite $(\int_1^n f)_{n \in \mathbb{N}}$.*

On peut appliquer ce résultat à l'étude de la série suivante [GOU 217-218] en fonction du paramètre $\alpha > 0$:

$$\sum_{n \in \mathbb{N}^*} \frac{e^{i\sqrt{n}}}{n^\alpha}$$

[CMP 260-261] En utilisant une comparaison série-intégrale, on peut établir l'existence d'une constante $K > 0$ telle que $n! \sim Kn^n e^{-n} \sqrt{n}$. La formule de Wallis permet alors d'identifier cette constante K pour en déduire :

Proposition 2.5.6 (Formule de Stirling) *Nous avons l'équivalent suivant lorsque n tend vers l'infini :*

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}.$$

Terminons ce chapitre en donnant une démonstration alternative de l'existence d'une constante $K > 0$ telle que $n! \sim Kn^n e^{-n} \sqrt{n}$. Nous allons d'abord établir la *règle de Raabe-Duhamel*, souvent utile pour étudier la nature d'une série positive lorsque la règle de D'Alembert ne permet pas de conclure.

Proposition 2.5.7 (Règle de Raabe-Duhamel) *Soit $(u_n) \in (\mathbb{R}_+^*)^{\mathbb{N}}$ une suite telle que :*

$$\exists(\alpha, \beta) \in \mathbb{R} \times]1, +\infty[, \quad \frac{u_{n+1}}{u_n} = 1 - \frac{\alpha}{n} + O\left(\frac{1}{n^\beta}\right).$$

Alors, il existe une constante $K > 0$ telle que $u_n \sim K/n^\alpha$.

Démonstration: On veut montrer l'existence d'une constante $K > 0$ telle que $n^\alpha u_n \rightarrow K$, ce qui équivaut, par passage au logarithme népérien, à :

$$\exists k \in \mathbb{R}, \quad a_n := \alpha \ln n + \ln u_n \rightarrow k.$$

Si nous définissons la suite $(v_n)_{n \in \mathbb{N}}$ par :

$$v_n := a_{n+1} - a_n = \alpha \ln \left(1 + \frac{1}{n}\right) + \ln \frac{u_{n+1}}{u_n},$$

cela équivaut encore à la convergence de la série $\sum v_n$.

En utilisant le développement limité $\ln(1+x) = x + O(x^2)$ au voisinage de 0, nous calculons :

$$v_n = \alpha \left(\frac{1}{n} + O\left(\frac{1}{n^2}\right)\right) - \frac{\alpha}{n} + O\left(\frac{1}{n^\beta}\right) + O\left(\frac{1}{n^2}\right) = O\left(\frac{1}{n^\gamma}\right),$$

avec $\gamma := \min(\beta, 2) > 1$, ce qui permet de conclure par comparaison avec l'exemple de Riemann. \square

Corollaire 2.5.8 *Il existe une constante $K > 0$ telle que $n! \sim Ke^{-n} n^{n+\frac{1}{2}}$.*

Démonstration: Définissons la suite $(u_n)_{n \in \mathbb{N}^*}$ par $u_n := n! e^n n^{-n}$, si bien qu'il s'agit de montrer l'existence de $K > 0$ telle que $u_n \sim K n^{1/2}$.

Afin d'appliquer la proposition précédente, nous calculons :

$$\forall n \in \mathbb{N}^*, \quad \frac{u_{n+1}}{u_n} = (n+1)e \left(1 + \frac{1}{n}\right)^{-n} (n+1)^{-1} = \exp \left(1 - n \ln \left(1 + \frac{1}{n}\right)\right).$$

En utilisant le développement limité $\ln(1+x) = x - x^2/2 + O(x^3)$ au voisinage de 0, nous obtenons :

$$\frac{u_{n+1}}{u_n} = \exp \left(1 - n \left(\frac{1}{n} - \frac{1}{2n^2} + O\left(\frac{1}{n^3}\right)\right)\right) = \exp \left(\frac{1}{2n} + O\left(\frac{1}{n^2}\right)\right).$$

Or, $\exp u = 1 + u + O(u^2)$ au voisinage de 0, si bien que :

$$\frac{u_{n+1}}{u_n} = 1 + \frac{1}{2n} + O\left(\frac{1}{n^2}\right).$$

Nous concluons en appliquant la règle de Raabe-Duhamel avec $\alpha = -1/2$. \square

2.6 Transformée de Laplace à variable réelle

Nous suivons ici la présentation de [POM 197-199]. Pour $f \in C^0(\mathbb{R}_+, \mathbb{C})$, nous notons :

$$\Delta_f = \left\{ \lambda \in \mathbb{R}, \text{ l'intégrale impropre } \int_0^{+\infty} e^{-\lambda t} f(t) dt \text{ converge} \right\}.$$

Définition 2.6.1 Si $\Delta_f \neq \emptyset$, on appelle transformée de Laplace de f l'application $\mathcal{L}f : \Delta_f \rightarrow \mathbb{C}$ définie par :

$$\forall \lambda \in \Delta_f, \quad \mathcal{L}f(\lambda) = \int_0^{+\infty} e^{-\lambda t} f(t) dt.$$

2.6.1 Quelques propriétés générales

Proposition 2.6.2 Si $\lambda \in \Delta_f$, alors pour tout $\mu > \lambda$, $\mu \in \Delta_f$.

Démonstration: Soit $\lambda \in \Delta_f$ et $\mu > \lambda$; posons $\alpha = \mu - \lambda > 0$. Définissons l'application $G : \mathbb{R}_+ \rightarrow \mathbb{C}$ par :

$$\forall x \in \mathbb{R}_+, \quad G(x) = \int_0^x e^{-\lambda t} f(t) dt.$$

Pour tout $x \in \mathbb{R}_+$, nous avons :

$$\int_0^x e^{-\mu t} f(t) dt = \int_0^x e^{-\alpha t} e^{-\lambda t} f(t) dt = \int_0^x e^{-\alpha t} G'(t) dt.$$

Par intégration par parties, nous en déduisons :

$$\int_0^x e^{-\mu t} f(t) dt = [e^{-\alpha t} G(t)]_0^x + \alpha \int_0^x e^{-\alpha t} G(t) dt = e^{-\alpha x} G(x) + \alpha \int_0^x e^{-\alpha t} G(t) dt.$$

Or, l'application G est continue et admet par hypothèse une limite en $+\infty$, donc elle est bornée. Nous en déduisons que $\int_0^{+\infty} e^{-\alpha t} G(t) dt$ converge et

$$\lim_{x \rightarrow +\infty} \int_0^x e^{-\mu t} f(t) dt = \alpha \int_0^{+\infty} e^{-\alpha t} G(t) dt.$$

□

Corollaire 2.6.3 *L'ensemble Δ_f est soit vide, soit égal à \mathbb{R} , soit de la forme $[a, +\infty[$, soit de la forme $]a, +\infty[$, avec $a \in \mathbb{R}$.*

Démonstration: Si $\Delta_f \neq \emptyset$, il suffit de considérer $a = \inf \Delta_f \in \mathbb{R} \cup \{-\infty\}$ pour conclure grâce à la proposition précédente. □

Proposition 2.6.4 *Si $0 \in \Delta_f$, alors nous avons l'égalité :*

$$\int_0^{+\infty} f(t) dt = \lim_{\lambda \rightarrow 0+} \mathcal{L}f(\lambda).$$

Démonstration: Puisque $0 \in \Delta_f$, nous pouvons définir $F : \mathbb{R}_+ \rightarrow \mathbb{C}$ par :

$$\forall x \in \mathbb{R}_+, \quad F(x) = \int_x^{+\infty} f(t) dt.$$

Pour tout $\lambda \in \mathbb{R}_+$ et tout $x \in \mathbb{R}_+$, nous avons :

$$\int_0^x e^{-\lambda t} f(t) dt = - \int_0^x e^{-\lambda t} F'(t) dt.$$

Par intégration par parties, nous en déduisons :

$$\int_0^x e^{-\lambda t} f(t) dt = [-e^{-\lambda t} F(t)]_0^x - \lambda \int_0^x e^{-\lambda t} F(t) dt,$$

ce qui s'écrit encore

$$\int_0^x e^{-\lambda t} f(t) dt = \int_0^{+\infty} f(t) dt - e^{-\lambda x} F(x) - \lambda \int_0^x e^{-\lambda t} F(t) dt.$$

L'application F est continue et de limite nulle en $+\infty$. Par passage à la limite lorsque $x \rightarrow +\infty$ dans l'égalité précédente, nous obtenons donc :

$$\forall \lambda \in \mathbb{R}_+, \quad \mathcal{L}f(\lambda) = \int_0^{+\infty} f(t) dt - \lambda \int_0^{+\infty} e^{-\lambda t} F(t) dt.$$

Il nous reste donc à prouver que :

$$\lim_{\lambda \rightarrow 0+} \lambda \int_0^{+\infty} e^{-\lambda t} F(t) dt = 0.$$

Soit $\delta > 0$ arbitraire. Nous pouvons choisir $A > 0$ tel que, pour tout $t > A$, $|F(t)| < \delta$. Nous en déduisons, pour tout $\lambda \in \mathbb{R}_+$,

$$\left| \lambda \int_0^{+\infty} e^{-\lambda t} F(t) dt \right| \leq \lambda \int_0^A |F(t)| dt + \delta \lambda \int_A^{+\infty} e^{-\lambda t} dt \leq \lambda \int_0^A |F(t)| dt + \delta.$$

Ainsi, il existe $\lambda_0 > 0$ tel que :

$$\forall \lambda \in [0, \lambda_0[, \quad \left| \lambda \int_0^{+\infty} e^{-\lambda t} F(t) dt \right| \leq 2\delta,$$

ce qui nous permet de conclure. \square

2.6.2 Transformée de Laplace et calcul des probabilités

Application 1. Soit X une variable aléatoire de loi $\gamma(a, \lambda)$. Sa transformée de Laplace se calcule facilement sur \mathbb{R}_+ :

$$\forall t \in \mathbb{R}_+, \quad E[e^{-tX}] = \left(\frac{\lambda}{\lambda + t} \right)^a.$$

Comme la valeur de la transformée de Laplace sur \mathbb{R}_+ suffit à caractériser une loi de probabilité sur \mathbb{R}_+ , nous en déduisons immédiatement :

$$\gamma(a, \lambda) \star \gamma(b, \lambda) = \gamma(a + b, \lambda).$$

Application 2. La transformée de Laplace d'une variable aléatoire réelle bornée par 1 vérifie l'inégalité :

$$\forall t \in \mathbb{R}, \quad E[e^{-tX}] \leq \exp\left(\frac{t^2}{2}\right). \quad (2.7)$$

Cette propriété est établie dans un problème extrait du concours commun Maths II MP 2015 : l'étude des variables aléatoires discrètes *sous-gaussiennes* [MEY 87-88].

On déduit de la relation (2.7) *l'inégalité de Hoeffding*, utile en statistique pour construire des intervalles de confiance [OUV² 132-135].

Chapitre 3

Intégration sur un intervalle quelconque

Dans tout ce chapitre, I désigne un intervalle réel non vide et nous notons encore $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

Rappelons qu'une application $f : I \rightarrow \mathbb{K}$ est dite continue par morceaux sur l'intervalle I , ce que l'on note $f \in C_M(I, \mathbb{K})$, si f est continue par morceaux sur tout segment inclus dans I .

3.1 Fonctions intégrables

Nous ne ferons qu'énoncer ici les définitions et résultats essentiels sur le sujet. Le lecteur est donc encouragé à étudier l'intégrabilité d'une fonction sur un intervalle quelconque dans [DAN 465-472] ou encore dans [CMP 154-177] qui la traite en détail, parfois d'une façon redondante avec ce que nous avons vu dans le chapitre sur les intégrales impropres.

3.1.1 Le cas positif

Définition 3.1.1 Soit $f \in C_M(I, \mathbb{R}_+)$; on appelle intégrale de l'application positive f sur l'intervalle I l'élément de \mathbb{R}_+ suivant :

$$\int_I f = \sup_{[a,b] \subset I} \int_a^b f(t) dt.$$

On dit que l'application f est intégrable sur I si $\int_I f < +\infty$.

La proposition suivante fait le lien entre cette notion d'intégrabilité sur un intervalle quelconque I et celle de convergence d'une intégrale impropre que nous avons étudiée dans le chapitre précédent.

Proposition 3.1.2 *Soit $f \in C_M(I, \mathbb{R}_+)$. Alors f est intégrable sur I si, et seulement si, l'intégrale impropre de f sur I est convergente. Dans ce cas, les deux intégrales sont égales.*

Pour fixer les idées, prenons par exemple le cas où $I =]c, +\infty[$. D'après cette proposition, nous avons l'équivalence :

$$\sup_{[a,b] \subset]c, +\infty[} \int_a^b f(t) dt < +\infty \iff \text{l'intégrale impropre } \int_c^{+\infty} f(t) dt \text{ converge}$$

et dans ce cas, nous avons l'égalité :

$$\sup_{[a,b] \subset]c, +\infty[} \int_a^b f(t) dt = \int_c^{+\infty} f(t) dt,$$

ce que nous notons :

$$\int_{]c, +\infty[} f = \int_c^{+\infty} f(t) dt. \quad (3.1)$$

Dans le cas contraire, l'intégrale impropre est divergente mais nous pouvons encore écrire, grâce à la Convention 2.2.2, page 50,

$$\int_{]c, +\infty[} f = \int_c^{+\infty} f(t) dt = +\infty.$$

Ainsi, lorsque l'intégrande f est positif, l'égalité (3.1) est toujours valable.

3.1.2 Le cas réel

Rappelons ici que, pour tout nombre réel x , nous définissons sa *partie positive* $x^+ = \sup(x, 0)$ et sa *partie négative* $x^- = \sup(-x, 0)$, d'où les égalités :

$$\forall x \in \mathbb{R}, \quad x = x^+ - x^- \text{ et } |x| = x^+ + x^-.$$

Par composition, nous définirons donc pour $f \in C_M(I, \mathbb{R})$ les deux applications $f^+ \in C_M(I, \mathbb{R}_+)$ et $f^- \in C_M(I, \mathbb{R}_+)$ par $f^+(x) := \sup(f(x), 0)$ et $f^-(x) := \sup(-f(x), 0)$.

Définition et proposition 3.1.3 *Une application $f \in C_M(I, \mathbb{R})$ est dite intégrable si l'application positive $|f|$ est intégrable sur I . Ceci équivaut à l'intégrabilité des deux applications positives f^+ et f^- sur I .*

Dans ce cas, nous appellerons intégrale de l'application f sur l'intervalle I le réel suivant.

$$\int_I f := \int_I f^+ - \int_I f^-.$$

Comme précédemment, nous allons faire le lien entre cette notion et celle de convergence d'une intégrale impropre.

Proposition 3.1.4 *Soit $f \in C_M(I, \mathbb{R})$. Alors f est intégrable sur I si, et seulement si, l'intégrale impropre de f sur I est absolument convergente. Dans ce cas, $\int_I f$ et l'intégrale impropre de f sur I sont égales.*

Considérons par exemple l'application $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ définie par $f(0) = 1$ et $f(t) = \frac{\sin t}{t}$ si $t \in \mathbb{R}_+^*$; nous savons que l'intégrale impropre $\int_0^{+\infty} f(t) dt$ est semi-convergente. L'application f n'est donc pas intégrable sur $I = \mathbb{R}_+$ et l'écriture $\int_{\mathbb{R}_+} f$ n'a pas de sens !

3.1.3 Le cas complexe

Pour tout nombre complexe z , nous noterons $\Re z$ sa partie réelle et $\Im z$ sa partie imaginaire.

Définition et proposition 3.1.5 *Une application $f \in C_M(I, \mathbb{C})$ est dite intégrable si l'application positive $|f|$ est intégrable sur I . Ceci équivaut à l'intégrabilité des deux applications réelles $\Re f$ et $\Im f$ sur I .*

Dans ce cas, nous appellerons intégrale de l'application f sur l'intervalle I le nombre complexe suivant.

$$\int_I f := \int_I \Re f + i \int_I \Im f.$$

Le lecteur pourra vérifier facilement l'équivalence qui vient d'être énoncée en utilisant les trois inégalités suivantes :

$$\forall z \in \mathbb{C}, \quad |\Re z| \leq |z|, \quad |\Im z| \leq |z|, \quad |z| \leq |\Re z| + |\Im z|.$$

Proposition 3.1.6 *Soit $f \in C_M(I, \mathbb{C})$. Alors f est intégrable sur I si, et seulement si, l'intégrale impropre de f sur I est absolument convergente. Dans ce cas, $\int_I f$ et l'intégrale impropre de f sur I sont égales.*

3.2 Théorèmes de convergence

Les deux théorèmes suivants, dont la démonstration relève de la théorie de la mesure et de l'intégration abstraite de Lebesgue, sont admis.

Théorème 3.2.1 (de convergence monotone) *Soit $(f_n) \in (C_M(I, \mathbb{R}_+))^{\mathbb{N}}$ une suite **croissante** d'applications positives intégrables sur I qui converge simplement vers $f \in C_M(I, \mathbb{R}_+)$. Alors, nous avons l'égalité*

$$\int_I f(t) dt = \lim_{n \rightarrow +\infty} \int_I f_n(t) dt \in \bar{\mathbb{R}}_+.$$

Remarques.

1. En particulier, l'application f est intégrable sur I si, et seulement si, la suite $(\int_I f_n(t) dt)_{n \in \mathbb{N}}$ est majorée.
2. On note parfois cette égalité sous la forme

$$\int_I \lim \uparrow f_n = \lim \uparrow \int_I f_n.$$

Exemple : calcul de l'intégrale de Gauss avec la formule de Wallis

1. Pour tout $n \in \mathbb{N}^*$, nous définissons l'application $f_n : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ par

$$f_n(t) := \mathbf{1}_{[0, 2^{n/2}]}(t) \left(1 - \frac{t^2}{2^n}\right)^{2^n}.$$

Montrer que la suite $(f_n)_{n \in \mathbb{N}^*}$ est croissante et converge simplement vers l'application $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ définie par $f(x) = \exp(-x^2)$.

En déduire que :

$$\lim_{n \rightarrow +\infty} \int_0^{2^{n/2}} \left(1 - \frac{t^2}{2^n}\right)^{2^n} dt = \int_0^{+\infty} e^{-t^2} dt.$$

2. Démontrer l'égalité suivante pour tout $n \in \mathbb{N}^*$:

$$\sqrt{2^n} \int_0^{\pi/2} \sin^{2^{n+1}+1} x dx = \int_0^{2^{n/2}} \left(1 - \frac{t^2}{2^n}\right)^{2^n} dt.$$

En utilisant la *formule de Wallis*

$$\int_0^{\pi/2} \sin^p x dx \sim \sqrt{\frac{\pi}{2p}},$$

en déduire la valeur de l'intégrale de Gauss :

$$\int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2}.$$

Cette égalité est utile, par exemple, en calcul des probabilités pour définir la *loi de Gauss*.

Le lecteur trouvera plus loin une autre application du théorème de convergence monotone : la *formule d'Euler-Gauss* (cf. Proposition 3.3.3, page 83).

Corollaire 3.2.2 Soit $(u_n) \in (C_M(I, \mathbb{R}_+))^{\mathbb{N}}$ une suite d'applications positives intégrables sur I telle que la série $\sum u_n$ converge simplement vers $S = \sum_{n=0}^{+\infty} u_n \in C_M(I, \mathbb{R}_+)$. Alors, on a l'égalité :

$$\int_I \sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} \int_I u_n.$$

Démonstration: Pour tout $n \in \mathbb{N}$, nous posons $f_n := \sum_{k=0}^n u_k$. Nous appliquons alors le théorème précédent avec $f := S$. \square

Remarque. En particulier, S est intégrable sur I si, et seulement si, la série $\sum \int_I u_n$ converge.

Dans les deux exemples qui suivent, nous utilisons la fonction zêta de Riemann, c'est-à-dire la fonction $\zeta :]1, +\infty[\rightarrow]1, +\infty[$ définie par :

$$\forall s > 1, \quad \zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}. \quad (3.2)$$

Les propriétés de cette fonction sont étudiées dans [SKA 57].

Dans le deuxième exemple, nous utilisons aussi par anticipation la fonction Γ d'Euler qui sera définie par la formule (3.3), page 81.

Exemple. En faisant deux intégrations par parties successives, on démontre facilement que, pour tout $n \in \mathbb{N}^*$,

$$I_n := \int_0^{+\infty} e^{-nx} x^2 dx = \frac{2}{n^3}.$$

Comme d'autre part la série d'applications $\sum e^{-nx} x^2$ converge simplement sur $]0, +\infty[$ vers $S \in C^0([0, +\infty[, \mathbb{R}_+)$ définie par $S(x) = x^2/(e^x - 1)$, nous déduisons du Corollaire 3.2.2 :

$$\int_0^{+\infty} \frac{x^2}{e^x - 1} dx = \sum_{n=1}^{+\infty} \frac{2}{n^3} = 2\zeta(3).$$

Exemple. En utilisant un changement de variable simple, on établit :

$$\int_0^{+\infty} \frac{t^{x-1}}{e^t - 1} dt = \int_0^{+\infty} \sum_{n=1}^{+\infty} e^{-nt} t^{x-1} dt = \Gamma(x)\zeta(x).$$

Exemple. Pour tous $a > 0$ et $b > 0$, nous posons :

$$I_{a,b} := \int_0^1 \frac{x^{a-1}}{1+x^b} dx.$$

On vérifie immédiatement, par comparaison avec l'exemple de Riemann, que cette intégrale est bien définie, pour tous $a > 0$ et $b > 0$.

En utilisant une série géométrique, nous pouvons écrire :

$$\forall x \in [0, 1[, \quad \frac{x^{a-1}}{1+x^b} = \sum_{n=0}^{+\infty} (-1)^n x^{a-1+nb}.$$

Afin de faire apparaître une série d'applications positives, nous regroupons les termes deux à deux comme suit :

$$\forall x \in [0, 1[, \quad \frac{x^{a-1}}{1+x^b} = \sum_{k=0}^{+\infty} (x^{a-1+2kb} - x^{a-1+(2k+1)b}),$$

avec $f_{a,b,k}(x) := x^{a-1+2kb} - x^{a-1+(2k+1)b} = x^{a-1+2kb}(1-x^b) \geq 0$, pour tout réel $x \in [0, 1[$ et pour tout entier k .

Nous pouvons donc appliquer le Corollaire 3.2.2 pour obtenir :

$$I_{a,b} = \sum_{k=0}^{+\infty} \int_{[0,1[} f_{a,b,k} = \sum_{k=0}^{+\infty} \left(\frac{1}{a+2kb} - \frac{1}{a+(2k+1)b} \right).$$

Finalement, cette égalité se réécrit, pour tous $a > 0$ et $b > 0$,

$$\int_0^1 \frac{x^{a-1}}{1+x^b} dx = \sum_{n=0}^{+\infty} \frac{(-1)^n}{a+nb}.$$

En prenant $a = b = 1$, nous en déduisons :

$$\sum_{n=0}^{+\infty} \frac{(-1)^n}{n+1} = \int_0^1 \frac{dx}{1+x} = \ln 2.$$

En prenant $a = 1$ et $b = 2$, nous obtenons :

$$\sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} = \int_0^1 \frac{dx}{1+x^2} = \frac{\pi}{4}.$$

Théorème 3.2.3 (de convergence dominée) Soit $(f_n) \in (C_M(I, \mathbb{C}))^{\mathbb{N}}$ une suite d'applications qui converge simplement vers $f \in C_M(I, \mathbb{C})$. On suppose qu'il existe une application $g \in C_M(I, \mathbb{R}_+)$ intégrable sur I telle que :

$$\forall n \in \mathbb{N}, \quad \forall t \in I, \quad |f_n(t)| \leq g(t).$$

Alors, l'application f est intégrable sur I et nous avons l'égalité

$$\int_I f = \lim_{n \rightarrow +\infty} \int_I f_n.$$

Remarques.

1. C'est l'existence d'une telle application g qui est l'hypothèse essentielle : on l'appelle « majorant intégrable » ou parfois « chapeau intégrable ».
2. Les intégrales du membre de droite sont bien définies car l'existence du majorant intégrable implique l'intégrabilité des applications f_n , $n \in \mathbb{N}$.

Exemple On peut reprendre l'exemple précédent en appliquant cette fois-ci le théorème de convergence dominée [DAN 475-476].

En effet, un raisonnement de concavité permet de prouver la majoration $\ln(1+u) \leq u$, pour tout $u > -1$, d'où :

$$\forall t \in [0, \sqrt{n}[, \quad \left(1 - \frac{t^2}{n}\right)^n \leq e^{-t^2} ,$$

ce qui nous fournit un majorant intégrable.

Corollaire 3.2.4 (Intégration terme à terme) Soit $(u_n) \in (C_M(I, \mathbb{C}))^{\mathbb{N}}$ une suite d'applications intégrables sur I telle que la série $\sum u_n$ converge simplement vers $S = \sum_{n=0}^{+\infty} u_n \in C_M(I, \mathbb{C})$ et telle que la série $\sum \int_I |u_n|$ converge. Alors, S est intégrable sur I et l'on a :

$$\int_I \sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} \int_I u_n.$$

Démonstration: L'idée est d'appliquer le théorème précédent en posant, pour tout $n \in \mathbb{N}$, $f_n := \sum_{k=0}^n u_k$ et en utilisant le majorant $g := \sum |u_n|$. Notons que nos hypothèses ne garantissent pas que g soit continue par morceaux sur I . Nous admettrons (cela relève de la théorie de la mesure) que l'on peut donner néanmoins un sens à $\int_I g$ et que ceci nous permet de conclure. \square

Exemple. En procédant comme dans l'exemple qui suit le Corollaire 3.2.2, on démontre l'égalité :

$$\int_0^{+\infty} \frac{\sin x}{e^x - 1} dx = \sum_{n=1}^{+\infty} \frac{1}{n^2 + 1}.$$

On est amené à prouver la convergence de la série $\sum_{n \geq 1} \int_0^{+\infty} |\sin x| e^{-nx} dx$ en utilisant la majoration $|\sin x| \leq x$. On termine le calcul grâce à deux intégrations par parties successives.

Exemple. [MOMP 344] prouve par intégration terme à terme l'égalité :

$$\forall x \in]1, +\infty[, \quad \forall \lambda \in [-1; 1], \quad \int_0^{+\infty} \frac{t^{x-1} e^{-t}}{1 - \lambda e^{-t}} dt = \sum_{n=0}^{+\infty} \frac{\lambda^n \Gamma(x)}{(n+1)^x}.$$

En fait, le théorème de convergence dominée reste valable sous l'hypothèse plus faible que toutes les applications en jeu sont réglées au lieu d'être continues par morceaux. Ceci justifie l'application suivante [BP 136].

Application 3.2.5 (Intégration d'une dérivée bornée) *Soit f une fonction partout dérivable sur $[a, b]$, à valeurs dans \mathbb{K} et de dérivée $f' \in \mathcal{R}([a, b], \mathbb{K})$. Alors, nous avons l'égalité :*

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

Démonstration: Nous notons d'abord que, quitte à introduire l'application $g : [0, 1] \rightarrow \mathbb{K}$ définie par

$$\forall x \in [0, 1], \quad g(x) = f(a + (b - a)x),$$

et à faire un simple changement de variable, nous pouvons supposer $a = 0$ et $b = 1$ sans perte de généralité.

Nous introduisons alors la suite $(f_n)_{n \in \mathbb{N}^*}$ d'applications de $[0, 1]$ dans \mathbb{K} définies par :

$$\forall x \in \left[0, 1 - \frac{1}{n}\right], \quad f_n(x) = n \left(f\left(x + \frac{1}{n}\right) - f(x) \right) \text{ et } \forall x \in \left]1 - \frac{1}{n}, 1\right], \quad f_n(x) = 0.$$

Pour tout $n \in \mathbb{N}^*$, nous vérifions facilement que $f_n \in C_M([0, 1], \mathbb{K})$ puisque f est dérivable donc *a fortiori* continue sur $[0, 1]$.

En outre, nous constatons immédiatement que la suite $(f_n)_{n \in \mathbb{N}^*}$ converge simplement sur $[0, 1[$ vers f' .

Comme $M := \sup_{[0, 1]} |f'|$ est fini par hypothèse, le théorème des accroissements finis nous donne :

$$\forall n \in \mathbb{N}^*, \quad \forall x \in [0, 1], \quad |f_n(x)| \leq M.$$

Puisque l'application constante égale à M est intégrable sur $[0, 1]$, le théorème de convergence dominée nous donne l'égalité

$$\int_0^1 f'(x) dx = \lim_{n \rightarrow \infty} \int_0^1 f_n(x) dx$$

Pour conclure, nous allons préciser la valeur du membre de droite. Nous calculons d'abord, à l'aide d'un simple changement de variable :

$$\begin{aligned}\int_0^1 f_n(x) \, dx &= n \int_0^{1-\frac{1}{n}} f\left(x + \frac{1}{n}\right) \, dx - n \int_0^{1-\frac{1}{n}} f(x) \, dx \\ &= n \int_{\frac{1}{n}}^1 f(x) \, dx - n \int_0^{1-\frac{1}{n}} f(x) \, dx.\end{aligned}$$

Introduisons maintenant l'application $F : [0, 1] \rightarrow \mathbb{K}$ définie par

$$\forall x \in [0, 1], \quad F(x) = \int_0^x f(t) \, dt,$$

qui est une primitive de l'application continue f . Nous avons alors :

$$\int_0^1 f_n(x) \, dx = n \left(F(1) - F\left(1 - \frac{1}{n}\right) \right) - n \left(F\left(\frac{1}{n}\right) - F(0) \right),$$

d'où l'égalité suivante :

$$\lim \int_0^1 f_n(x) \, dx = F'(1) - F'(0) = f(1) - f(0),$$

ce qui nous permet de conclure. \square

Remarque. On déduit de cette application que, si $f \in \mathcal{R}([a, b], \mathbb{K})$ admet une primitive F sur $[a, b]$, alors celle-ci est nécessairement de la forme :

$$\forall x \in [a, b], \quad F(x) = \int_a^x f(t) \, dt + C, \text{ où } C \text{ est une constante.}$$

En effet, notons pour l'instant $G : [a, b] \rightarrow \mathbb{K}$ l'application définie par

$$\forall x \in [a, b], \quad G(x) = \int_a^x f(t) \, dt.$$

Puisque F est une primitive de f sur $[a, b]$, nous avons $F' = f \in \mathcal{R}([a, b], \mathbb{K})$ si bien que, d'après l'application précédente, pour tout $x \in [a, b]$

$$G(x) = \int_a^x f(t) \, dt = \int_a^x F'(t) \, dt = F(x) - F(a),$$

d'où la conclusion. \square

Les deux théorèmes qui vont apparaître dans le paragraphe suivant sont des exemples importants d'application du théorème de convergence dominée.

3.3 Intégrale dépendant d'un paramètre

Dans toute cette section, I et J désignent des intervalles réels d'intérieurs non vides (i.e. ces intervalles ne sont ni vides ni réduits à un point).

Nous allons étudier la régularité d'une application $F : I \rightarrow \mathbb{C}$ de la forme

$$\forall x \in I, \quad F(x) = \int_J f(x, t) dt,$$

définie sous des hypothèses appropriées.

3.3.1 Étude de la continuité

Théorème 3.3.1 (de continuité) *Considérons $f : I \times J \rightarrow \mathbb{C}$ telle que, pour tout $t \in J$ fixé, l'application $x \mapsto f(x, t)$ est continue sur I et, pour tout $x \in I$ fixé, l'application $t \mapsto f(x, t)$ est continue par morceaux sur J .*

Supposons que l'hypothèse suivante, dite de domination, est satisfaite :

$$\exists g \in C_M(J, \mathbb{R}_+) \text{ intégrable sur } J \text{ telle que } \forall (x, t) \in I \times J, \quad |f(x, t)| \leq g(t).$$

Nous pouvons alors définir l'application $F : I \rightarrow \mathbb{C}$ par :

$$\forall x \in I, \quad F(x) = \int_J f(x, t) dt$$

et cette application F est continue sur l'intervalle I .

Démonstration: Notons d'abord que l'hypothèse de domination implique l'intégrabilité de $t \mapsto f(x, t)$ sur J , ce qui nous permet de définir F .

Soit $x \in I$ arbitraire fixé. Nous allons montrer que F est continue au point x en utilisant la caractérisation séquentielle de la continuité.

Considérons donc une suite $(x_n) \in I^{\mathbb{N}}$ qui converge vers x et définissons, pour tout $n \in \mathbb{N}$, une application $f_n : J \rightarrow \mathbb{C}$ par $f_n(t) = f(x_n, t)$.

Nos hypothèses impliquent que $f_n \in C_M(J, \mathbb{C})$ pour tout entier n .

De même, définissons $f^x \in C_M(J, \mathbb{C})$ par $f^x(t) = f(x, t)$. Nos hypothèses entraînent que la suite d'applications $(f_n)_{n \in \mathbb{N}}$ converge simplement sur J vers l'application f^x .

En outre, l'application g nous fournit un majorant intégrable puisque :

$$\forall n \in \mathbb{N} \quad \forall t \in J \quad |f_n(t)| \leq g(t).$$

Les conditions sont donc réunies pour appliquer le théorème de convergence dominée, qui nous donne :

$$\lim_{n \rightarrow +\infty} \int_J f_n(t) dt = \int_J f^x(t) dt,$$

ce qui s'écrit encore :

$$\lim_{n \rightarrow +\infty} F(x_n) = F(x).$$

Nous en déduisons la continuité de F au point x . Finalement, comme x avait été fixé arbitrairement dans I , nous concluons : $F \in C^0(I, \mathbb{C})$. \square

Exemple. [CMP 191] Soit $f \in C^0(\mathbb{R}, \mathbb{C})$, supposée intégrable sur \mathbb{R} , et soit $g \in C^0(\mathbb{R}, \mathbb{C})$, supposée bornée. Le *produit de convolution* de f et g est l'application $f * g : \mathbb{R} \rightarrow \mathbb{C}$ définie par :

$$\forall x \in \mathbb{R}, \quad f * g(x) = \int_{-\infty}^{+\infty} f(t)g(x-t) dt.$$

Le théorème précédent implique immédiatement que $f * g \in C^0(\mathbb{R}, \mathbb{C})$.

Exercice. Pb 5.2 de [CMP 346] paragraphe I.

Pour toute application $f \in C_M(\mathbb{R}, \mathbb{C})$ intégrable sur $] -\infty, +\infty[$, nous définissons sa *transformée de Fourier* $\mathcal{F}f : \mathbb{R} \rightarrow \mathbb{C}$ par :

$$\forall x \in \mathbb{R} \quad \mathcal{F}f(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-ixt} dt$$

Montrer que $\mathcal{F}f$ est continue et bornée sur \mathbb{R} .

Remarque. La continuité étant une propriété locale, nous obtenons la même conclusion sous une hypothèse plus faible, dite *de domination locale* : Pour tout segment K inclus dans I , il existe une application $g_K \in C_M(J, \mathbb{R}_+)$ intégrable sur J et telle que

$$\forall (x, t) \in K \times J \quad |f(x, t)| \leq g_K(t).$$

En effet, ceci entraîne la continuité de l'application F sur tout segment K inclus dans I et nous en déduisons sa continuité en tout point de I . Le lecteur trouvera la preuve formelle dans [CMP 191] mais nous pourrions aussi simplement en appliquer l'idée au cas par cas, comme nous allons le voir dans l'exemple suivant.

Application à la fonction Γ d'Euler Nous allons justifier la définition de la fonction $\Gamma : \mathbb{R}_+^* \rightarrow \mathbb{R}$ par

$$\forall x \in \mathbb{R}_+^*, \quad \Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt \quad (3.3)$$

et montrer que cette fonction est continue sur \mathbb{R}_+^* .

Soit $x > 0$ fixé ; montrons que $\Gamma(x)$ est bien définie. Puisque l'application $t \mapsto t^{x-1}e^{-t}$ est continue sur $]0, +\infty[$, il reste à étudier la convergence de l'intégrale impropre en 0 et $+\infty$. Comme nous avons l'équivalence suivante lorsque $t \rightarrow 0+$:

$$t^{x-1}e^{-t} \sim t^{x-1},$$

avec $x-1 > -1$, nous obtenons par comparaison avec l'exemple de Riemann la convergence de notre intégrale impropre en 0.

Pour étudier la convergence de l'intégrale impropre en $+\infty$, nous utilisons

$$\lim_{t \rightarrow +\infty} t^{x+1}e^{-t} = 0,$$

pour en déduire

$$\exists T > 0, \quad \forall t \geq T, \quad t^{x-1}e^{-t} \leq \frac{1}{t^2},$$

ce qui nous donne la convergence de notre intégrale impropre en $+\infty$ par comparaison avec l'exemple de Riemann.

Considérons maintenant un segment $[a, b] \subset \mathbb{R}_+^*$. L'application $(x, t) \mapsto t^{x-1}e^{-t}$ est clairement continue sur $[a, b] \times \mathbb{R}_+^*$. En outre, nous avons la majoration suivante :

$$\forall (x, t) \in [a, b] \times \mathbb{R}_+^* \quad |t^{x-1}e^{-t}| \leq (t^{a-1} + t^{b-1})e^{-t},$$

cette dernière application étant intégrable sur $]0, +\infty[$ d'après ce qui précède. Le Théorème 3.3.1 nous donne alors la continuité de l'application Γ sur $[a, b]$. Enfin, ce segment $[a, b] \subset \mathbb{R}_+^*$ ayant été choisi de façon arbitraire, nous en déduisons la continuité de la fonction Gamma d'Euler sur \mathbb{R}_+^* . □

Remarque : Il n'était pas possible de prouver la continuité de la fonction Γ sur \mathbb{R}_+^* par domination globale car, pour tout $t \in]0, 1[$, nous avons :

$$\sup_{x \in \mathbb{R}_+^*} t^{x-1}e^{-t} = \frac{e^{-t}}{t},$$

et cette dernière application n'est pas intégrable sur $]0, +\infty[$ puisqu'elle est équivalente à $1/t$ au voisinage de 0. □

Le lecteur montrera facilement à l'aide d'une intégration par parties généralisée l'égalité suivante :

$$\forall x \in \mathbb{R}_+^* \quad \Gamma(x+1) = x \Gamma(x)$$

et il en déduira, par une simple récurrence, les valeurs particulières prises par la fonction Gamma d'Euler sur les entiers :

$$\forall n \in \mathbb{N}^* \quad \Gamma(n) = (n-1)!$$

Nous résumons ces résultats dans la proposition suivante.

Proposition 3.3.2 *On définit bien une application $\Gamma : \mathbb{R}_+^* \rightarrow \mathbb{R}$ en posant :*

$$\forall x \in \mathbb{R}_+^*, \quad \Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt.$$

En outre, $\Gamma \in C^0(\mathbb{R}_+^, \mathbb{R})$, $\Gamma(1) = 1$ et on a l'égalité :*

$$\forall x \in \mathbb{R}_+^*, \quad \Gamma(x+1) = x \Gamma(x).$$

Proposition 3.3.3 (Formule d'Euler-Gauss)

$$\forall x \in \mathbb{R}_+^*, \quad \Gamma(x) = \lim_{n \rightarrow +\infty} \frac{n^x n!}{x(x+1) \cdots (x+n)}.$$

Démonstration: [CMP 306] Pour tout $n \in \mathbb{N}^*$, soit $g_n : [0, n[\rightarrow \mathbb{R}$ l'application définie par :

$$\forall x \in [0, n[, \quad g_n(x) := (n+1) \ln \left(1 - \frac{x}{n+1}\right) - n \ln \left(1 - \frac{x}{n}\right).$$

On vérifie facilement que $g'_n \geq 0$, puis que $g_n \geq 0$, d'où :

$$\forall x \in [0, n[\quad \left(1 - \frac{x}{n}\right)^n \leq \left(1 - \frac{x}{n+1}\right)^{n+1}.$$

Nous pouvons alors appliquer le thm de convergence monotone pour obtenir :

$$\forall x \in \mathbb{R}_+^*, \quad \int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1} dt \xrightarrow{n \rightarrow +\infty} \Gamma(x).$$

Pour tout $n \in \mathbb{N}^*$, le changement de variable $u = t/n$ nous donne :

$$\int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1} dt = n^x \int_0^1 (1-u)^n u^{x-1} du.$$

Il nous reste à effectuer des intégrations par parties successives dans cette dernière intégrale pour conclure. □

La formule d'Euler-Gauss permet d'établir le lien suivant entre la constante γ d'Euler et la fonction Γ d'Euler [CMP 306].

Corollaire 3.3.4 (Formule de Weierstrass) *Pour tout $x \in \mathbb{R}_+^*$, on a :*

$$\frac{1}{\Gamma(x)} = x \exp(\gamma x) \prod_{n=1}^{+\infty} \left[\left(1 + \frac{x}{n}\right) \exp\left(-\frac{x}{n}\right) \right].$$

3.3.2 Étude de la dérivabilité

Théorème 3.3.5 (de dérivation) *Considérons $f : I \times J \rightarrow \mathbb{C}$ telle que, pour tout $x \in I$ fixé, l'application $t \mapsto f(x, t)$ est continue par morceaux et intégrable sur J .*

Nous pouvons alors définir l'application $F : I \rightarrow \mathbb{C}$ par :

$$\forall x \in I, \quad F(x) = \int_J f(x, t) dt.$$

Supposons en outre que, pour tout $t \in J$ fixé, l'application $x \mapsto f(x, t)$ est dérivable (resp. de classe C^1) sur I et que, pour tout $x \in I$ fixé, l'application $t \mapsto \frac{\partial f}{\partial x}(x, t)$ est continue par morceaux sur J .

Si l'hypothèse (de domination) suivante est satisfaite :

$$\exists g \in C_M(J, \mathbb{R}_+) \text{ intégrable sur } J \text{ t.q. } \forall (x, t) \in I \times J, \quad \left| \frac{\partial f}{\partial x}(x, t) \right| \leq g(t),$$

alors l'application F est dérivable (resp. de classe C^1) sur I , de dérivée :

$$\forall x \in I, \quad F'(x) = \int_J \frac{\partial f}{\partial x}(x, t) dt. \quad (3.4)$$

Démonstration: Soit $x \in I$ fixé. Comme dans le théorème précédent, c'est par une caractérisation séquentielle que nous allons démontrer la dérivabilité de l'application F au point x et exprimer sa dérivée.

Nous considérons donc une suite $(x_n) \in (I \setminus \{x\})^{\mathbb{N}}$ qui converge vers x et, pour tout $n \in \mathbb{N}$, nous définissons une application $h_n : J \rightarrow \mathbb{C}$ par

$$\forall t \in J, \quad h_n(t) = \frac{f(x_n, t) - f(x, t)}{x_n - x},$$

de sorte que

$$\frac{F(x_n) - F(x)}{x_n - x} = \int_J h_n(t) dt. \quad (3.5)$$

Nous aimerions appliquer le théorème de convergence dominée à l'intégrale du membre de droite afin de prouver que le membre de gauche admet une limite. La suite (x_n) étant arbitraire, ceci entraînera la dérivabilité de l'application F au point x par caractérisation séquentielle d'une limite.

Pour cela, nous remarquons d'abord que les hypothèses sur f impliquent :

$$\forall t \in J, \quad \lim_{n \rightarrow +\infty} h_n(t) = \frac{\partial f}{\partial x}(x, t).$$

Nous cherchons maintenant un majorant intégrable. L'hypothèse de domination et *l'inégalité des accroissements finis* nous donnent :

$$\forall t \in J, \quad \forall n \in \mathbb{N}, \quad |h_n(t)| \leq g(t).$$

Nous pouvons donc appliquer le théorème de convergence dominée pour passer à la limite lorsque n tend vers l'infini dans l'égalité (3.5) et obtenir

$$\lim_{n \rightarrow +\infty} \frac{F(x_n) - F(x)}{x_n - x} = \lim_{n \rightarrow +\infty} \int_J h_n(t) dt = \int_J \frac{\partial f}{\partial x}(x, t) dt,$$

ce qui prouve la dérivabilité de l'application F au point x et nous donne l'égalité (3.4).

Si, en outre, l'application $x \mapsto \frac{\partial f}{\partial x}(x, t)$ est continue sur I , notre hypothèse de domination permet de démontrer la continuité de F' grâce à l'égalité (3.4) et au Théorème 3.3.1.

□

Application : Transformée de Fourier. Pb 5.2 de [CMP 346] V 1)3).

Soit $n \in \mathbb{N}^*$ et soit $f \in C_M(\mathbb{R}, \mathbb{C})$ telle que les applications $f, f_1 : t \mapsto tf(t), \dots, f_n : t \mapsto t^n f(t)$ soient toutes intégrables sur $] -\infty, +\infty[$.

Alors, $\mathcal{F}f \in C^n(\mathbb{R}, \mathbb{C})$ et nous avons les égalités :

$$\forall k \in \llbracket 1; n \rrbracket, \quad \forall x \in \mathbb{R}, \quad (\mathcal{F}f)^{(k)}(x) = (-i)^k \mathcal{F}f_k(x).$$

Remarque. La dérivabilité étant une propriété locale, nous obtenons la même conclusion sous une hypothèse plus faible, dite de domination locale : Pour tout segment K inclus dans I , il existe une application $g_K \in C_M(J, \mathbb{R}_+)$ intégrable sur J et telle que

$$\forall (x, t) \in K \times J, \quad \left| \frac{\partial f}{\partial x}(x, t) \right| \leq g_K(t).$$

En effet, ceci entraîne la dérivabilité de l'application F sur tout segment K inclus dans I et l'égalité :

$$\forall x \in K, \quad F'(x) = \int_J \frac{\partial f}{\partial x}(x, t) dt.$$

Nous en déduisons la dérivabilité de F en tout point de I et l'égalité

$$\forall x \in I, \quad F'(x) = \int_J \frac{\partial f}{\partial x}(x, t) dt.$$

□

Application : Intégrale de Gauss [DAN 479-480] calcule l'intégrale de Gauss en établissant l'égalité $F' + G' = 0$ pour les applications $F, G : \mathbb{R} \rightarrow \mathbb{R}$ définies par :

$$\forall x \in \mathbb{R}, \quad F(x) = \int_0^1 \frac{\exp(-x^2(1+t^2))}{1+t^2} dt \text{ et } G(x) = \left(\int_0^x \exp(-t^2) dt \right)^2.$$

Application : Intégrale de Poisson [MOMP 203] Cette intégrale est définie, pour tout $x \in \mathbb{R} \setminus \{-1; 1\}$, par :

$$I(x) := \int_0^\pi \ln(1 - 2x \cos t + x^2) dt.$$

On démontre que, pour tout $x \in]-1; 1[\setminus \{0\}$, $I(1/x) = I(x) - 2\pi \ln|x|$, ce qui permet de restreindre l'étude de la valeur de I au cas $x \in]-1; 1[$. Le théorème de dérivation implique que $I \in C^1(]-1; 1[, \mathbb{R})$, avec

$$\forall x \in]-1; 1[, \quad I'(x) = \int_0^\pi \frac{2x - 2 \cos t}{1 - 2x \cos t + x^2} dt.$$

On effectue dans cette dernière intégrale le changement de variable $u = \tan(t/2)$ et l'on obtient, après quelques lignes de calcul, sa nullité. Il en résulte que I est constante sur $]-1; 1[$ et, comme $I(0) = 0$, $I|_{]-1; 1[} \equiv 0$. Finalement,

$$|x| < 1 \Rightarrow I(x) = 0 \quad ; \quad |x| > 1 \Rightarrow I(x) = 2\pi \ln|x|.$$

Application : Théorème de Weierstrass trigonométrique Le lecteur trouvera une démonstration de ce théorème utilisant le produit de convolution dans l'ouvrage Oraux X-ENS, Analyse 2, Exercice 4.26.

Application : Transformée de Laplace Voir section 3.5.

3.4 La fonction Γ d'Euler

Bibliographie : Le « sujet d'étude » numéro 1 de [GOU 295] établit de nombreuses propriétés de la fonction Γ d'Euler.

Nous avons introduit la fonction Γ d'Euler dans la proposition 3.3.2. En appliquant le théorème de dérivation 3.3.5 sous une hypothèse de domination locale, nous obtenons :

Proposition 3.4.1 *La fonction Γ d'Euler est de classe C^∞ sur \mathbb{R}_+^* et*

$$\forall k \in \mathbb{N} \quad \forall x \in \mathbb{R}_+^* \quad \Gamma^{(k)}(x) = \int_0^{+\infty} (\ln t)^k t^{x-1} e^{-t} dt.$$

Le lecteur trouvera la démonstration détaillée dans [CMP 200-202], qui établit des propriétés supplémentaires et trace l'allure de la courbe représentative de la fonction Γ .

En particulier, nous constatons immédiatement que $\Gamma'' \geq 0$ ce qui implique la convexité de Γ . En fait, nous avons même mieux :

Définition et proposition 3.4.2 *Soient I un intervalle réel non trivial et $f : I \rightarrow \mathbb{R}_+^*$. On dit que l'application f est logarithmiquement convexe si l'application $\ln f$ est convexe.*

Toute application logarithmiquement convexe est convexe.

Démonstration: Soient $(x, y) \in I^2$ et $\lambda \in [0, 1]$. Par hypothèse, on a :

$$\ln f(\lambda x + (1 - \lambda)y) \leq \lambda \ln f(x) + (1 - \lambda) \ln f(y).$$

Le logarithme étant une fonction concave, on en déduit :

$$\ln f(\lambda x + (1 - \lambda)y) \leq \ln (\lambda f(x) + (1 - \lambda)f(y)).$$

On obtient alors le résultat voulu en passant à l'exponentielle dans cette dernière inégalité. □

Remarque. Le lecteur constatera que la réciproque est fautive en prenant $f(x) = x^2$.

Proposition 3.4.3 *La fonction Γ d'Euler est logarithmiquement convexe.*

Démonstration: Comme nous avons les égalités

$$(\ln \Gamma)'' = \left(\frac{\Gamma'}{\Gamma} \right)' = \frac{\Gamma \Gamma'' - (\Gamma')^2}{\Gamma^2},$$

il suffit de prouver que $(\Gamma'(x))^2 \leq \Gamma(x)\Gamma''(x)$ pour tout $x > 0$.

Or ceci résulte immédiatement de l'inégalité de Cauchy-Schwarz appliquée à $f(t) = t^{\frac{x-1}{2}} e^{-\frac{t}{2}}$ et $g(t) = \ln t \, t^{\frac{x-1}{2}} e^{-\frac{t}{2}}$:

$$\left(\int_0^{+\infty} \ln t \, t^{x-1} e^{-t} dt \right)^2 \leq \int_0^{+\infty} t^{x-1} e^{-t} dt \int_0^{+\infty} (\ln t)^2 t^{x-1} e^{-t} dt$$

□

Nous renvoyons le lecteur à [GOU 295] pour établir le tableau de variations suivant, l'équivalent $\Gamma(x) \sim 1/x$ lorsque $x \rightarrow 0+$, l'existence d'une branche parabolique donnée par $\Gamma(x)/x \rightarrow +\infty$ lorsque $x \rightarrow +\infty$, ainsi que l'allure du graphe de Γ .

x	0	1	c	2	$+\infty$
$\Gamma'(x)$		−	0	+	
$\Gamma(x)$	$+\infty$	1	m	1	$+\infty$

La notion de convexité logarithmique nous permet d'écrire une caractérisation de la fonction Γ d'Euler.

Théorème 3.4.4 (Bohr-Mollerup) *Soit $f : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ une application logarithmiquement convexe telle que $f(1) = 1$ et :*

$$\forall x \in \mathbb{R}_+^*, \quad f(x+1) = xf(x). \quad (3.6)$$

Alors, $f = \Gamma$.

Démonstration: À partir de (3.6), une récurrence facile nous donne :

$$\forall n \in \mathbb{N}^*, \quad \forall x \in \mathbb{R}_+^*, \quad f(n+x) = f(x) \prod_{0 \leq k < n} (x+k). \quad (3.7)$$

En particulier, avec la condition $f(1) = 1$, nous en déduisons $f(n+1) = n!$. La convexité de l'application $\ln f$ nous donne, par passage à l'exponentielle,

$$\forall (u, v) \in (\mathbb{R}_+^*)^2, \quad \forall x \in [0, 1], \quad f(xu + (1-x)v) \leq f(u)^x f(v)^{1-x}. \quad (3.8)$$

Pour tout $n \in \mathbb{N}^*$, en écrivant $n+x = x(n+1) + (1-x)n$, nous en déduisons :

$$\forall n \in \mathbb{N}^*, \quad \forall x \in [0, 1], \quad f(n+x) \leq f(n+1)^x f(n)^{1-x} = n^x (n-1)!$$

D'autre part, en écrivant $n+1 = x(n+x) + (1-x)(n+1+x)$, nous déduisons aussi de la formule (3.8) :

$$n! = f(n+1) \leq f(n+x)^x f(n+1+x)^{1-x} = (n+x)^{1-x} f(n+x).$$

Nous avons ainsi établi l'encadrement $n!(n+x)^{x-1} \leq f(n+x) \leq (n-1)!n^x$, qui s'écrit encore, d'après l'égalité (3.7),

$$\forall n \in \mathbb{N}^*, \quad \forall x \in]0, 1], \quad \frac{n!(n+x)^{x-1}}{\prod_{0 \leq k < n} (x+k)} \leq f(x) \leq \frac{(n-1)!n^x}{\prod_{0 \leq k < n} (x+k)}. \quad (3.9)$$

L'application Γ satisfaisant les conditions de l'énoncé, nous avons donc :

$$\forall n \in \mathbb{N}^*, \quad \forall x \in]0, 1], \quad \frac{n!(n+x)^{x-1}}{\prod_{0 \leq k < n} (x+k)} \leq \Gamma(x) \leq \frac{(n-1)!n^x}{\prod_{0 \leq k < n} (x+k)}.$$

Or, nous constatons facilement que les termes extrêmes de cet encadrement sont équivalents lorsque n tend vers l'infini. Il est facile d'en déduire qu'ils convergent tous deux vers $\Gamma(x)$:

$$\forall x \in]0, 1], \quad \lim_{n \rightarrow +\infty} \frac{n!(n+x)^{x-1}}{\prod_{0 \leq k < n} (x+k)} = \lim_{n \rightarrow +\infty} \frac{(n-1)!n^x}{\prod_{0 \leq k < n} (x+k)} = \Gamma(x). \quad (3.10)$$

Revenant à l'encadrement (3.9) et passant à la limite lorsque n tend vers l'infini, nous en déduisons que $f(x) = \Gamma(x)$ pour tout $x \in]0, 1]$.

Finalement, la condition (3.6) nous permet de conclure que cette égalité reste vraie sur \mathbb{R}_+^* .

□

Remarque. Nous retrouvons au passage, en utilisant (3.10), la *formule d'Euler-Gauss* (cf. Proposition 3.3.3).

Corollaire 3.4.5 (Fonction Bêta) *Nous pouvons bien définir une fonction $B : (\mathbb{R}_+^*)^2 \rightarrow \mathbb{R}_+^*$ en posant :*

$$\forall (x, y) \in (\mathbb{R}_+^*)^2, \quad B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt. \quad (3.11)$$

Nous avons alors l'égalité :

$$\forall (x, y) \in (\mathbb{R}_+^*)^2, \quad B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}. \quad (3.12)$$

Le lecteur vérifiera facilement la première phrase par comparaison avec l'exemple de Riemann. Nous le renvoyons à [RUD 180-181] pour déduire l'égalité (3.12) du théorème précédent.

La formule de Stirling qui nous donne un équivalent de $n!$ lorsque $n \rightarrow +\infty$ admet la généralisation suivante [CMP 306].

Proposition 3.4.6 (Formule de Stirling) *Nous avons l'équivalence suivante lorsque $x \rightarrow +\infty$:*

$$\Gamma(x+1) \sim \left(\frac{x}{e}\right)^x \sqrt{2\pi x}.$$

[GOU 160-162] retrouve ce résultat en appliquant la *méthode de Laplace* mais cette preuve est plus compliquée.

3.5 Transformation de Laplace à variable complexe

Référence : chapitre 8 de *Distributions, analyse de Fourier et transformation de Laplace* par Ahmed LESFARI chez Ellipses.

Pour $f \in C^0(\mathbb{R}_+, \mathbb{C})$, nous noterons :

$$D_f = \left\{ z \in \mathbb{C}, \quad \int_0^{+\infty} |e^{-zt} f(t)| dt < +\infty \right\}.$$

Définition 3.5.1 *Soit $f \in C^0(\mathbb{R}_+, \mathbb{C})$ telle que $D_f \neq \emptyset$. On appelle transformée de Laplace de f l'application $\mathcal{L}f : D_f \rightarrow \mathbb{C}$ définie par :*

$$\forall z \in D_f, \quad \mathcal{L}f(z) = \int_0^{+\infty} e^{-zt} f(t) dt.$$

Si $z_0 \in D_f$, on constate que $\Re z \geq \Re z_0 \Rightarrow z \in D_f$; en effet, il suffit d'écrire $|\exp(-zt)| = \exp(-\Re z t)$ puis d'appliquer un théorème de comparaison. Cette remarque nous conduit à introduire la définition suivante.

Définition 3.5.2 *On appelle abscisse de sommabilité de $f \in C^0(\mathbb{R}_+, \mathbb{C})$ l'élément de $\bar{\mathbb{R}}$ suivant (avec la convention $\inf \emptyset = +\infty$) :*

$$\alpha_f = \inf \left\{ s \in \mathbb{R}, \quad \int_0^{+\infty} e^{-st} |f(t)| dt < +\infty \right\}.$$

En utilisant la remarque précédente, le lecteur montrera les implications :

$$\forall z \in \mathbb{C}, \quad \Re z > \alpha_f \Rightarrow z \in D_f \quad \text{et} \quad \Re z < \alpha_f \Rightarrow z \notin D_f.$$

Pour $\Re z = \alpha_f$, on ne peut rien dire de façon générale, il faut examiner la situation au cas par cas. Résumons-nous :

3.5. TRANSFORMATION DE LAPLACE À VARIABLE COMPLEXE 91

Proposition 3.5.3 *Le demi-plan ouvert de sommabilité de $f \in C^0(\mathbb{R}_+, \mathbb{C})$ est défini par $\mathcal{P}_f = \{z \in \mathbb{C}, \Re z > \alpha_f\}$. Nous avons les inclusions :*

$$\mathcal{P}_f \subseteq D_f \subseteq \overline{\mathcal{P}_f}.$$

Proposition 3.5.4 *Soit $f \in C^0(\mathbb{R}_+, \mathbb{C})$ une application bornée. Alors, on a l'inclusion $\mathbb{R}_+^* \subseteq D_f$ et la restriction de l'application $\mathcal{L}f$ à \mathbb{R}_+^* est de classe C^∞ , de dérivées successives données par*

$$\forall n \in \mathbb{N}^*, \quad \forall \lambda \in \mathbb{R}_+^*, \quad (\mathcal{L}f)^{(n)}(\lambda) = (-1)^n \int_0^{+\infty} e^{-\lambda t} t^n f(t) dt. \quad (3.13)$$

Démonstration: Pour tout $\lambda > 0$, l'inégalité

$$\forall t \in \mathbb{R}_+, \quad |\exp(-\lambda t)f(t)| \leq \|f\|_\infty \exp(-\lambda t) \quad (3.14)$$

implique l'absolue convergence de l'intégrale impropre $\int_0^{+\infty} e^{-\lambda t} f(t) dt$. On en déduit que $\mathbb{R}_+^* \subseteq D_f$.

Soit $\lambda_0 > 0$ arbitraire fixé. La dérivation étant une notion locale, il suffit de prouver (3.13) pour tout $\lambda \in]\lambda_0, +\infty[$.

On procède par récurrence sur $n \in \mathbb{N}^*$ en appliquant le Théorème 3.3.5 avec le majorant intégrable $t \mapsto t^n \exp(-\lambda_0 t) \|f\|_\infty$ sur \mathbb{R}_+ . \square

Application au calcul de l'intégrale de Dirichlet. Nous allons établir :

$$\int_0^{+\infty} \frac{\sin t}{t} dt = \frac{\pi}{2}. \quad (3.15)$$

Soit $f \in C^0(\mathbb{R}_+, \mathbb{R})$ l'application *sinus cardinal* définie par $f(0) = 1$ et, pour tout $t > 0$, $f(t) = (\sin t)/t$.

Nous savons que l'intégrale impropre $\int_0^{+\infty} f(t) dt$ converge, donc $0 \in \Delta_f$, ce qui nous permet d'appliquer la Proposition 2.6.4, page 68, pour obtenir :

$$\int_0^{+\infty} \frac{\sin t}{t} dt = \lim_{\lambda \rightarrow 0^+} \mathcal{L}f(\lambda). \quad (3.16)$$

Par ailleurs, l'application f est bornée (avec $\|f\|_\infty = 1$) donc nous pouvons appliquer la proposition précédente pour obtenir $(\mathcal{L}f)|_{\mathbb{R}_+^*} \in C^\infty(\mathbb{R}_+^*, \mathbb{R})$ et

$$\forall \lambda \in \mathbb{R}_+^*, \quad (\mathcal{L}f)'(\lambda) = - \int_0^{+\infty} e^{-\lambda t} \sin t dt.$$

On calcule facilement cette dernière intégrale par deux intégrations par parties successives, d'où :

$$\forall \lambda \in \mathbb{R}_+^*, \quad (\mathcal{L}f)'(\lambda) = \frac{-1}{1 + \lambda^2}.$$

Nous en déduisons l'existence d'une constante réelle C telle que :

$$\forall \lambda \in \mathbb{R}_+^*, \quad \mathcal{L}f(\lambda) = C - \arctan \lambda.$$

En utilisant l'inégalité (3.14), on montre facilement que $\mathcal{L}f(\lambda) \rightarrow 0$ lorsque $\lambda \rightarrow +\infty$. Nous en déduisons que $C = \pi/2$. Ainsi, nous avons l'égalité :

$$\forall \lambda \in \mathbb{R}_+^*, \quad \mathcal{L}f(\lambda) = \frac{\pi}{2} - \arctan \lambda.$$

Nous concluons alors par la formule (3.16). □

Exemple : Transformée de Laplace de la densité de Gauss.

1^{re} méthode : équation différentielle. Exercice 3.5.10 de [CMP 204], légèrement modifié.

On rappelle que la densité de Gauss est l'application $g : \mathbb{R} \rightarrow \mathbb{R}_+$, définie par $g(t) = (1/\sqrt{2\pi}) \exp(-t^2/2)$. On vérifie que $D_g = \mathbb{C}$ par théorème de comparaison. Ensuite, par changement de variable, on réduit le problème au calcul de $\mathcal{L}g(z)$ lorsque z est imaginaire pur. On étudie donc la fonction $F : \mathbb{R} \rightarrow \mathbb{R}$, définie par :

$$\forall y \in \mathbb{R}, \quad F(y) = \int_{-\infty}^{+\infty} e^{iyt} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt,$$

qui n'est autre que la transformée de Fourier de g .

Grâce au Théorème 3.3.5, on montre qu'elle satisfait l'équation différentielle $F'(y) = -yF(y)$, avec la condition initiale $F(0) = 1$. On en déduit que $F(y) = \exp(-y^2/2)$, pour tout $y \in \mathbb{R}$.

Revenant au cas général, on obtient finalement :

$$\forall z \in \mathbb{C}, \quad \mathcal{L}g(z) = e^{-\frac{z^2}{2}}.$$

2^e méthode : théorème des zéros isolés. Nous renvoyons le lecteur à l'Application 11.4.3, page 206.

Chapitre 4

Intégrales multiples

Nous ne donnerons dans ce chapitre que les définitions et résultats essentiels. Pour alléger les écritures, nous ne traiterons que le cas de l'intégrale double : tous les énoncés se généralisent sans difficulté aux intégrales multiples.

Le programme officiel précise que tous les théorèmes relatifs aux intégrales multiples (notamment Fubini et changement de variable) sont admis.

Le lecteur trouvera dans le chapitre 12 de [CMPSI 395-419] les énoncés qui se ramènent à des intégrales simples sur des segments, à savoir ceux qui concernent des intégrales multiples définies sur un sous-ensemble borné (et géométriquement simple) de \mathbb{R}^2 ou \mathbb{R}^3 ; des exercices sur ce thème se trouvent dans le chapitre 12 de [EMPSI 183-194].

Les énoncés qui font apparaître des fonctions intégrables sur des intervalles quelconques pourront être étudiés dans le chapitre 3 de [CMP 206-216] : ils permettent d'intégrer une application continue sur certains sous-ensembles non bornés de \mathbb{R}^2 ou \mathbb{R}^3 .

4.1 Intégrale double sur un domaine simple borné

Proposition et définition 4.1.1 *Soit $f : [a, b] \times [c, d] \rightarrow \mathbb{K}$ continue. Alors nous avons l'égalité :*

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

La valeur commune de ces deux intégrales est appelée intégrale double de f sur le pavé $[a, b] \times [c, d]$ et notée :

$$\iint_{[a, b] \times [c, d]} f(x, y) \, dx \, dy$$

Définition 4.1.2 On appelle domaine simple du plan toute partie D de \mathbb{R}^2 qui vérifie l'une des deux propriétés suivantes.

1. Il existe $\phi_1, \phi_2 \in C_M([a, b], \mathbb{R})$ telles que :

$$D = \{(x, y) \in \mathbb{R}^2 ; a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x)\} \quad (4.1)$$

2. Il existe $\psi_1, \psi_2 \in C_M([c, d], \mathbb{R})$ telles que :

$$D = \{(x, y) \in \mathbb{R}^2 ; c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\} \quad (4.2)$$

Proposition et définition 4.1.3 Considérons un domaine simple D du plan et une application continue $f : D \rightarrow \mathbb{K}$ continue. Nous définissons l'intégrale double de l'application f sur le domaine D comme suit :

1. Si D est de la forme (4.1),

$$\iint_D f(x, y) \, dx \, dy = \int_a^b \left(\int_{\phi_1(x)}^{\phi_2(x)} f(x, y) \, dy \right) dx$$

2. Si D est de la forme (4.2),

$$\iint_D f(x, y) \, dx \, dy = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x, y) \, dx \right) dy$$

Cette définition est cohérente en ce sens que si un domaine D peut s'écrire à la fois sous les formes (4.1) et (4.2), alors nous avons l'égalité :

$$\int_a^b \left(\int_{\phi_1(x)}^{\phi_2(x)} f(x, y) \, dy \right) dx = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x, y) \, dx \right) dy$$

L'intégrale double ainsi définie possède les propriétés de linéarité, positivité, croissance et satisfait les inégalités triangulaire et de Cauchy-Schwarz.

4.2 Intégrale double sur un produit d'intervalles quelconques

Dans toute la suite, I et I' désignent des intervalles réels non vides ni réduits à un point.

4.2.1 Le cas positif

Définition 4.2.1 Soit $f \in C(I \times I', \mathbb{R}_+)$. On appelle intégrale double de f sur $I \times I'$ l'élément de \mathbb{R}_+ :

$$\iint_{I \times I'} f = \sup \left\{ \iint_{[a,b] \times [c,d]} f(x,y) \, dx \, dy; [a,b] \subset I, [c,d] \subset I' \right\} \leq +\infty$$

On dit que f est intégrable sur $I \times I'$ si $\iint_{I \times I'} f < +\infty$.

Remarque. Une application $f \in C(I \times I', \mathbb{R}_+)$ est donc intégrable s'il existe $M > 0$ tel que pour tous segments $[a,b] \subset I$ et $[c,d] \subset I'$,

$$\iint_{[a,b] \times [c,d]} f(x,y) \, dx \, dy \leq M$$

Théorème 4.2.2 Soit $f \in C(I \times I', \mathbb{R}_+)$. On suppose que, pour tout $y \in I'$, l'application $x \mapsto f(x,y)$ est intégrable sur I , ce qui nous permet de définir l'application $y \mapsto \int_I f(x,y) \, dx$, que l'on suppose intégrable sur I' . Alors, f est intégrable sur $I \times I'$ et on a l'égalité :

$$\iint_{I \times I'} f = \int_{I'} \left(\int_I f(x,y) \, dx \right) dy$$

Symétriquement, nous avons aussi le résultat suivant.

Soit $f \in C(I \times I', \mathbb{R}_+)$. On suppose que, pour tout $x \in I$, l'application $y \mapsto f(x,y)$ est intégrable sur I' , ce qui nous permet de définir l'application $x \mapsto \int_{I'} f(x,y) \, dy$, que l'on suppose intégrable sur I .

Alors, f est intégrable sur $I \times I'$ et on a l'égalité :

$$\iint_{I \times I'} f = \int_I \left(\int_{I'} f(x,y) \, dy \right) dx$$

Voici ce qu'il est important de retenir en substance :

Si toutes les conditions précédentes sont satisfaites, c'est-à-dire si l'on peut légitimement définir les deux intégrales suivantes :

$$\int_{I'} \left(\int_I f(x,y) \, dx \right) dy \text{ et } \int_I \left(\int_{I'} f(x,y) \, dy \right) dx$$

alors elles sont égales. L'intégrale double correspondante est encore égale à leur valeur commune. C'est le plus souvent sous cette forme que l'on utilisera ce théorème, dit de *Fubini-Tonelli*.

4.2.2 Le cas réel ou complexe

On adopte exactement la même démarche que celle utilisée dans le chapitre précédent pour définir l'intégrale d'une application sur un intervalle quelconque en partant du cas positif.

Définition et proposition 4.2.3 Une application $f \in C(I \times I', \mathbb{R})$ est dite intégrable si l'application positive $|f|$ est intégrable sur $I \times I'$. Ceci équivaut à l'intégrabilité des deux applications positives f^+ et f^- sur $I \times I'$.

Dans ce cas, nous appellerons intégrale de l'application f sur $I \times I'$ le réel :

$$\int_{I \times I'} f = \int_{I \times I'} f^+ - \int_{I \times I'} f^-$$

Définition et proposition 4.2.4 Une application $f \in C(I \times I', \mathbb{C})$ est dite intégrable si l'application positive $|f|$ est intégrable sur $I \times I'$. Ceci équivaut à l'intégrabilité des deux applications réelles $\Re f$ et $\Im f$ sur $I \times I'$.

Dans ce cas, nous appellerons intégrale de l'application f sur $I \times I'$ le nombre complexe suivant :

$$\int_{I \times I'} f = \int_{I \times I'} \Re f + i \int_{I \times I'} \Im f$$

L'intégrale double ainsi définie est linéaire et satisfait l'inégalité triangulaire. Dans le cas $\mathbb{K} = \mathbb{R}$, elle possède les propriétés de positivité et croissance et satisfait l'inégalité de Cauchy-Schwarz.

Théorème 4.2.5 Soit $f \in C(I \times I', \mathbb{K})$. Supposons que, pour tout $y \in I'$, l'application $x \mapsto f(x, y)$ est intégrable sur I , ce qui nous permet de définir l'application positive $y \mapsto \int_I |f(x, y)| dx$, que nous supposons intégrable sur I' . Alors f est intégrable sur $I \times I'$.

Si nous supposons en outre que l'application $y \mapsto \int_I f(x, y) dx$ est continue par morceaux sur I' , alors elle est intégrable sur I' et nous avons l'égalité :

$$\iint_{I \times I'} f = \int_{I'} \left(\int_I f(x, y) dx \right) dy$$

Symétriquement, nous avons aussi le résultat suivant :

Soit $f \in C(I \times I', \mathbb{K})$. On suppose que, pour tout $x \in I$, l'application $y \mapsto f(x, y)$ est intégrable sur I' , ce qui nous permet de définir l'application positive $x \mapsto \int_{I'} |f(x, y)| dy$, que l'on suppose intégrable sur I . Alors f est intégrable sur $I \times I'$.

4.2. INTÉGRALE DOUBLE SUR UN PRODUIT D'INTERVALLES QUELCONQUES 97

Si, en outre, l'application $x \mapsto \int_{I'} f(x, y) dy$ est continue par morceaux sur I , alors elle est intégrable sur I et on a l'égalité :

$$\iint_{I \times I'} f = \int_I \left(\int_{I'} f(x, y) dy \right) dx$$

Voici ce qu'il est important de retenir en substance :

Si l'une des deux intégrales positives suivantes est finie

$$\int_{I'} \left(\int_I |f(x, y)| dx \right) dy \text{ ou } \int_I \left(\int_{I'} |f(x, y)| dy \right) dx$$

alors il y a égalité entre les deux intégrales suivantes

$$\int_{I'} \left(\int_I f(x, y) dx \right) dy \text{ et } \int_I \left(\int_{I'} f(x, y) dy \right) dx$$

sous réserve des conditions de continuité par morceaux assurant que ces intégrales sont bien définies. L'intégrale double correspondante est alors égale à cette valeur commune. C'est le théorème de *Fubini(-Lebesgue)*.

Remarque importante. Il faut vérifier d'abord la condition d'intégrabilité sur $|f|$, sans quoi la dernière égalité peut tomber en défaut.

Par exemple, nous avons l'égalité :

$$\int_1^{+\infty} \left(\int_1^{+\infty} \frac{x-y}{(x+y)^3} dx \right) dy = \frac{1}{2} \quad (4.3)$$

En effet,

$$\int_1^{+\infty} \frac{x+y-2y}{(x+y)^3} dx = \int_1^{+\infty} \frac{dx}{(x+y)^2} - y \int_1^{+\infty} \frac{2}{(x+y)^3} dx ,$$

ce qui s'écrit encore, en passant aux primitives :

$$\left[\frac{-1}{x+y} \right]_{x=1}^{x \rightarrow +\infty} - y \left[\frac{-1}{(x+y)^2} \right]_{x=1}^{x \rightarrow +\infty} = \frac{1}{1+y} - \frac{y}{(1+y)^2} = \frac{1}{(1+y)^2}$$

L'intégrale qui nous intéresse vaut donc :

$$\int_1^{+\infty} \frac{dy}{(1+y)^2} = \left[\frac{-1}{1+y} \right]_{y=1}^{y \rightarrow +\infty} = \frac{1}{2}$$

L'égalité (4.3) étant maintenant prouvée, nous pouvons y intervertir les variables (muettes) x et y pour obtenir :

$$\int_1^{+\infty} \left(\int_1^{+\infty} \frac{x-y}{(x+y)^3} dy \right) dx = -\frac{1}{2} \neq \int_1^{+\infty} \left(\int_1^{+\infty} \frac{x-y}{(x+y)^3} dx \right) dy$$

Nous en concluons que l'application f définie par $f(x, y) := \frac{x-y}{(x+y)^3}$ n'est pas intégrable sur $[1, +\infty]^2$.

4.3 Changement de variable

Nous admettons le théorème suivant :

Théorème 4.3.1 *Soient U et V deux ouverts de \mathbb{R}^2 , $\Phi : U \rightarrow V$ un C^1 -difféomorphisme et $f : V \rightarrow \mathbb{R}$ une application positive et continue. Alors nous avons l'égalité suivante dans \mathbb{R}_+ :*

$$\iint_V f(x, y) \, dx \, dy = \iint_U f(\Phi(s, t)) |J_\Phi(s, t)| \, ds \, dt \quad ,$$

où J_Φ est le déterminant jacobien du C^1 -difféomorphisme Φ .

Ce théorème se démontre grâce à la théorie de la mesure et c'est dans le cadre de celle-ci que l'on peut définir les deux intégrales ci-dessus pour des ouverts quelconques U et V de \mathbb{R}^2 . Dans le cadre de notre programme, nous ne pourrons l'appliquer que lorsque ces ouverts seront de l'une des deux formes suivantes :

- un produit de deux intervalles ouverts,
- l'intérieur d'un domaine simple du plan.

Dans ce dernier cas, par exemple si $V = \overset{\circ}{D}$ avec D domaine simple du plan, nous admettrons l'égalité :

$$\iint_V f(x, y) \, dx \, dy = \iint_D f(x, y) \, dx \, dy \quad ,$$

correspondant à l'idée intuitive que la frontière ∂D étant de «surface nulle», elle n'apporte aucune contribution à cette dernière intégrale double.

Nous admettons encore le théorème suivant.

Théorème 4.3.2 *Soient U et V deux ouverts de \mathbb{R}^2 , $\Phi : U \rightarrow V$ un C^1 -difféomorphisme et $f : V \rightarrow \mathbb{R}$ une application continue telle que :*

$$\iint_V |f(x, y)| \, dx \, dy < +\infty \tag{4.4}$$

Alors, nous avons l'égalité suivante dans \mathbb{R} :

$$\iint_V f(x, y) \, dx \, dy = \iint_U f(\Phi(s, t)) |J_\Phi(s, t)| \, ds \, dt.$$

Remarquons que la condition d'intégrabilité (4.4) assure que l'intégrale du membre de droite est bien définie. En effet, en appliquant le théorème 4.3.1 avec $|f|$, nous obtenons :

$$\iint_U |f(\Phi(s, t))| |J_\Phi(s, t)| \, ds \, dt = \iint_V |f(x, y)| \, dx \, dy < +\infty.$$

Applications. Le lecteur trouvera des exemples d'utilisation d'intégrales multiples pour calculer des « grandeurs » (aire, volume, moment d'inertie ...) dans les sections 3 à 5 du chapitre 6.

Chapitre 5

Méthodes de calcul approché d'une intégrale. Majoration de l'erreur.

Références.

- Chapitre 24 du livre de DANTZER.
- Chapitre 3 de l'ouvrage *Analyse numérique et équations différentielles* par Jean-Pierre DEMAILLY, collection Grenoble Sciences, Presses Universitaires de Grenoble.
- [SKA 31-36] pour l'utilisation de polynômes orthogonaux.

Notation. Pour tout $p \in \mathbb{N}$, nous noterons $\mathbb{R}_p[X]$ l'espace vectoriel des polynômes à coefficients réels, de degré inférieur ou égal à p .

Introduction

Dans les trois premières sections de ce chapitre, nous considérerons un intervalle réel d'intérieur non vide I et une application $f \in C(I, \mathbb{R})$. Étant donné un segment $[a, b] \subset I$ et x_0, x_1, \dots, x_p des points distincts de $[a, b]$, avec $p \in \mathbb{N}$, nous chercherons une formule d'approximation de la forme :

$$\int_a^b f(t) dt \simeq (b-a) \sum_{j=0}^p \omega_j f(x_j),$$

où les $\omega_j, 0 \leq j \leq p$ sont des réels tels que $\sum_{j=0}^p \omega_j = 1$, si bien que la formule précédente est une égalité dans le cas où f est constante sur $[a, b]$.

Une telle formule sera appelée *méthode de quadrature* et nous dirons que l'entier p est le *rang* de cette méthode.

Dans les trois premières sections de ce chapitre, nous nous laisserons guider par quelques idées simples :

Idée 1. Soit p un entier fixé. Dans le cas particulier où $f = P \in \mathbb{R}_p[X]$, nous disposons de formules simples de calcul exact de $\int_a^b P$.

Idée 2. Nous pouvons approcher une fonction suffisamment régulière f par un polynôme P et utiliser alors — c'est ce qu'on appelle une *méthode élémentaire* — l'approximation $\int_a^b f \simeq \int_a^b P$.
Il sera bien sûr essentiel de contrôler l'erreur commise lors de cette approximation.

Idée 3. Par relation de Chasles, nous pouvons découper l'intervalle $[a, b]$ en « petits » sous-intervalles et appliquer sur chacun d'eux la méthode élémentaire. Nous obtenons ainsi une *méthode composée* donnant une approximation de $\int_a^b f$ plus précise que la méthode élémentaire.

Dans la quatrième section, nous chercherons à approcher des intégrales d'une forme plus générale, éventuellement impropres, en utilisant les propriétés des polynômes orthogonaux.

Enfin, une brève cinquième section mentionnera une méthode probabiliste de calcul approché d'une intégrale.

5.1 Intégrale d'une fonction polynomiale.

5.1.1 Rang et ordre d'une méthode de quadrature

Soit $[a, b]$ un segment réel fixé.

Définition 5.1.1 Soit $p \in \mathbb{N}$. On appelle système interpolateur de rang p tout $(p+1)$ -uplet (x_0, x_1, \dots, x_p) constitué d'éléments **distincts** de $[a, b]$.

Un tel système interpolateur étant donné, considérons, pour tout $j \in \llbracket 0; p \rrbracket$, la forme linéaire $\delta_{x_j} : \mathbb{R}_p[X] \rightarrow \mathbb{R}$ définie par $\delta_{x_j}(P) = P(x_j)$.

Proposition 5.1.2 Soit (x_0, x_1, \dots, x_p) un système interpolateur de rang p . Alors, le $(p+1)$ -uplet $(\delta_{x_0}, \dots, \delta_{x_p})$ forme une base de l'espace dual $(\mathbb{R}_p[X])^*$.

Démonstration: cf. [DAN 488]. □

Puisque $P \mapsto \frac{1}{b-a} \int_a^b P(t) dt$ est une forme linéaire sur $\mathbb{R}_p[X]$, on en déduit immédiatement le résultat suivant.

Corollaire 5.1.3 *Soit (x_0, x_1, \dots, x_p) un système interpolateur de rang p . Alors, il existe un unique $(\omega_0, \dots, \omega_p) \in \mathbb{R}^{p+1}$, appelé $(p+1)$ -uplet de coefficients associé au système interpolateur (x_0, x_1, \dots, x_p) , tel que :*

$$\forall P \in \mathbb{R}_p[X], \quad \int_a^b P(t) dt = (b-a) \sum_{j=0}^p \omega_j P(x_j).$$

Remarques.

1. Il suffit de prendre $P \equiv 1$ pour constater que la condition $\sum_{j=0}^p \omega_j = 1$ est alors bien vérifiée.
2. En fait, les résultats précédents subsistent même si les réels distincts x_0, x_1, \dots, x_p ne sont pas tous éléments de $[a, b]$. Néanmoins, notre but étant d'obtenir une méthode de quadrature pour toute application f suffisamment régulière (pas nécessairement polynomiale!), il est naturel de faire cette hypothèse supplémentaire.

Définition 5.1.4 *Soit (x_0, x_1, \dots, x_p) un système interpolateur de rang p . Une méthode de quadrature de la forme*

$$\int_a^b f(t) dt \simeq (b-a) \sum_{j=0}^p \omega_j f(x_j), \quad \text{avec} \quad \sum_{j=0}^p \omega_j = 1, \quad (5.1)$$

est dite aussi de rang p . Une telle méthode est dite d'ordre q (avec $q \in \mathbb{N}$) si la formule approchée est exacte pour toute application polynomiale $f \in \mathbb{R}_q[X]$ et inexacte pour au moins une application $f \in \mathbb{R}_{q+1}[X]$.

Si $(\omega_0, \dots, \omega_p)$ est le $(p+1)$ -uplet de coefficients associé au système interpolateur, le corollaire précédent nous dit que la méthode de quadrature (5.1) est d'ordre $q \geq p$. Nous verrons plus loin sur des exemples qu'il est possible d'avoir $q > p$, ce qui est particulièrement intéressant.

Une question essentielle dans la pratique est bien sûr de déterminer explicitement les coefficients $\omega_0, \dots, \omega_p$ associés au système interpolateur. Dans ce but, nous introduisons la définition suivante.

Définition 5.1.5 *Soit (x_0, x_1, \dots, x_p) un système interpolateur de rang p . On appelle polynômes élémentaires de Lagrange relatifs à ce système interpolateur les polynômes $l_j \in \mathbb{R}_p[X]$, $0 \leq j \leq p$, définis par :*

$$\forall j \in \llbracket 0; p \rrbracket, \quad l_j(X) = \prod_{\substack{0 \leq i \leq p \\ i \neq j}} \frac{X - x_i}{x_j - x_i}. \quad (5.2)$$

Le lecteur constatera que, pour tout $j \in \llbracket 0; p \rrbracket$, l_j est l'unique polynôme de $\mathbb{R}_p[X]$ vérifiant (avec $\delta_{i,j}$ symbole de Kronecker) :

$$\forall i \in \llbracket 0; p \rrbracket, \quad l_j(x_i) = \delta_{i,j}.$$

On déduit immédiatement du Corollaire 5.1.3 le résultat suivant.

Proposition 5.1.6 *Soit (x_0, x_1, \dots, x_p) un système interpolateur de rang p . Les coefficients $(\omega_0, \dots, \omega_p)$ associés à ce système interpolateur sont donnés par les égalités :*

$$\forall j \in \llbracket 0; p \rrbracket, \quad \omega_j = \frac{1}{b-a} \int_a^b l_j(t) dt. \quad (5.3)$$

Un simple changement de variable nous permet d'en déduire le lemme suivant, utile pour se ramener à un segment fixé (par exemple $[0, 1]$ ou $[-1, 1]$) sans perdre de généralité.

Lemme 5.1.7 *Soit (x_0, x_1, \dots, x_p) un système interpolateur de rang p relatif à l'intervalle $[a, b]$, soit $(\omega_0, \dots, \omega_p)$ le $(p+1)$ -uplet de coefficients associé et soit enfin $A : \mathbb{R} \rightarrow \mathbb{R}$ une transformation affine strictement croissante.*

Alors, $(A(x_0), A(x_1), \dots, A(x_p))$ est un système interpolateur relatif à l'intervalle $[A(a), A(b)]$ et le $(p+1)$ -uplet de coefficients qui lui est associé est encore $(\omega_0, \dots, \omega_p)$. En outre, les deux méthodes de quadrature

$$\int_a^b f(t) dt \simeq (b-a) \sum_{j=0}^p \omega_j f(x_j) \text{ et } \int_{A(a)}^{A(b)} f(t) dt \simeq (A(b)-A(a)) \sum_{j=0}^p \omega_j f(A(x_j))$$

admettent le même ordre.

5.1.2 Exemples

Méthodes des rectangles ($p = 0$, $q = 0$)

La *méthode des rectangles à gauche* correspond au cas $p = 0$ et $x_0 = a$.

On a évidemment $\omega_0 = 1$ de sorte que la formule s'écrit :

$$\int_a^b f(t) dt \simeq (b-a) f(a).$$

Symétriquement, la *méthode des rectangles à droite* correspond au cas $p = 0$ et $x_0 = b$, avec la formule :

$$\int_a^b f(t) dt \simeq (b-a) f(b).$$

On vérifie immédiatement que ces méthodes sont d'ordre $q = p = 0$ (prendre $a = 0$, $b = 1$ et $f(x) = x$).

Méthode du point milieu ($p = 0, q = 1$)

Elle correspond au cas $p = 0$ et $x_0 = \frac{a+b}{2}$. On a bien sûr toujours $\omega_0 = 1$ et la formule s'écrit donc :

$$\int_a^b f(t) dt \simeq (b-a) f\left(\frac{a+b}{2}\right).$$

En remarquant que, pour $f(t) = t$, nous avons :

$$\int_a^b t dt = \frac{b^2 - a^2}{2} = (b-a) f\left(\frac{a+b}{2}\right),$$

nous en déduisons par linéarité que la méthode du point milieu est d'ordre $q \geq 1 > p = 0$. En fait, le lecteur vérifiera facilement que $q = 1$.

Méthode des trapèzes ($p = 1, q = 1$)

Elle correspond au cas $p = 1, x_0 = a$ et $x_1 = b$.

Pour calculer ω_1 , nous pouvons nous ramener au cas $a = 0$ et $b = 1$ en utilisant le Lemme 5.1.7, avec $\varphi(x) := a + (b-a)x$. On a alors $l_1(X) = X$ d'où, d'après (5.3), $\omega_1 = 1/2$ et donc $\omega_0 = \omega_1 = 1/2$. Finalement, nous obtenons la formule de quadrature :

$$\int_a^b f(t) dt \simeq (b-a) \frac{f(a) + f(b)}{2}.$$

Nous laissons le soin au lecteur de vérifier que cette méthode est d'ordre $q = p = 1$.

En fait, la méthode des trapèzes est un cas particulier des méthodes à subdivision régulière que nous traitons maintenant.

5.1.3 Méthodes de Newton-Cotes (subdivision régulière)

On parle de méthode de Newton-Cotes lorsque le système interpolateur forme une subdivision régulière de l'intervalle $[a, b]$. Autrement dit, dans la méthode de Newton-Cotes de rang p (en abrégé NC_p), avec $p \in \mathbb{N}^*$, on choisit le système interpolateur (x_0, \dots, x_p) comme suit.

$$\forall j \in \llbracket 0; p \rrbracket, \quad x_j = a + (b-a)j/p.$$

Le calcul des coefficients associés à ce système interpolateur a l'avantage d'être facilité par la propriété de symétrie suivante.

Proposition 5.1.8 *Dans la méthode NC_p , nous avons les égalités :*

$$\forall j \in \llbracket 0; p \rrbracket, \quad \omega_j = \omega_{p-j}.$$

Démonstration: D'après le Lemme 5.1.7 avec $\varphi(x) := \frac{a+b}{2} + \frac{b-a}{2}x$, nous pouvons nous ramener au cas $a = -1$ et $b = 1$ sans perte de généralité.

Notre système interpolateur est alors défini par les points

$$x_j = -1 + 2j/p, \quad 0 \leq j \leq p,$$

et l_j est l'unique polynôme de $\mathbb{R}_p[X]$ vérifiant :

$$\forall i \in \llbracket 0; p \rrbracket, \quad l_j(-1 + 2i/p) = \delta_{i,j}. \quad (5.4)$$

Par un simple changement d'indice, ceci s'écrit encore :

$$\forall i \in \llbracket 0; p \rrbracket, \quad l_j(-1 + 2(p-i)/p) = \delta_{p-i,j},$$

ce que nous réécrivons sous la forme :

$$\forall i \in \llbracket 0; p \rrbracket, \quad l_j(1 - 2i/p) = \delta_{i,p-j}.$$

Définissant le polynôme $\check{l}_j \in \mathbb{R}_p[X]$ par $\check{l}_j(x) := l_j(-x)$, nous avons donc :

$$\forall i \in \llbracket 0; p \rrbracket, \quad \check{l}_j(-1 + 2i/p) = \delta_{i,p-j},$$

ce qui prouve, d'après (5.4), l'égalité $\check{l}_j = l_{p-j}$. Revenant aux relations (5.3), nous en déduisons par un simple changement de variable :

$$\omega_j = \frac{1}{2} \int_{-1}^1 l_j(t) dt = \frac{1}{2} \int_{-1}^1 \check{l}_j(t) dt = \frac{1}{2} \int_{-1}^1 l_{p-j}(t) dt = \omega_{p-j}. \quad \square$$

Nous avons déjà rencontré la méthode NC_1 , qui n'est autre que la méthode des trapèzes. Notons que grâce à la proposition précédente, nous retrouvons immédiatement $\omega_0 = \omega_1 = 1/2$.

Voyons maintenant deux autres exemples de méthodes de Newton-Cotes.

Méthode de Simpson (NC_2)

D'après la proposition précédente, nous avons $\omega_0 = \omega_2$. En se ramenant à $a = -1$ et $b = 1$ par le Lemme 5.1.7, nous calculons $l_1(X) = 1 - X^2$ d'où $\omega_1 = \frac{1}{2} \int_{-1}^1 (1 - t^2) dt = 2/3$ et $\omega_0 = \omega_2 = 1/6$.

Finalement, nous obtenons la formule de quadrature :

$$\int_a^b f(t) dt \simeq \frac{b-a}{6} \left(f(a) + f(b) + 4f\left(\frac{a+b}{2}\right) \right).$$

Remarque. Dans un exercice corrigé, [GOU 79-81] retrouve la méthode de Simpson « à la main » en appliquant le théorème des valeurs intermédiaires et le théorème des accroissements finis.

Méthode de Boole-Villarceau (NC_4)

On se ramène à $a = -1$ et $b = 1$ par le Lemme 5.1.7, si bien que le système interpolateur est défini par : $x_0 = -1, x_1 = -\frac{1}{2}, x_2 = 0, x_3 = \frac{1}{2}, x_4 = 1$.

Nous calculons alors :

$$l_2(X) = \frac{(X+1)(X+\frac{1}{2})(X-\frac{1}{2})(X-1)}{1 \cdot \frac{1}{2} \cdot (-\frac{1}{2}) \cdot (-1)} = (X^2-1)(4X^2-1),$$

$$\text{d'où } \omega_2 = \frac{1}{2} \int_{-1}^1 (4t^4 - 5t^2 + 1) dt = 2/15;$$

$$l_1(X) = \frac{(X+1)X(X-\frac{1}{2})(X-1)}{\frac{1}{2} \cdot (-\frac{1}{2}) \cdot (-1) \cdot (-\frac{3}{2})} = -\frac{8}{3}X(X^2-1)(X-\frac{1}{2})$$

$$\text{d'où, avec des considérations de parité, } \omega_1 = -\frac{8}{3} \int_0^1 (t^4 - t^2) dt = \frac{16}{45},$$

$$\omega_3 = \omega_1 = \frac{16}{45} \quad \text{et} \quad \omega_0 = \omega_4 = \frac{1}{2} \left(1 - \frac{2}{15} - \frac{32}{45}\right) = \frac{7}{90}.$$

La formule de quadrature par la méthode de Boole-Villarceau s'écrit donc :

$$\int_a^b f \simeq \frac{b-a}{90} \left(7f(a) + 7f(b) + 32f\left(\frac{3a+b}{4}\right) + 32f\left(\frac{a+3b}{4}\right) + 12f\left(\frac{a+b}{2}\right) \right).$$

À part dans le cas $p = 1$ (méthode des trapèzes), on n'utilise les méthodes de Newton-Cotes que lorsque p est pair, pour la raison suivante.

Proposition 5.1.9 *Soit $p \in \mathbb{N}^*$. Si p est pair, alors l'ordre de NC_p est $(p+1)$. Si p est impair, alors l'ordre de NC_p est p .*

Démonstration: D'après le Lemme 5.1.7, on peut se ramener au cas $a = -1$ et $b = 1$ sans perdre de généralité. On a donc $x_j = -1 + 2j/p$, si bien que $x_{p-j} = -x_j$ pour tout $j \in \llbracket 0; p \rrbracket$.

Nous nous contenterons de montrer que, si p est pair, alors l'ordre de NC_p est au moins $(p+1)$, en admettant le reste de la proposition.

Si $p = 2k$, nous avons $x_k = 0$ d'après l'égalité précédente. En outre, nous savons déjà que l'ordre de NC_{2k} est au moins $2k$, c'est-à-dire que :

$$\forall P \in \mathbb{R}_{2k}[X], \quad \int_{-1}^1 P(t) dt = 2 \sum_{j=0}^{2k} \omega_j P(x_j). \quad (5.5)$$

Si nous démontrons que cette égalité reste vraie pour $P_0(X) := X^{2k+1}$, alors par linéarité elle sera aussi satisfaite pour tout polynôme de $\mathbb{R}_{2k+1}[X]$, ce qui

nous permettra de conclure.

Or, le membre de gauche de (5.5) avec $P_0(t) = t^{2k+1}$ est nul.

Le membre de droite peut être réécrit, avec un changement d'indice :

$$2\omega_k P_0(x_k) + 2 \sum_{j=0}^{k-1} (\omega_j P_0(x_j) + \omega_{2k-j} P_0(x_{2k-j})).$$

Nous concluons grâce à la Proposition 5.1.8 et aux égalités :

$$P_0(x_{2k-j}) = P_0(-x_j) = -P_0(x_j) \quad ; \quad P_0(x_k) = P_0(0) = 0.$$

□

5.2 Interpolation polynomiale.

On appelle interpolation polynomiale un procédé permettant d'approcher (dans un sens à préciser) une fonction suffisamment régulière $f : I \rightarrow \mathbb{R}$, où I est un intervalle réel, par une fonction polynomiale.

Comme précédemment, nous fixons donc un intervalle réel I , un segment $[a, b] \subset I$ et nous considérons un système interpolateur (x_0, x_1, \dots, x_p) constitué de $(p+1)$ points du segment $[a, b]$. En outre, nous introduisons désormais un $(p+1)$ -uplet d'entiers non nuls $(\alpha_0, \dots, \alpha_p) \in (\mathbb{N}^*)^{p+1}$ et nous posons :

$$n = \sum_{j=0}^p \alpha_j.$$

Nous avons donc toujours $n \geq p+1$.

5.2.1 Interpolation de Lagrange et Hermite

Notre résultat fondamental est le suivant (cf. [DAN 486]), dans lequel nous utilisons la convention d'écriture $f^{(0)}(x) = f(x)$.

Proposition 5.2.1 *Nous supposons que, pour tout $j \in \llbracket 0; p \rrbracket$, l'application $f : I \rightarrow \mathbb{R}$ admet au point x_j une dérivée d'ordre $(\alpha_j - 1)$. Alors, il existe un unique polynôme $H \in \mathbb{R}_{n-1}[X]$ tel que :*

$$\forall j \in \llbracket 0; p \rrbracket, \quad \forall k \in \llbracket 0; \alpha_j - 1 \rrbracket, \quad H^{(k)}(x_j) = f^{(k)}(x_j).$$

On dit que H est le polynôme d'interpolation d'Hermite de f relativement au système interpolateur (x_0, \dots, x_p) et aux entiers $(\alpha_0, \dots, \alpha_p)$.

Notons qu'il y a en tout $\sum_{j=0}^p \alpha_j = n$ conditions sur le polynôme H , ce qui correspond à la dimension de l'espace vectoriel $\mathbb{R}_{n-1}[X]$.

Remarque. Bien qu'il n'y ait aucune condition sur la fonction f au voisinage de x_j lorsque $\alpha_j = 1$, des conditions de régularité vont apparaître dès que nous voudrions contrôler l'erreur commise dans l'approximation $\int_a^b f \simeq \int_a^b H$, comme nous le verrons dans la suite.

Dans le cas particulier $\alpha_0 = \alpha_1 = \dots = \alpha_p = 1$ (et donc $n = p + 1$), nous obtenons le *polynôme d'interpolation de Lagrange*, noté $L \in \mathbb{R}_p[X]$. Nous pouvons l'explicitier simplement grâce aux polynômes élémentaires de Lagrange introduits en (5.2) :

$$L = \sum_{j=0}^p f(x_j) l_j.$$

5.2.2 Évaluation de l'erreur

Pour poursuivre notre démarche, nous aurons besoin de la définition suivante.

Définition 5.2.2 On appelle *noyau d'interpolation associé au système interpolateur* (x_0, x_1, \dots, x_p) et aux entiers $(\alpha_0, \dots, \alpha_p) \in (\mathbb{N}^*)^{p+1}$ le *polynôme* $N \in \mathbb{R}_n[X]$ défini par :

$$N(X) = \prod_{j=0}^p (X - x_j)^{\alpha_j}.$$

Avant d'établir la proposition qui nous permettra de contrôler l'erreur commise dans l'approximation $\int_a^b f \simeq \int_a^b H$, rappelons un théorème d'analyse.

Théorème 5.2.3 (Second théorème de Rolle) Supposons que $f : I \rightarrow \mathbb{R}$ est n fois dérivable et admet au moins $(n + 1)$ zéros, en comptant les ordres de multiplicité. Alors, l'application $f^{(n)}$ admet au moins un zéro sur I .

Proposition 5.2.4 Soit $f \in C^n(I, \mathbb{R})$. Soient $H \in \mathbb{R}_{n-1}[X]$ le polynôme d'interpolation d'Hermite de f , resp. $N \in \mathbb{R}_n[X]$ le noyau interpolateur, associés au système interpolateur (x_0, \dots, x_p) et aux entiers $(\alpha_0, \dots, \alpha_p) \in (\mathbb{N}^*)^{p+1}$. Alors, nous avons :

$$\forall t \in I, \quad \exists y \in I, \quad f(t) - H(t) = \frac{1}{n!} N(t) f^{(n)}(y).$$

Démonstration: Nous remarquons d'abord que la propriété précédente est immédiate dans le cas où t est un élément du système interpolateur (x_0, \dots, x_p) car on a alors $f(t) = H(t)$ et $N(t) = 0$, donc tout $y \in I$ convient. Considérons désormais $t \in I \setminus \{x_0, \dots, x_p\}$ arbitraire fixé. Comme $N(t) \neq 0$, nous pouvons choisir $\lambda \in \mathbb{R}$ tel que l'application $g : I \rightarrow \mathbb{R}$ définie par :

$$\forall x \in I, \quad g(x) = f(x) - H(x) - \lambda N(x), \quad (5.6)$$

admette t pour zéro. Puisque $f \in C^n(I, \mathbb{R})$, nous avons $g \in C^n(I, \mathbb{R})$ et nous constatons que g admet $1 + \sum_{j=0}^p \alpha_j = (n+1)$ zéros, en tenant compte des ordres de multiplicité.

Le théorème précédent nous assure l'existence de $y \in I$ tel que $g^{(n)}(y) = 0$. Comme $H \in \mathbb{R}_{n-1}[X]$, on a $H^{(n)} \equiv 0$ et, le terme dominant du polynôme N étant X^n , $N^{(n)} \equiv n!$, d'où $g^{(n)}(y) = f^{(n)}(y) - \lambda n! = 0$.

Nous avons donc $\lambda = f^{(n)}(y)/(n!)$, si bien que (5.6) se réécrit :

$$\forall x \in I, \quad g(x) = f(x) - H(x) - \frac{f^{(n)}(y)}{n!} N(x).$$

Il reste à écrire $g(t) = 0$ pour conclure. □

Remarque. La valeur de y dépend de celle de $t \in I$.

Cette proposition nous donne immédiatement une majoration de l'erreur commise dans l'approximation $\int_a^b f \simeq \int_a^b H$, sous l'hypothèse $f \in C^n([a, b], \mathbb{R})$.

Corollaire 5.2.5 Avec les notations de la proposition précédente, dans laquelle nous prenons $I = [a, b]$, nous avons la majoration :

$$\left| \int_a^b f(t) dt - \int_a^b H(t) dt \right| \leq \frac{1}{n!} \sup_{t \in [a, b]} |f^{(n)}(t)| \int_a^b |N(t)| dt.$$

Pour calculer cette dernière intégrale, on peut se ramener au segment $[-1, 1]$, comme le montre le lemme suivant, dans lequel nous reprenons les notations du Lemme 5.1.7.

Lemme 5.2.6 Supposons que $A : x \mapsto cx + d$, avec $c > 0$ et $d \in \mathbb{R}$. Soit $N \in \mathbb{R}_n[X]$, resp. $N_1 \in \mathbb{R}_n[X]$, le noyau interpolateur associé au système interpolateur (x_0, \dots, x_p) , resp. $(A(x_0), A(x_1), \dots, A(x_p))$, et aux entiers $(\alpha_0, \dots, \alpha_p) \in (\mathbb{N}^*)^{p+1}$. Alors,

$$\int_{A(a)}^{A(b)} |N_1(t)| dt = c^{n+1} \int_a^b |N(t)| dt.$$

Démonstration: Par définition, on a :

$$N(X) = \prod_{j=0}^p (X - x_j)^{\alpha_j} \quad ; \quad N_1(X) = \prod_{j=0}^p (X - A(x_j))^{\alpha_j}.$$

Nous calculons donc, par le changement de variable $u = A^{-1}(t) \Leftrightarrow t = A(u)$,

$$\int_{A(a)}^{A(b)} \prod_{j=0}^p |t - A(x_j)|^{\alpha_j} dt = \int_a^b \prod_{j=0}^p |A(u) - A(x_j)|^{\alpha_j} c du,$$

d'où :

$$\int_{A(a)}^{A(b)} \prod_{j=0}^p |t - A(x_j)|^{\alpha_j} dt = c^{n+1} \int_a^b \prod_{j=0}^p |u - x_j|^{\alpha_j} du,$$

ce qui nous permet de conclure. \square

Par exemple, pour passer de l'intervalle $[a, b]$ à l'intervalle $[-1, 1]$, on choisit $c = 2/(b - a)$ et $d = -(a + b)/(b - a)$ et le lemme précédent nous donne :

$$\int_a^b |N(t)| dt = \frac{(b - a)^{n+1}}{2^{n+1}} \int_{-1}^1 |N_1(t)| dt, \quad \text{où } n = \sum_{j=0}^p \alpha_j. \quad (5.7)$$

5.3 Intégration numérique

5.3.1 Avec un polynôme de Lagrange

Soit $p \in \mathbb{N}$, soit (x_0, x_1, \dots, x_p) un système interpolateur constitué de points du segment $[a, b]$, et soit $(\omega_0, \dots, \omega_p) \in \mathbb{R}^{p+1}$ le $(p + 1)$ -uplet de coefficients qui est associé à ce système interpolateur de rang p (cf. Corollaire 5.1.3).

Considérons $f : [a, b] \rightarrow \mathbb{R}$ et notons $L \in \mathbb{R}_p[X]$ le polynôme d'interpolation de Lagrange de f associé au système interpolateur (x_0, x_1, \dots, x_p) .

Par définition des coefficients $(\omega_0, \dots, \omega_p)$, comme $L \in \mathbb{R}_p[X]$, on a :

$$\int_a^b L = (b - a) \sum_{j=0}^p \omega_j L(x_j) = (b - a) \sum_{j=0}^p \omega_j f(x_j).$$

Ainsi, l'approximation $\int_a^b f \simeq \int_a^b L$ se réécrit :

$$\int_a^b f \simeq (b - a) \sum_{j=0}^p \omega_j f(x_j). \quad (5.8)$$

Remarque. Il est inutile de calculer explicitement le polynôme de Lagrange L pour écrire cette formule de quadrature.

En outre, le polynôme d'interpolation de Lagrange correspondant au cas $\alpha_0 = \dots = \alpha_p = 1$, nous avons $n = p + 1$ et donc le Corollaire 5.2.5 nous permet de majorer l'erreur comme suit.

Proposition 5.3.1 Soit $f \in C^{p+1}([a, b], \mathbb{R})$. Alors, on a la majoration :

$$\left| \int_a^b f - (b-a) \sum_{j=0}^p \omega_j f(x_j) \right| \leq \frac{1}{(p+1)!} \sup_{t \in [a, b]} |f^{(p+1)}(t)| \int_a^b |N(t)| dt.$$

Exemple. Pour la méthode *des rectangles à gauche*, le noyau d'interpolation est donné par $N(t) = t - a$, si bien que $\int_a^b |N(t)| dt = (b-a)^2/2$. Nous avons donc, sous l'hypothèse $f \in C^1([a, b], \mathbb{R})$,

$$\left| \int_a^b f(t) dt - (b-a)f(a) \right| \leq \frac{(b-a)^2}{2} \sup_{t \in [a, b]} |f'(t)|.$$

Cette majoration reste valable pour la méthode *des rectangles à droite*.

5.3.2 Avec un polynôme d'Hermite

Soit $p \in \mathbb{N}$, soit (x_0, x_1, \dots, x_p) un système interpolateur constitué de points du segment $[a, b]$, soit $(\omega_0, \dots, \omega_p) \in \mathbb{R}^{p+1}$ le $(p+1)$ -uplet de coefficients qui est associé à ce système interpolateur de rang p , et soit enfin un $(p+1)$ -uplet d'entiers non nuls $(\alpha_0, \dots, \alpha_p) \in (\mathbb{N}^*)^{p+1}$. Nous notons $n = \sum_{j=0}^p \alpha_j$ et, pour $f : [a, b] \rightarrow \mathbb{R}$, nous considérons le polynôme d'interpolation d'Hermite $H \in \mathbb{R}_{n-1}[X]$ de f , associé au système interpolateur (x_0, \dots, x_p) et aux entiers $(\alpha_0, \dots, \alpha_p)$. Nous voudrions pouvoir écrire :

$$\int_a^b f \simeq \int_a^b H = (b-a) \sum_{j=0}^p \omega_j H(x_j) = (b-a) \sum_{j=0}^p \omega_j f(x_j),$$

mais il apparaît ici une difficulté particulière. En effet, puisque $H \in \mathbb{R}_{n-1}[X]$, l'égalité précédente n'est justifiée que si notre méthode de quadrature est d'ordre $q \geq n-1$. Or, en dehors du cas $\alpha_0 = \dots = \alpha_p = 1$ où l'on considère un polynôme d'interpolation de Lagrange, nous avons $n-1 > p$, ce qui rend indispensable de choisir une méthode de quadrature d'ordre $q > p$.

Remarque. La formule de quadrature ci-dessus est identique à la formule (5.8) obtenue à l'aide du polynôme d'interpolation de Lagrange. L'intérêt d'introduire un polynôme d'interpolation d'Hermite — qu'il est d'ailleurs inutile de calculer — ne réside donc pas dans cette formule mais dans la majoration de l'erreur que nous pouvons déduire du Corollaire 5.2.5.

Ainsi, il est exact de dire qu'on obtient la formule de quadrature de Simpson en remplaçant la fonction par un polynôme de degré inférieur ou égal à 2 qui l'interpole aux points a , $(a+b)/2$ et b . Mais, en réalité, pour majorer l'erreur, c'est un polynôme d'Hermite de degré inférieur ou égal à 3 que l'on utilise, ce qui est possible car on a dans ce cas $q = 3 > p = 2$.

Avant de donner des exemples, nous reformulons ce qui précède en énonçant le résultat essentiel suivant.

Théorème 5.3.2 *Soit $p \in \mathbb{N}$, soit (x_0, x_1, \dots, x_p) un système interpolateur constitué de points du segment $[a, b]$, soit $(\omega_0, \dots, \omega_p) \in \mathbb{R}^{p+1}$ le $(p+1)$ -uplet de coefficients qui est associé à ce système interpolateur de rang p . Considérons un entier $n \geq p+1$ et une application $f \in C^n([a, b], \mathbb{R})$. Alors, nous avons la majoration suivante dans la formule de quadrature (5.8).*

$$\left| \int_a^b f - (b-a) \sum_{j=0}^p \omega_j f(x_j) \right| \leq \frac{M_n}{n!} \int_a^b |N(t)| dt,$$

où l'on a noté $M_n = \max_{t \in [a, b]} |f^{(n)}(t)|$ et où N est tout polynôme de la forme $N(X) = \prod_{j=0}^p (X - x_j)^{\alpha_j}$, avec $(\alpha_0, \dots, \alpha_p) \in (\mathbb{N}^*)^{p+1}$ tel que $\alpha_0 + \dots + \alpha_p = n$.

Démonstration: Pour $n \geq p+1$ et $f \in C^n([a, b], \mathbb{R})$ donnés, soient $\alpha_0, \dots, \alpha_p$ des entiers non nuls tels que $\alpha_0 + \dots + \alpha_p = n$. Nous pouvons considérer le polynôme d'interpolation d'Hermite H de l'application f relativement au système interpolateur (x_0, \dots, x_p) et aux entiers $(\alpha_0, \dots, \alpha_p)$, de sorte que N est le noyau d'interpolation associé, ce qui nous permet de conclure grâce au Corollaire 5.2.5. \square

Remarques.

1. Dès que $n > p+1$, il y a plusieurs choix possibles pour le polynôme N . On a intérêt à retenir un polynôme pour lequel la valeur de l'intégrale $\int_a^b |N(t)| dt$ est minimale.
2. Pour simplifier le calcul de cette dernière intégrale, on a souvent intérêt à se ramener à l'intervalle $[-1, 1]$ et à utiliser la formule (5.7).

Exemple. (Méthode du point milieu) On a $p = 0$, $x_0 = (a + b)/2$, $\omega_0 = 1$ et $q = 1 > p$. Ceci nous permet d'utiliser le polynôme d'Hermite H associé au point x_0 et à l'entier $\alpha_0 = 2$.

En nous ramenant à l'intervalle $[-1, 1]$, grâce au Lemme 5.2.6, nous calculons $N_1(X) = X^2$ puis

$$\int_{-1}^1 |N_1(t)| dt = \frac{2}{3} \quad \text{d'où} \quad \int_a^b |N(t)| dt = \frac{(b-a)^3}{12},$$

cette dernière égalité résultant de la formule (5.7).

Sous l'hypothèse $f \in C^2([a, b], \mathbb{R})$, le Théorème 5.3.2 nous permet d'en déduire la majoration de l'erreur suivante.

$$\left| \int_a^b f(t) dt - (b-a)f\left(\frac{a+b}{2}\right) \right| \leq \frac{(b-a)^3}{24} \sup_{t \in [a, b]} |f''(t)|.$$

Pour les méthodes de Newton-Cotes, nous savons que le cas $q > p$ ne se présente que si p est pair et nous avons alors $q = p + 1$. C'est pourquoi, comme nous l'avons déjà mentionné précédemment, hormis NC_1 (méthode des trapèzes), seules les méthodes NC_p avec p pair sont utilisées. Attardons-nous un peu sur la méthode NC_2 .

Exemple. (Méthode de Simpson) On a $x_0 = a$, $x_1 = (a + b)/2$, $x_2 = b$, et $q = 3 > p$, ce qui nous donne la *formule des trois niveaux* :

$$\forall P \in \mathbb{R}_3[X], \quad \int_a^b P(t) dt = \frac{b-a}{6} \left(P(a) + P(b) + 4P\left(\frac{a+b}{2}\right) \right).$$

Cette formule nous permet de considérer le polynôme d'Hermite relatif aux entiers $\alpha_0 = 1$, $\alpha_1 = 2$ et $\alpha_2 = 1$ ($n = 4$ donc $H \in \mathbb{R}_3[X]$).

En nous ramenant à l'intervalle $[-1, 1]$, grâce au Lemme 5.2.6, nous calculons $N_1(X) = (X + 1)X^2(X - 1) = X^4 - X^2$ puis

$$\int_{-1}^1 |N_1(t)| dt = 2 \int_0^1 (t^2 - t^4) dt = \frac{4}{15} \quad \text{d'où} \quad \int_a^b |N(t)| dt = \frac{(b-a)^5}{120},$$

cette dernière égalité résultant de la formule (5.7).

Finalement, le Théorème 5.3.2 nous donne la majoration d'erreur suivante, valable pour toute application $f \in C^4([a, b], \mathbb{R})$:

$$\left| \int_a^b f(t) dt - \frac{b-a}{6} \left(f(a) + f(b) + 4f\left(\frac{a+b}{2}\right) \right) \right| \leq \frac{(b-a)^5}{2880} \sup_{[a, b]} |f^{(4)}|.$$

Remarques.

1. Le lecteur vérifiera que le choix $\alpha_0 = 2, \alpha_1 = 1$ et $\alpha_2 = 1$, resp. le choix symétrique $\alpha_0 = 1, \alpha_1 = 1$ et $\alpha_2 = 2$, conduit à $\int_{-1}^1 |N_1(t)| dt = 1/2$, ce qui est moins intéressant que $4/15$ pour majorer l'erreur de quadrature.
2. Pour $p \geq 8$, il apparaît dans la méthode de Newton-Cotes de rang p des coefficients $\omega_j < 0$, qui peuvent amplifier les erreurs d'arrondis [DEM 63-64] car on a alors :

$$\sum_{j=0}^p |\omega_j| > \sum_{j=0}^p \omega_j = 1.$$

C'est pourquoi, outre les cas NC_2 (Simpson) et NC_4 (Boole-Villarceau) déjà présentés, seule NC_6 (Weddle-Hardy), est aussi utilisée dans la pratique.

5.3.3 Méthodes composées.

Elles consistent à subdiviser l'intervalle $[a, b]$ en sous-intervalles de la forme $[a + k(b-a)/N, a + (k+1)(b-a)/N]$, $k \in \llbracket 0; N-1 \rrbracket$, où $N \in \mathbb{N}^*$ est fixé, puis à appliquer une méthode élémentaire sur chacun des sous-intervalles, et enfin à additionner les résultats obtenus pour trouver une valeur approchée de l'intégrale de f sur $[a, b]$.

Méthode composée des rectangles

L'approximation

$$\int_a^b f(t) dt \simeq R_N(f) = \frac{b-a}{N} \sum_{k=0}^{N-1} f\left(a + k \frac{b-a}{N}\right)$$

donne lieu à une erreur qui, sous l'hypothèse $f \in C^1([a, b], \mathbb{R})$, peut être majorée comme suit :

$$\left| \int_a^b f(t) dt - R_N(f) \right| \leq \frac{(b-a)^2}{2N} \sup_{t \in [a, b]} |f'(t)|.$$

Méthode composée du point milieu

L'approximation

$$\int_a^b f(t) dt \simeq PM_N(f) = \frac{b-a}{N} \sum_{k=0}^{N-1} f\left(a + \left(k + \frac{1}{2}\right) \frac{b-a}{N}\right)$$

donne lieu à une erreur qui, sous l'hypothèse $f \in C^2([a, b], \mathbb{R})$, peut être majorée comme suit :

$$\left| \int_a^b f(t) dt - PM_N(f) \right| \leq \frac{(b-a)^3}{24N^2} \sup_{t \in [a, b]} |f''(t)|.$$

Tableau récapitulatif

En procédant comme précédemment, on obtient un exemple de méthode composée pour chaque méthode élémentaire présentée plus haut. Dans le tableau récapitulatif suivant, nous supposons $f \in C^{q+1}([a, b], \mathbb{R})$ pour une méthode d'ordre q et nous notons $M_{q+1} = \sup_{t \in [a, b]} |f^{(q+1)}(t)|$.

Méthode	Majorant de l'erreur
rectangles	$M_1(b-a)^2/(2N)$
point milieu	$M_2(b-a)^3/(24N^2)$
trapèzes	$M_2(b-a)^3/(12N^2)$
Simpson	$M_4(b-a)^5/(2880N^4)$

5.3.4 Lien avec la méthode de Romberg-Richardson

On peut montrer que la méthode d'accélération de convergence de Romberg-Richardson s'applique à la suite obtenue par la méthode des trapèzes lorsqu'on subdivise régulièrement $[a, b]$ en 2^m sous-intervalles [DEM 85-86].

On retrouve ainsi les méthodes de Simpson et de Boole-Villarceau. Ce procédé permet aussi de faire apparaître des méthodes de quadrature qui ne sont pas des méthodes de Newton-Cotes.

Une majoration précise de l'erreur est établie, sous forme d'exercices corrigés, dans l'ouvrage de CHAMBERT-LOIR et FERMIGIER intitulé *Exercices de mathématiques pour l'agrégation - Analyse 2* chez Masson, pages 186 à 193.

5.4 Méthodes de quadrature de Gauss.

Nous nous intéressons maintenant à une forme d'intégrale plus générale que celle étudiée précédemment :

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

où $f \in C^0([a, b], \mathbb{R})$ et $\pi \in C^0(]a, b[, \mathbb{R}_+^*)$, appelé *poids* sur $]a, b[$, est intégrable sur cet intervalle.

Pour obtenir une valeur approchée de cette intégrale, l'idée est de considérer le produit scalaire sur $C^0([a, b], \mathbb{R})$ défini par :

$$\langle f, g \rangle = \int_a^b f(t)g(t) \pi(t) dt.$$

En partant de la base canonique de $\mathbb{R}[X]$ et en lui appliquant le procédé d'orthogonalisation de Gram-Schmidt, nous obtenons une suite $(P_n)_{n \in \mathbb{N}}$ de polynômes orthogonaux pour le produit scalaire que nous venons de définir. On peut montrer que pour tout $n \in \mathbb{N}$, P_n est un polynôme de degré n dont le coefficient de plus haut degré vaut 1 et qui admet n racines simples appartenant à l'intervalle $]a, b[$ (cf. [SKA 31-33]).

Pour tout $p \in \mathbb{N}$, la méthode de Gauss de rang p consiste à prendre pour système interpolateur (x_0, \dots, x_p) les racines du polynôme P_{p+1} .

L'application $P \mapsto \int_a^b P(x) \pi(x) dx$ étant une forme linéaire sur $\mathbb{R}_p[X]$, la Proposition 5.1.2 implique le résultat suivant.

Proposition 5.4.1 *Il existe un unique $(p+1)$ -uplet de coefficients réels $(\lambda_0, \dots, \lambda_p)$ tel que :*

$$\forall P \in \mathbb{R}_p[X], \quad \int_a^b P(t) \pi(t) dt = \sum_{j=0}^p \lambda_j P(x_j).$$

L'intérêt principal des méthodes de Gauss est que, pour un rang donné $p \in \mathbb{N}$, elles maximisent l'ordre. En effet, [DAN 498-499] établit le résultat suivant.

Proposition 5.4.2 *La méthode de Gauss de rang p est d'ordre $(2p+1)$. C'est l'unique méthode de quadrature de rang p admettant un ordre supérieur ou égal $(2p+1)$.*

Remarque. En particulier, si q est l'ordre d'une méthode de quadrature quelconque de rang p , on a $p \leq q \leq 2p+1$.

En utilisant alors le polynôme d'Hermite correspondant à $\alpha_0 = \dots = \alpha_p = 2$, nous obtenons le résultat suivant, grâce à la Proposition 5.2.4.

Proposition 5.4.3 *Soit $f \in C^{2p+2}([a, b], \mathbb{R})$. On peut majorer comme suit l'erreur relative à la méthode de Gauss à $(p+1)$ points :*

$$\left| \int_a^b f(t) \pi(t) dt - \sum_{j=0}^p \lambda_j f(x_j) \right| \leq \frac{\sup_{t \in [a, b]} |f^{(2p+2)}(t)|}{(2p+2)!} \int_a^b P_{p+1}^2(t) \pi(t) dt.$$

Il y a un bémol à ce beau résultat théorique : les polynômes orthogonaux sont difficiles à calculer dans la pratique. C'est pour cette raison que les méthodes de Gauss ne sont utilisées que pour les deux poids suivants :

- $\pi \equiv 1$ sur $[-1, 1]$, c'est la méthode de Gauss-Legendre [DEM 76] ;
- $\pi :]-1, 1[\rightarrow \mathbb{R}_+^*$ défini par :

$$\forall t \in]-1, 1[, \quad \pi(t) = \frac{1}{\sqrt{1-t^2}}.$$

C'est la méthode de Gauss-Tchebychev [DAN 369-376, exo corrigé]. Cette méthode d'ordre $(2p+1)$ s'écrit :

$$\int_{-1}^1 f(t) \frac{dt}{\sqrt{1-t^2}} \simeq \frac{\pi}{p+1} \sum_{j=0}^p f\left(\cos\left(\frac{2j+1}{2p+2}\pi\right)\right).$$

5.5 Une méthode probabiliste

Dans le tome 2 de son livre *Probabilités* pages 128-129, Jean-Yves Ouvrard présente le principe des *méthodes de Monte-Carlo* dans un cas simple, sous forme de l'exercice corrigé 10.8, utilisant l'inégalité de Bienaymé-Tchebychev.

Chapitre 6

Applications de l'analyse au calcul des grandeurs

Avertissement : Ce chapitre ne prétend nullement à l'exhaustivité !

Bibliographie :

FLO Topologie, Analyse (exercices avec solutions) tome 3 par Georges FLORY chez Vuibert

LT Intégrales curvilignes et de surfaces (niveau L2) par LOFFICIAL et TANRÉ chez Ellipses

RDO Cours de Mathématiques Spéciales tome 5 (applications de l'analyse à la géométrie) par RAMIS, DESCHAMPS, ODOUX chez Masson

6.1 Longueur d'un arc de classe C^1

6.1.1 Rectification d'un arc paramétré.

Définitions générales

Définition 6.1.1 Soit $d \in \mathbb{N}^*$ fixé et $f \in C([a, b], \mathbb{R}^d)$. On dit que le couple $\Gamma = ([a, b], f)$ est un arc paramétré continu de \mathbb{R}^d . Cet arc est dit rectifiable si

$$\sup \left\{ \sum_{i=0}^{n-1} \|f(a_{i+1}) - f(a_i)\| \right\} < +\infty,$$

cette borne supérieure portant sur toutes les subdivisions $(a_i)_{0 \leq i \leq n}$ ($n \in \mathbb{N}^*$) de l'intervalle $[a, b]$ et $\|\cdot\|$ désignant la norme euclidienne dans \mathbb{R}^d .

Dans ce cas, cette borne supérieure est appelée longueur de l'arc paramétré continu Γ et notée $l(\Gamma)$.

Un exemple d'arc continu non rectifiable nous est donné par l'application $f : [0, 1] \rightarrow \mathbb{R}^2$ définie par

$$f(0) = (0, 0) \text{ et } \forall t \in]0, 1[\quad f(t) = (t, t \sin \frac{1}{t}).$$

Le lecteur pourra démontrer que la borne supérieure ci-dessus est infinie en choisissant judicieusement des subdivisions et en utilisant la divergence de la série harmonique (i.e. $\sum_n \frac{1}{n} = +\infty$).

Remarque. Le fait d'être rectifiable et la longueur d'un arc sont des propriétés *géométriques* pour un arc continu, en ce sens qu'un changement de paramètre admissible (homéomorphisme) ne les modifie pas.

Pour $a \leq t < t' \leq b$, notons $\Gamma_{t,t'}$ le sous-arc $([t, t'], f|_{[t,t']})$ de Γ .

Définition 6.1.2 Soit $\Gamma = ([a, b], f)$ un arc paramétré continu et rectifiable de \mathbb{R}^d . Nous dirons qu'une application $s : [a, b] \rightarrow \mathbb{R}$ est une abscisse curviligne de Γ si pour tout $(t, t') \in [a, b]^2$ tel que $t < t'$, on a :

$$l(\Gamma_{t,t'}) = |s(t') - s(t)|$$

Si, en outre, il existe $t_0 \in [a, b]$ tel que $s(t_0) = 0$, nous dirons que s est une abscisse curviligne d'origine t_0 .

Proposition 6.1.3 Soient $\Gamma = ([a, b], f)$ un arc paramétré rectifiable de \mathbb{R}^d , s_1 et s_2 deux abscisses curvilignes de Γ . Alors il existe $\epsilon = \pm 1$ et $c \in \mathbb{R}$ tels que $s_2 = \epsilon s_1 + c$.

Réciproquement, si s est une abscisse curviligne de Γ , alors toute application de la forme $\epsilon s + c$ est une abscisse curviligne de Γ .

Démonstration: RDO 92

□

Arc de classe C^1

Dans toute la suite, nous supposerons que l'arc considéré est de classe C^1 , conformément au titre de la leçon. Nous avons alors le résultat essentiel suivant :

Théorème 6.1.4 Soit $f \in C^1([a, b], \mathbb{R}^d)$ et $\Gamma = ([a, b], f)$ l'arc paramétré associé. Alors cet arc est rectifiable, de longueur

$$l(\Gamma) = \int_a^b \|f'(t)\| dt.$$

Démonstration: DANTZER 203-205 ou RDO 94-95

□

Corollaire 6.1.5 *Sous les hypothèses du théorème précédent, l'application $s : [a, b] \rightarrow \mathbb{R}$ définie par*

$$\forall t \in [a, b] \quad s(t) = \int_a^t \|f'(u)\| du$$

est une abscisse curviligne de Γ d'origine a .

6.1.2 Calcul pratique

Nous nous plaçons maintenant en dimension $d = 3$ et nous établissons des formules pratiques pour calculer la longueur d'un arc suivant le système de coordonnées utilisé : cartésiennes, cylindriques ou sphériques.

Coordonnées cartésiennes

Pour tout $t \in [a, b]$, nous noterons $f(t) = (x(t), y(t), z(t))$ si bien que $\|f'(t)\| = \sqrt{x'^2(t) + y'^2(t) + z'^2(t)}$ et le théorème précédent nous donne :

$$l(\Gamma) = \int_a^b \sqrt{x'^2(t) + y'^2(t) + z'^2(t)} dt$$

Pour obtenir une formule mnémotechnique, nous remarquons que l'abscisse curviligne s qui apparaît dans le corollaire précédent vérifie :

$$\left(\frac{ds}{dt}\right)^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2$$

ce que nous retenons sous la forme :

$$\boxed{ds^2 = dx^2 + dy^2 + dz^2}$$

Remarque. Tout ce qui vient d'être fait en dimension 3 se généralise immédiatement en dimension d .

Exercice : [FLO 153]

Calculer la longueur totale de l'astroïde définie sous forme paramétrée par

$$f(t) = (a \cos^3 t, a \sin^3 t) \quad , \quad 0 \leq t \leq 2\pi \quad (a > 0).$$

Exercice : [RDO 97] ou [LT 143]

Calculer la longueur d'une arche de cycloïde définie sous forme paramétrée par

$$f(t) = (a(t - \sin t), a(1 - \cos t)) \quad , \quad 0 \leq t \leq 2\pi.$$

Coordonnées cylindriques

Pour tout $t \in [a, b]$, nous avons $f(t) = (\rho(t) \cos \theta(t), \rho(t) \sin \theta(t), z(t))$ et nous calculons facilement $\|f'(t)\| = \sqrt{\rho'^2(t) + \rho^2(t)\theta'^2(t) + z'^2(t)}$.

En procédant comme précédemment, nous obtenons donc la formule mnémotechnique :

$$\boxed{ds^2 = d\rho^2 + \rho^2 d\theta^2 + dz^2}$$

Exercice : [RDO 104] ou [LT 145]

Calculer la longueur totale de l'hélice circulaire définie en coordonnées cylindriques par

$$\rho \equiv a > 0 \quad , \quad z = b\theta \quad , \quad 0 \leq \theta \leq 2\pi \quad (b \in \mathbb{R}^*)$$

Remarque. En faisant $z = 0$ dans la formule précédente, nous obtenons le cas d'un arc représenté en **coordonnées polaires** dans le plan euclidien orienté :

$$\boxed{ds^2 = d\rho^2 + \rho^2 d\theta^2}$$

Exercice : [RDO 97] ou [LT 141]

Calculer la longueur totale de la cardioïde définie en coordonnées polaires par

$$\rho = a(1 + \cos \theta) \quad , \quad -\pi \leq \theta \leq \pi \quad (a > 0).$$

Exercice : [FLO 157]

Sur l'arc d'une lemniscate de Bernoulli défini en coordonnées polaires par

$$\rho = a\sqrt{\cos(2\theta)} \quad , \quad 0 \leq \theta \leq \frac{\pi}{4}$$

nous considérons le point A obtenu pour $\theta = 0$, ainsi que deux points M et M' obtenus respectivement pour $\theta = t$ et $\theta = t'$, où t et t' sont des éléments de l'intervalle $[0, \frac{\pi}{4}]$ vérifiant la relation :

$$\cos t \cos t' = \frac{1}{\sqrt{2}}$$

Montrer que la longueur de l'arc de la lemniscate entre l'origine O (obtenue pour $\theta = \pi/4$) et le point M est égale à celle de l'arc de la lemniscate entre le point M' et le point A.

Exercice : [EMPSI 390]

Calculer l'abscisse curviligne en tout point $M(\theta)$ de l'arc défini en coordonnées polaires par

$$\rho = \tanh \frac{\theta}{2}$$

en prenant comme origine des abscisses curvilignes le point de cet arc correspondant à $\theta = 0$.

Coordonnées sphériques

Pour tout $t \in [a, b]$, nous avons

$$f(t) = (r(t) \sin \theta(t) \cos \varphi(t), r(t) \sin \theta(t) \sin \varphi(t), r(t) \cos \theta(t)),$$

ce qui nous donne $\|f'(t)\| = \sqrt{r'^2(t) + r^2(t)\theta'^2(t) + r^2(t)\sin^2 \theta(t)\varphi'^2(t)}$.

Nous retiendrons donc la formule :

$$\boxed{ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2}$$

Exercice : [FLO 155]

Déterminer l'application $T : [-\pi/4, \pi/4] \rightarrow \mathbb{R}$ pour que la tangente à l'arc défini par

$$f(\varphi) = (a \cos T(\varphi) \cos \varphi, a \sin T(\varphi) \cos \varphi, a \sin \varphi) \quad , \quad -\frac{\pi}{4} \leq \varphi \leq \frac{\pi}{4}$$

fasse un angle constant et égal à $\pi/4$ avec l'axe (Oz).

Calculer la longueur de cet arc.

Le premier théorème de Guldin

Théorème 6.1.6 *Considérons une surface de révolution \mathcal{S} engendrée par la rotation d'un arc plan Γ autour d'un axe situé dans le plan de Γ et ne coupant pas cet arc. L'aire de \mathcal{S} est égale à la longueur $l(\Gamma)$ multipliée par la longueur du chemin parcouru par le centre d'inertie G de Γ lors de la révolution autour de l'axe.*

L'application classique est le calcul de l'aire latérale du tore à collier [LT 153] ou [RDO 295] mais on peut le généraliser en prenant pour demi-méridienne une ellipse à la place d'un cercle.

Inégalité isopérimétrique

Terminons par ce résultat (plus difficile que les précédents) qui peut trouver sa place dans la leçon “Ex. de calcul de la longueur d'un arc de classe C^1 ”.

Théorème 6.1.7 *Dans le plan, si un arc de classe C^1 et de longueur l délimite une portion du plan d'aire A , alors on a l'inégalité $4\pi A \leq l^2$ et l'égalité est obtenue si et seulement si l'arc paramétré définit un cercle parcouru une fois.*

La démonstration utilise les séries de Fourier et se trouve dans la livre *Analyse pour l'agrégation* par QUEFFELEC et ZUILY chez Dunod, pages 103-105.

6.2 Produit mixte et produit vectoriel dans \mathbb{R}^n

6.2.1 Orientation de \mathbb{R}^n

Si nous considérons deux bases $e = (e_1, \dots, e_n)$ et $e' = (e'_1, \dots, e'_n)$ de \mathbb{R}^n et si nous notons $P_e^{e'}$ la matrice de passage de e à e' , nous définissons une relation d'équivalence sur l'ensemble des bases de \mathbb{R}^n par :

$$e \sim e' \iff \det P_e^{e'} > 0$$

Ceci résulte en effet des égalités suivantes :

$$P_e^e = I_n \quad ; \quad P_e^{e'} = (P_e^{e'})^{-1} \quad ; \quad P_e^{e''} = P_e^{e'} P_{e'}^{e''}$$

Il n'y a que deux classes d'équivalence. En effet, si $e = (e_1, e_2, \dots, e_n)$ est une base arbitraire fixée, alors $e' = (-e_1, e_2, \dots, e_n)$ est une base telle que $\det P_e^{e'} = -1$ et donc toute base f de \mathbb{R}^n vérifie soit $f \sim e$ soit $f \sim e'$.

On dit qu'on *oriente* \mathbb{R}^n en choisissant l'une de ces deux classes d'équivalence. De façon plus pratique, nous orienterons \mathbb{R}^n en choisissant une base e et en disant que les bases qui sont dans la classe d'équivalence de e sont *positives*, les autres étant qualifiées de *negatives*.

Convention : Dans toute la suite, nous orienterons systématiquement \mathbb{R}^n en choisissant sa base canonique.

Considérons un n -uplet $(x^1, \dots, x^n) \in (\mathbb{R}^n)^n$ tel que

$$\forall j = 1, \dots, n \quad x^j = (x_1^j, \dots, x_n^j)$$

et notons X sa matrice par rapport à la base canonique, c'est-à-dire :

$$X = \left[x_i^j \right]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$$

Nous constatons alors facilement que (x^1, \dots, x^n) est une base positive de \mathbb{R}^n si et seulement si $\det X > 0$.

6.2.2 Produit mixte dans \mathbb{R}^n euclidien orienté

Nous nous plaçons désormais dans l'espace euclidien \mathbb{R}^n muni du produit scalaire usuel noté $(x | y)$.

La base canonique e constitue une base orthonormée de \mathbb{R}^n .

Rappelons qu'une base e' est alors orthonormée si et seulement si (en notant M^* la transposée de la matrice M) :

$$(P_e^{e'})^* P_e^{e'} = P_e^{e'} (P_e^{e'})^* = I_n$$

Nous en déduisons en particulier : $(\det P_e^{e'})^2 = 1$.

Ainsi, si e' est une base orthonormée positive de \mathbb{R}^n , nous avons $\det P_e^{e'} = 1$; si e' est une base orthonormée négative de \mathbb{R}^n , nous avons $\det P_e^{e'} = -1$.

Proposition et définition 6.2.1 Soit $f = (f_1, \dots, f_n)$ une base orthonormée positive de \mathbb{R}^n (orienté suivant sa base canonique).

Considérons un n -uplet $(x^1, \dots, x^n) \in (\mathbb{R}^n)^n$ et notons $M(x^1, \dots, x^n; f)$ la matrice de ce n -uplet par rapport à la base f , c'est-à-dire la matrice carrée

$$\left[m_i^j \right]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \quad \text{telle que} \quad \forall j = 1, \dots, n \quad x^j = \sum_{i=1}^n m_i^j f_i$$

Alors le réel $\det M(x^1, \dots, x^n; f)$ ne dépend pas du choix de la base orthonormée positive f .

On l'appelle produit mixte de (x^1, \dots, x^n) et on le note $[x^1, \dots, x^n]$. Le produit mixte est une application n -linéaire alternée.

Démonstration: Soient f et f' deux bases orthonormées positives de \mathbb{R}^n et $P_f^{f'}$ la matrice de passage de f à f' .

Puisqu'il s'agit d'une matrice de passage entre bases orthonormées, nous savons qu'elle est orthogonale et donc que $(\det P_f^{f'})^2 = 1$. Comme f et f' ont la même orientation, nous en déduisons $\det P_f^{f'} = 1$.

L'égalité matricielle :

$$M(x^1, \dots, x^n; f) = P_f^{f'} M(x^1, \dots, x^n; f')$$

implique donc l'égalité $\det M(x^1, \dots, x^n; f) = \det M(x^1, \dots, x^n; f')$.

La dernière phrase résulte immédiatement des propriétés des déterminants. \square

Si nous prenons pour base orthonormée positive f la base canonique, la dernière phrase du paragraphe précédent nous donne l'équivalence :

$$(x^1, \dots, x^n) \text{ base positive de } \mathbb{R}^n \Leftrightarrow [x^1, \dots, x^n] > 0$$

Dans le plan ($n = 2$), un simple raisonnement géométrique permet de constater que $|[x, y]|$ est l'aire du parallélogramme construit à partir des vecteurs non colinéaires x et y .

6.2.3 Produit vectoriel dans \mathbb{R}^n euclidien orienté

Théorème et définition 6.2.2 *Dans \mathbb{R}^n orienté suivant sa base canonique, considérons $n - 1$ vecteurs x^1, \dots, x^{n-1} . Il existe un unique vecteur $v \in \mathbb{R}^n$ tel que*

$$\forall x^n \in \mathbb{R}^n \quad [x^1, \dots, x^{n-1}, x^n] = (v \mid x^n)$$

On l'appelle produit vectoriel de x^1, \dots, x^{n-1} et on le note $v = x^1 \wedge \dots \wedge x^{n-1}$. L'application $(x^1, \dots, x^{n-1}) \mapsto x^1 \wedge \dots \wedge x^{n-1}$ est $(n-1)$ -linéaire et alternée.

Démonstration: Fixons $n - 1$ vecteurs x^1, \dots, x^{n-1} de \mathbb{R}^n et considérons l'application $l : \mathbb{R}^n \rightarrow \mathbb{R}$ définie par

$$\forall x^n \in \mathbb{R}^n \quad l(x^n) = [x^1, \dots, x^{n-1}, x^n]$$

Le produit mixte étant n -linéaire, l'application l est clairement une forme linéaire sur \mathbb{R}^n , donc un élément de son dual : $l \in (\mathbb{R}^n)^*$.

Rappelons que l'application $\varphi : v \mapsto (\cdot \mid v)$ est un isomorphisme de \mathbb{R}^n sur son dual $(\mathbb{R}^n)^*$ appelé isomorphisme canonique. Nous constatons alors que l'unique vecteur de \mathbb{R}^n vérifiant la condition de l'énoncé est $\varphi^{-1}(l)$.

La dernière phrase résulte des propriétés du produit mixte. \square

Le lecteur montrera à titre d'exercice, en utilisant le théorème de la base incomplète, que le produit vectoriel $x^1 \wedge \dots \wedge x^{n-1}$ est nul si et seulement si la famille (x^1, \dots, x^{n-1}) est liée.

En outre, si la famille (x^1, \dots, x^{n-1}) est libre, il constatera que la famille $(x^1, \dots, x^{n-1}, x^1 \wedge \dots \wedge x^{n-1})$ forme alors une base positive de \mathbb{R}^n puisque :

$$[x^1, \dots, x^{n-1}, x^1 \wedge \dots \wedge x^{n-1}] = \|x^1 \wedge \dots \wedge x^{n-1}\|^2 > 0$$

Calcul pratique du produit vectoriel Soit $f = (f_1, \dots, f_n)$ une base orthonormée positive de \mathbb{R}^n . Donnons-nous $(x^1, \dots, x^{n-1}) \in (\mathbb{R}^n)^{n-1}$ par :

$$\forall j = 1, \dots, n-1 \quad x^j = \sum_{i=1}^n \xi_i^j f_i$$

et posons $x^1 \wedge \dots \wedge x^{n-1} = \sum_{i=1}^n \alpha_i f_i$.

L'égalité des formes linéaires $(\xi_1, \dots, \xi_n) \mapsto \alpha_1 \xi_1 + \dots + \alpha_n \xi_n$ et

$$(\xi_1, \dots, \xi_n) \mapsto \begin{vmatrix} \xi_1^1 & \dots & \xi_1^{n-1} & \xi_1 \\ \vdots & \vdots & \vdots & \vdots \\ \xi_n^1 & \dots & \xi_n^{n-1} & \xi_n \end{vmatrix}$$

équivalent à celle des coefficients de ξ_1, \dots, ξ_n . Nous en déduisons la proposition suivante.

Proposition 6.2.3 *Pour tout $i = 1, \dots, n$, la i -ème coordonnée du vecteur $x^1 \wedge \dots \wedge x^{n-1}$ dans la base orthonormée positive f est égale au cofacteur du terme en position (i, n) dans la matrice, par rapport à la base f , de la famille (x^1, \dots, x^{n-1}, x) , où $x \in \mathbb{R}^n$ est arbitrairement choisi.*

Soit (x^1, \dots, x^{n-1}) une famille orthogonale dans \mathbb{R}^n . Si nous posons

$$\forall j = 1, \dots, n-1 \quad f_j = \frac{x^j}{\|x^j\|}$$

nous pouvons trouver $f_n \in \mathbb{R}^n$ tel que la famille $f = (f_1, \dots, f_n)$ constitue une base orthonormée positive de \mathbb{R}^n . En appliquant la proposition précédente, nous obtenons facilement :

$$(x^1, \dots, x^{n-1}) \text{ orthogonale} \implies \|x^1 \wedge \dots \wedge x^{n-1}\| = \prod_{j=1}^{n-1} \|x^j\|$$

6.2.4 Interprétations géométriques en dimension 3

L'interprétation suivante nous sera utile pour les calculs d'aires de surfaces paramétrées dans \mathbb{R}^3 .

Considérons deux vecteurs non colinéaires u et v dans \mathbb{R}^3 et notons π la projection orthogonale sur la droite $\mathbb{R}u$, de sorte que $(u, v - \pi v)$ forme une famille orthogonale. Nous obtenons alors :

$$\|u \wedge v\| = \|u \wedge (v - \pi v)\| = \|u\| \|v - \pi v\|$$

Un simple dessin nous permet alors d'interpréter $\|u \wedge v\|$ comme l'aire du parallélogramme construit à partir des vecteurs non colinéaires u et v . La propriété suivante est souvent utilisée comme définition du produit vectoriel dans \mathbb{R}^3 dans une présentation élémentaire.

Proposition 6.2.4 *Soient u et v deux vecteurs de \mathbb{R}^3 . Si (u, v) est lié, alors $u \wedge v$ est le vecteur nul. Sinon, $u \wedge v$ est l'unique vecteur orthogonal au plan engendré par (u, v) , tel que $(u, v, u \wedge v)$ soit une base positive et que :*

$$\|u \wedge v\| = \|u\| \|v\| \sin \theta,$$

où θ est l'écart angulaire de u et v .

Dans le même ordre d'idée, si (u, v, w) est un système libre (et donc une base) de \mathbb{R}^3 , le produit mixte $[u, v, w] = (u \wedge v \mid w)$ peut être interprété (au signe près) comme le produit de l'aire du parallélogramme construit à partir de (u, v) par $\|\pi w\|$, où π est la projection sur la droite supplémentaire orthogonal de $\text{Vect}(u, v)$.

Si nous considérons le parallélotope construit à partir de (u, v, w) , la quantité $\|\pi w\|$ apparaît comme la "hauteur" de ce parallélotope. Finalement, $|[u, v, w]|$ est le produit de l'aire du parallélogramme qui constitue la "base" du parallélotope par la hauteur de ce dernier, c'est-à-dire son volume, comme on pourrait le montrer par un découpage savant ramenant ce parallélotope à un parallélépipède rectangle...

Retenons que $|[u, v, w]|$ est le volume du parallélotope :

$$\mathcal{P} = \{\alpha u + \beta v + \gamma w, 0 < \alpha < 1, 0 < \beta < 1, 0 < \gamma < 1\}$$

Comme le produit mixte est par définition un déterminant dans une base orthonormale positive, cette interprétation géométrique nous permet de comprendre intuitivement pourquoi la valeur absolue du déterminant jacobien apparaît dans le théorème du changement de variable, comme nous allons le voir dans la section suivante.

6.3 Théorème du changement de variables

6.3.1 Énoncé général

Nous rappelons ici cet énoncé en dimension 3 en vue de l'interprétation géométrique annoncée ci-dessus mais le théorème du changement de variables est valable en dimension $d \in \mathbb{N}^*$.

Théorème 6.3.1 Soit $\Phi : U \rightarrow V$ un C^1 -difféomorphisme entre deux ouverts de \mathbb{R}^3 et $f : V \rightarrow \mathbb{R}$ une application continue. Alors l'application f est intégrable sur V si et seulement si l'application $f \circ \Phi |J_\Phi|$ est intégrable sur U et, dans ce cas, nous avons l'égalité :

$$\int_V f(x, y, z) \, dx \, dy \, dz = \int_U f(\Phi(s, t, u)) |J_\Phi(s, t, u)| \, ds \, dt \, du$$

D'un point de vue pratique, on dit que l'on “pose le changement de variables $(x, y, z) = \Phi(s, t, u)$ ” et l'on écrit l'égalité entre “éléments différentiels” $dx \, dy \, dz = |J_\Phi(s, t, u)| \, ds \, dt \, du$.

Intuitivement, ce résultat admet l'interprétation géométrique suivante. Le difféomorphisme de U sur V noté $\Phi : (s, t, u) \mapsto (x, y, z)$ se comporte localement comme une application linéaire qui transforme le parallélépipède rectangle élémentaire $[s, s + ds] \times [t, t + dt] \times [u, u + du]$ de U en un parallélépipède élémentaire de V , défini (de façon naturelle) par le point (x, y, z) et les trois vecteurs :

$$\left(\frac{\partial x}{\partial s}, \frac{\partial y}{\partial s}, \frac{\partial z}{\partial s} \right) ds \quad , \quad \left(\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}, \frac{\partial z}{\partial t} \right) dt \quad , \quad \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right) du .$$

Selon le paragraphe précédent, son volume est donné par le produit mixte de ces trois vecteurs, autrement dit par l'élément différentiel $ds \, dt \, du$ multiplié par la valeur absolue du déterminant suivant :

$$\begin{vmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} & \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} & \frac{\partial y}{\partial u} \\ \frac{\partial z}{\partial s} & \frac{\partial z}{\partial t} & \frac{\partial z}{\partial u} \end{vmatrix}$$

qui n'est autre que $J_\Phi(s, t, u)$ par définition.

L'intégrale sur V apparaît comme une somme sur chacun de ces parallélépipèdes élémentaires.

Application : Volume d'une boule dans \mathbb{R}^d

Référence : *Topologie, Analyse par Georges FLORY, tome 3, éditions Vuibert, pages 177-178*

Pour $r > 0$, nous notons $B_d(r) = \{x \in \mathbb{R}^d, \|x\| < r\}$. Le volume de cette boule sera noté $V_d(r)$.

1. En appliquant le théorème du changement de variables dans \mathbb{R}^d , montrer l'égalité :

$$V_d(r) = r^d V_d(1)$$

2. Nous utiliserons la notation suivante pour les intégrales de Wallis :

$$I_d = \int_0^{\frac{\pi}{2}} \sin^d u \, du$$

En appliquant le théorème de Fubini puis un changement de variable simple, montrer les égalités

$$V_d(r) = V_{d-1}(1) \int_{-r}^r (r^2 - x^2)^{\frac{d-1}{2}} dx = 2r^d I_d V_{d-1}(1)$$

3. On rappelle les valeurs des intégrales de Wallis :

$$I_{2n} = \frac{(2n)!}{(2^n n!)^2} \frac{\pi}{2}, \quad I_{2n+1} = \frac{(2^n n!)^2}{(2n+1)!} \quad (n \in \mathbb{N})$$

Déduire de ce qui précède les égalités

$$V_{2n}(1) = \frac{\pi^n}{n!}, \quad V_{2n+1}(1) = \frac{2^{2n+1} n! \pi^n}{(2n+1)!}$$

4. La boule $B_d(1)$ est inscrite dans l'hypercube :

$$C_d = \{x \in \mathbb{R}^d, \max_{1 \leq i \leq d} |x_i| < 1\}$$

Comparer le comportement asymptotique de leurs volumes respectifs lorsque d tend vers l'infini.

Remarque : Il est facile de vérifier numériquement que le volume de la boule unité est maximal en dimension 5. Le volume d'une boule de rayon 2 est quant à lui maximal en dimension 24... Tout cela n'est pas tellement intuitif!

6.3.2 Quelques cas particuliers

Coordonnées polaires

L'application $\Phi : \mathbb{R}_+^* \times]0, 2\pi[\rightarrow \mathbb{R}^2 / (\mathbb{R}_+ \times \{0\})$ définie par

$$\Phi(\rho, \theta) = (x(\rho, \theta), y(\rho, \theta)) = (\rho \cos \theta, \rho \sin \theta)$$

est un C^1 -difféomorphisme de déterminant jacobien

$$J_\Phi(\rho, \theta) = \begin{vmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{vmatrix} = \rho$$

Nous retiendrons donc la formule :

$$\boxed{dx dy = \rho d\rho d\theta}$$

Une application très classique est le calcul de l'intégrale de Gauss :

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$

Coordonnées cylindriques

L'application $\Phi : \mathbb{R}_+^* \times]0, 2\pi[\times \mathbb{R} \rightarrow \mathbb{R}^3 / (\mathbb{R}_+ \times \{0\} \times \mathbb{R})$ définie par

$$\Phi(\rho, \theta, z) = (x(\rho, \theta, z), y(\rho, \theta, z), z(\rho, \theta, z)) = (\rho \cos \theta, \rho \sin \theta, z)$$

est un C^1 -difféomorphisme de déterminant jacobien

$$J_\Phi(\rho, \theta, z) = \begin{vmatrix} \cos \theta & -\rho \sin \theta & 0 \\ \sin \theta & \rho \cos \theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = \rho$$

Nous retiendrons donc la formule :

$$\boxed{dx dy dz = \rho d\rho d\theta dz}$$

Application : Volume d'un solide de révolution

Réf : *Intégrales curvilignes et de surfaces, L'OFFICIAL et TANRE, Ellipses* noté [LT] dans la suite, page 93.

Soit $f :]a, b[\rightarrow \mathbb{R}_+^*$ continue et S le solide de révolution autour de l'axe Oz défini par

$$S := \{(x, y, z) \in \mathbb{R}^3 / x^2 + y^2 < f^2(z), a < z < b\}$$

Calculer son volume $V(S)$.

Définition 6.3.2 Soit $\mathcal{P} \subset \mathbb{R}^2$ (supposée géométriquement simple par exemple) une plaque homogène dans le plan yOz , d'aire $\mathcal{A}(\mathcal{P}) \in \mathbb{R}_+^*$. Son centre d'inertie est le point $G = (y_G, z_G) \in \mathbb{R}^2$ tel que

$$\begin{pmatrix} y_G \\ z_G \end{pmatrix} := \frac{1}{\mathcal{A}(\mathcal{P})} \iint_{\mathcal{P}} \begin{pmatrix} y \\ z \end{pmatrix} dy dz$$

Remarque. La notion de centre d'inertie généralise celle de barycentre (ou centre de gravité) de n points pondérés dans \mathbb{R}^2 . L'hypothèse d'homogénéité de la plaque correspond au cas où tous les points ont même poids.

Application : Un théorème de Guldin

Réf : [LT 96]

Avec les notations de l'application précédente, considérons

$$\mathcal{P} := \{0 < y < f(z), a < z < b\}$$

une plaque homogène de centre d'inertie G et soit S le solide de révolution qu'elle engendre par rotation autour de l'axe Oz . Montrer l'égalité :

$$V(S) = 2\pi y_G \mathcal{A}(\mathcal{P})$$

Remarque. Ce théorème de Guldin se généralise immédiatement à une plaque homogène

$$\mathcal{P} := \{f_1(z) < y < f_2(z), a < z < b\}$$

où f_1 et f_2 sont deux applications telles que $0 < f_1 < f_2$.

Coordonnées sphériques

L'application $\Phi : \mathbb{R}_+^* \times]0, \pi[\times]0, 2\pi[\rightarrow \mathbb{R}^3 / (\mathbb{R}_+ \times \{0\} \times \mathbb{R})$ définie par

$$\Phi(r, \theta, \phi) = (x(r, \theta, \phi), y(r, \theta, \phi), z(r, \theta, \phi)) = (r \sin \theta \cos \phi, r \sin \theta \sin \phi, r \cos \theta)$$

est un C^1 -difféomorphisme de déterminant jacobien

$$J_\Phi(r, \theta, \phi) = \begin{vmatrix} \sin \theta \cos \phi & r \cos \theta \cos \phi & -r \sin \theta \sin \phi \\ \sin \theta \sin \phi & r \cos \theta \sin \phi & r \sin \theta \cos \phi \\ \cos \theta & -r \sin \theta & 0 \end{vmatrix} = r^2 \sin \theta$$

Nous retiendrons donc la formule :

$$\boxed{dx dy dz = r^2 \sin \theta dr d\theta d\phi}$$

Définition 6.3.3 Soit $S \subset \mathbb{R}^3$ (supposé géométriquement simple par exemple) un solide muni d'une "fonction de masse volumique" continue $\rho : S \rightarrow \mathbb{R}_+$. Le moment d'inertie du solide S par rapport à l'origine O est l'intégrale :

$$\iiint_S (x^2 + y^2 + z^2) \rho(x, y, z) dx dy dz$$

Exercice :*Réf :* [LT 106]

Calculer le moment d'inertie par rapport à l'origine des solides homogènes de masse volumique 1 définis par :

$$S_1 := \{(x, y, z) \in \mathbb{R}^3 / x > 0, y > 0, z > 0, x^2 + y^2 + z^2 < 1\}$$

$$S_2 := \left\{ (x, y, z) \in \mathbb{R}^3 / x > 0, y > 0, z > 0, \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1 \right\},$$

avec $(a, b, c) \in (\mathbb{R}_+^*)^3$ fixés.

6.4 Formule de Green-Riemann

Cette formule permet de calculer une intégrale double (par exemple celle qui définit une aire) sous la forme d'une intégrale curviligne (notée \oint) le long d'un chemin du plan.

Si \mathcal{D} est un sous-ensemble de \mathbb{R}^2 «géométriquement simple», son bord $\partial\mathcal{D}$ est une courbe fermée que nous orienterons dans le sens trigonométrique : ceci signifie intuitivement que si l'on se déplace sur $\partial\mathcal{D}$ dans ce sens, on a le sous-ensemble \mathcal{D} sur sa gauche. Nous noterons $\partial\mathcal{D}^+$ le bord de \mathcal{D} ainsi orienté.

Théorème 6.4.1 *Soit $P(x, y) dx + Q(x, y) dy$ une forme différentielle de classe C^1 définie sur un ouvert contenant le sous-ensemble géométriquement simple \mathcal{D} du plan. Alors, on a l'égalité :*

$$\oint_{\partial\mathcal{D}^+} P(x, y) dx + Q(x, y) dy = \iint_{\mathcal{D}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy$$

Démonstration: [LT 170-171]

□

En prenant en particulier $P(x, y) = -y$ et $Q(x, y) = x$, nous obtenons :

Corollaire 6.4.2 *L'aire d'un sous-ensemble géométriquement simple \mathcal{D} du plan est donnée par l'égalité :*

$$\mathcal{A}(\mathcal{D}) = \frac{1}{2} \oint_{\partial\mathcal{D}^+} x dy - y dx$$

On obtient grâce à cette formule l'aire délimitée par un *astroïde* (arc paramétré défini par $x(t) = \cos^3(t), y(t) = \sin^3(t), t \in [0, \pi/2]$) : elle vaut $3\pi/8$ [LT 172].

Si l'arc est paramétré en coordonnées polaires, *i.e.* par $x(\theta) = \rho(\theta) \cos \theta, y(\theta) = \rho(\theta) \sin \theta$, la forme différentielle $\frac{1}{2}x dy - y dx$ s'écrit $\frac{1}{2}\rho^2 d\theta$. On en déduit par exemple l'aire délimitée par la *lemniscate de Bernoulli* $\rho^2 = \cos(2\theta)$: celle-ci vaut 1 (cf. [LT 172] ou [GOU 342]).

6.5 Aire d'une nappe géométrique

6.5.1 Nappe paramétrée, nappe géométrique

Définition 6.5.1 On appelle C^k -nappe paramétrée de \mathbb{R}^3 tout couple (D, F) , où D est un domaine (*i.e.* un ouvert connexe) de \mathbb{R}^2 et $F \in C^k(D, \mathbb{R}^3)$. On dit que $F(D)$ est le support de la nappe ; c'est un connexe de \mathbb{R}^3 .

Le lecteur vérifiera que l'on définit une relation d'équivalence sur l'ensemble des C^k -nappes paramétrées de \mathbb{R}^3 par :

$$(D_1, F_1) \sim (D_2, F_2) \iff \exists \Phi : D_2 \rightarrow D_1, C^k\text{-difféomorphisme tel que } F_2 = F_1 \circ \Phi$$

Les classes d'équivalence sont appelées C^k -nappes géométriques de \mathbb{R}^3 .

Les représentants d'une C^k -nappe géométrique Σ sont appelés *paramétrisations* de cette nappe. C'est pourquoi un C^k -difféomorphisme Φ vérifiant la condition précédente est appelé *changement de paramétrisation*.

On vérifie immédiatement que toutes les paramétrisations d'une C^k -nappe géométrique Σ fixée ont même support, qui sera donc appelé support de cette nappe géométrique et noté $\text{supp } \Sigma$.

Exemple. Soit $D =]0, \pi[\times]0, 2\pi[$ et $F : D \rightarrow \mathbb{R}^3$ définie par

$$F(\theta, \phi) = (x(\theta, \phi), y(\theta, \phi), z(\theta, \phi)) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$$

Le couple (D, F) définit une C^∞ -nappe paramétrée de \mathbb{R}^3 dont le support est la sphère unité privée du demi-cercle correspondant à son intersection avec le demi-plan $y = 0, x \geq 0$.

Nous noterons S^2 la C^∞ -nappe géométrique correspondante.

Soit Σ une C^k -nappe géométrique. Le lecteur vérifiera que l'on définit une relation d'équivalence sur l'ensemble des triplets (D, F, X) tels que (D, F) est une paramétrisation de Σ et $X \subset D$ par :

$$(D_1, F_1, X_1) \sim (D_2, F_2, X_2) \iff \exists \Phi : D_2 \rightarrow D_1, C^k\text{-difféo. t.q. } F_2 = F_1 \circ \Phi \text{ et } X_1 = \Phi(X_2)$$

Définition 6.5.2 On appelle morceau \mathcal{X} de la C^k -nappe géométrique Σ une classe pour la relation d'équivalence précédente. On dit que $F(X)$, qui est une partie de $\text{supp } \Sigma$ indépendante de la paramétrisation choisie, est le support de ce morceau de nappe géométrique et on le note $\text{supp } \mathcal{X}$.

Un cas particulier important dans la suite est celui où, pour un représentant (D, F, X) du morceau \mathcal{X} , X est un compact de D . Cette propriété reste alors vraie pour n'importe quel représentant de \mathcal{X} , comme le lecteur le vérifiera à titre d'exercice.

Nous dirons dans ce cas que \mathcal{X} est un *morceau compact* de Σ .

Exemple : Fenêtre de Viviani

On appelle fenêtre de Viviani la partie de la sphère unité de \mathbb{R}^3 qui est incluse dans le solide (cylindrique) d'équation $x^2 + (y - \frac{1}{2})^2 \leq \frac{1}{4}$.

En reprenant les notations de l'exemple précédent, faire apparaître la fenêtre de Viviani comme un morceau compact \mathcal{V} de S^2 .

6.5.2 Aire d'une nappe géométrique

Théorème et définition 6.5.3 Soit \mathcal{X} un morceau compact d'une C^k -nappe géométrique Σ . Alors, pour tout représentant (D, F, X) de \mathcal{X} , la valeur de l'intégrale

$$\iint_X \left\| \frac{\partial F}{\partial u}(u, v) \wedge \frac{\partial F}{\partial v}(u, v) \right\| du dv$$

est indépendante du choix du représentant.

Nous l'appellerons aire du morceau compact \mathcal{X} et nous la noterons $\mathcal{A}(\mathcal{X})$.

Exercice : Aire de la fenêtre de Viviani

Réf : [LT 164]

Avec les notations de l'exemple précédent, établir l'égalité :

$$\mathcal{A}(\mathcal{V}) = 2\pi - 4$$

Chapitre 7

Espaces de Banach, suites et séries

Stefan Banach (1892-1945), mathématicien polonais, est un des fondateurs de *l'analyse fonctionnelle*, étude des espaces de fonctions. Comme d'habitude, nous noterons $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

7.1 Définition et premiers exemples

Définition 7.1.1 *On appelle espace de Banach tout espace vectoriel normé (e.v.n.) complet.*

Ceci signifie que l'espace E muni de la distance d associée à la norme $\|\cdot\|$ (par la relation $d(x, y) = \|x - y\|$) est un espace métrique complet.

Exemple. $(\mathbb{R}, |\cdot|)$ est un espace de Banach.

Rappel : Si (E_i, d_i) , $1 \leq i \leq n$ sont des espaces métriques, nous pouvons munir l'espace produit $E = \prod_{i=1}^n E_i$ des trois distances suivantes :

$$\delta_1(x, y) := \sum_{i=1}^n d_i(x_i, y_i), \quad \delta_2(x, y) := \left(\sum_{i=1}^n d_i^2(x_i, y_i) \right)^{\frac{1}{2}}, \quad \delta_\infty(x, y) := \max_{1 \leq i \leq n} d_i(x_i, y_i).$$

Ces trois distances sont équivalentes ($\delta_\infty \leq \delta_1 \leq n\delta_\infty$ et $\delta_\infty \leq \delta_2 \leq \sqrt{n}\delta_\infty$). Si tous les (E_i, d_i) , $1 \leq i \leq n$ sont complets, alors l'espace produit (E, d) , où $d \in \{\delta_1, \delta_2, \delta_3\}$, est complet.

Nous en déduisons immédiatement la proposition suivante.

Proposition 7.1.2 *Soient (E_i, N_i) , $1 \leq i \leq n$, des \mathbb{K} -espaces vectoriels normés. Nous pouvons munir l'espace produit $E = \prod_{i=1}^n E_i$ des trois normes suivantes, qui sont équivalentes :*

$$\|x\|_1 := \sum_{i=1}^n N_i(x_i) \quad , \quad \|x\|_2 := \left(\sum_{i=1}^n N_i^2(x_i) \right)^{\frac{1}{2}} \quad , \quad \|x\|_\infty := \max_{1 \leq i \leq n} N_i(x_i) \quad ,$$

où $x = (x_1, \dots, x_n)$ est un élément quelconque de l'espace produit E .
Si tous les (E_i, N_i) , $1 \leq i \leq n$ sont des espaces de Banach, alors l'espace produit $(E, \|\cdot\|)$, où $\|\cdot\|$ est l'une quelconque des trois normes précédentes, est un espace de Banach.

Exemples. En identifiant (par isométrie canonique) l'espace \mathbb{R}^2 muni de la norme euclidienne $\|\cdot\|_2$ et l'espace $(\mathbb{C}, |\cdot|)$, nous constatons que ce dernier est un espace de Banach. Plus généralement, les espaces vectoriels \mathbb{R}^n et \mathbb{C}^n munis d'une norme quelconque sont des espaces de Banach (rappelons ici que sur un espace de dimension finie, toutes les normes sont équivalentes). Plus généralement encore, tout \mathbb{K} -e.v.n. E de dimension finie est un espace de Banach. En effet, si (e_1, \dots, e_n) est une base de E , on peut munir cet espace de la norme définie par $\|x\| := \sum_{i=1}^n |x_i|$ pour $x = \sum_{i=1}^n x_i e_i$; ceci implique que l'application $\varphi : (\mathbb{K}^n, \|\cdot\|_1) \rightarrow (E, \|\cdot\|)$ définie par $\varphi(x_1, \dots, x_n) := \sum_{i=1}^n x_i e_i$ est une isométrie, ce qui nous permet de conclure.

Nous verrons un peu plus loin des exemples de Banach de dimension infinie. Avant de considérer des séries à valeurs dans un espace de Banach, nous consacrons la section suivante à des rappels sur les séries positives.

7.2 Séries à termes réels positifs

Dans toute cette section, nous noterons $(u_n) \in \mathbb{R}_+^{\mathbb{N}}$ et $(v_n) \in \mathbb{R}_+^{\mathbb{N}}$ des suites à valeurs réelles positives.

7.2.1 Définitions et théorèmes de comparaison

Définition 7.2.1 *Soit $(u_n) \in \mathbb{R}_+^{\mathbb{N}}$. On appelle série de terme général u_n et l'on note $\sum u_n$ la suite (S_n) définie par :*

$$\forall n \in \mathbb{N}, \quad S_n = \sum_{k=0}^n u_k.$$

On dit que S_n est la somme partielle d'ordre n associée à la série $\sum u_n$.

Définition 7.2.2 On dit que $\sum u_n$ converge si (S_n) converge. Dans ce cas,

- on appelle sa limite la somme de la série et on la note $\sum_{n=0}^{+\infty} u_n$;
- on appelle reste d'ordre n de la série le réel positif $R_n = \sum_{n=0}^{+\infty} u_n - S_n$, que l'on note encore $R_n = \sum_{k=n+1}^{+\infty} u_k$.

Nous avons la condition nécessaire de convergence suivante.

$$\sum u_n \text{ converge} \Rightarrow \lim u_n = 0. \quad (7.1)$$

Ce n'est pas une condition suffisante : un contre-exemple simple est donné par la *série harmonique* $\sum_{n \geq 1} 1/n$ qui diverge puisqu'on a pour cette série $S_{2n} - S_n \geq 1/2$.

Exemple. Soit $r \in \mathbb{R}_+$. La *série géométrique* $\sum r^n$ converge si, et seulement si, $r < 1$. Dans ce cas, sa somme est donnée par :

$$\sum_{n=0}^{+\infty} r^n = \frac{1}{1-r}.$$

Proposition 7.2.3 La série $\sum u_n$ converge si, et seulement si, la suite (S_n) de ses sommes partielles est majorée.

Corollaire 7.2.4 S'il existe $n_0 \in \mathbb{N}$ tel que, pour tout $n \geq n_0$, $0 \leq u_n \leq v_n$, alors nous avons l'implication : $\sum v_n$ converge $\Rightarrow \sum u_n$ converge.

Sous la même hypothèse, nous avons bien sûr, par contraposition, l'implication suivante : $\sum u_n$ diverge $\Rightarrow \sum v_n$ diverge.

Exemple. La série $\sum_{n \geq 1} 1/n^2 = \sum_{n \geq 0} 1/(n+1)^2 = 1 + \sum_{n \geq 1} 1/(n+1)^2$ converge car on a la majoration

$$\forall n \geq 1, \quad \frac{1}{(n+1)^2} \leq \frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1},$$

et les sommes partielles associées à ce dernier terme général se calculent facilement car elles sont « télescopiques » :

$$\sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k+1} \right) = 1 - \frac{1}{n+1} \xrightarrow{n \rightarrow +\infty} 1.$$

Application 7.2.5 Nous définissons deux sous-espaces vectoriels de l'espace des suites à valeurs réelles $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$ en posant :

$$l^1(\mathbb{N}) = \{(u_n) \in \mathbb{R}^{\mathbb{N}}, \sum |u_n| \text{ converge}\} ; \quad l^2(\mathbb{N}) = \{(u_n) \in \mathbb{R}^{\mathbb{N}}, \sum u_n^2 \text{ converge}\}.$$

(utiliser $(a+b)^2 \leq 2(a^2+b^2)$, pour tout $(a, b) \in \mathbb{R}^2$, afin de vérifier que $l^2(\mathbb{N})$ est bien stable par addition)

Grâce à l'implication (7.1) et au corollaire précédent, nous prouvons facilement l'inclusion $l^1(\mathbb{N}) \subset l^2(\mathbb{N})$.

Application 7.2.6 Grâce au corollaire précédent, on vérifie immédiatement la convergence de la série donnant le développement décimal d'un réel positif.

$$\forall x \in \mathbb{R}_+, \quad \exists (d_n) \in \mathbb{N}^{\mathbb{N}}, \quad \forall n \in \mathbb{N}^*, \quad 0 \leq d_n \leq 9 \quad \text{et} \quad x = \sum_{n=0}^{+\infty} \frac{d_n}{10^n}.$$

Rappelons que pour obtenir l'unicité de ce développement décimal, on interdit à la suite (d_n) de ne prendre que la valeur 9 à partir d'un certain rang.

Dans le théorème suivant, nous utilisons les notations de Landau.

Théorème 7.2.7 Si $u_n = O(v_n)$, c'est-à-dire si (u_n) est dominée par (v_n) , alors nous avons les propriétés suivantes.

- si $\sum v_n$ converge, alors $\sum u_n$ converge et $R_n^u = O(R_n^v)$;
- si $\sum u_n$ diverge, alors $\sum v_n$ diverge et $S_n^u = O(S_n^v)$.

Nous avons un énoncé analogue en remplaçant O par o .

Si nous avons l'équivalence $u_n \sim v_n$, alors les séries $\sum v_n$ et $\sum u_n$ sont de même nature (convergente ou divergente). Si elles convergent, alors $R_n^u \sim R_n^v$ (de limite nulle) ; si elles divergent, alors $S_n^u \sim S_n^v$ (de limite $+\infty$).

Exemple. (de Riemann) Par comparaison à la série harmonique, qui diverge, nous obtenons la divergence de la série $\sum_{n \geq 1} 1/n^\alpha$, pour tout $\alpha \leq 1$. Supposons désormais $\alpha > 1$. Grâce au développement limité, au voisinage de 0, $(1+x)^{1-\alpha} = 1 + (1-\alpha)x + o(x)$, nous prouvons l'équivalence :

$$\frac{1}{n^{\alpha-1}} - \frac{1}{(n+1)^{\alpha-1}} \sim \frac{\alpha-1}{n^\alpha}. \quad (7.2)$$

La série associée au terme général qui apparaît dans le membre de gauche est clairement convergente (les sommes partielles sont « télescopiques ») ; par

conséquent, il en est de même de la série $\sum_{n \geq 1} 1/n^\alpha$. Finalement, nous avons établi la condition nécessaire et suffisante suivante.

$$\sum_{n \in \mathbb{N}^*} \frac{1}{n^\alpha} \text{ converge} \iff \alpha > 1.$$

Notons enfin que, dans le cas $\alpha > 1$, si nous notons R_n le reste d'ordre n associé à la série $\sum_{n \geq 1} 1/n^\alpha$, nous obtenons l'équivalent

$$R_n \sim \frac{1}{(\alpha - 1)n^{\alpha-1}},$$

en utilisant l'équivalent (7.2).

Application 7.2.8 (Constante d'Euler) Pour tout $n \in \mathbb{N}^*$, notons H_n la somme partielle d'ordre n associée à la série harmonique. Nous allons montrer l'existence d'une constante $\gamma \in [0, 1]$ telle que $H_n = \ln n + \gamma + o(1)$. Il s'agit donc de prouver la convergence de la suite $(x_n)_{n \geq 1}$, que l'on définit par $x_n = H_n - \ln n$, vers une telle constante.

Or, la convergence de (x_n) équivaut à celle de la série $\sum_{n \geq 2} u_n$, où l'on a posé $u_n = x_n - x_{n-1}$, pour tout $n \geq 2$. Grâce à un développement limité de $\ln(1-x)$ au voisinage de 0, nous avons :

$$u_n = \frac{1}{n} - \ln n + \ln(n-1) = \frac{1}{n} + \ln\left(1 - \frac{1}{n}\right) = O\left(\frac{1}{n^2}\right)$$

ce qui nous permet de conclure par comparaison à l'exemple de Riemann. Par concavité du logarithme, nous constatons que $u_n \leq 0$, pour tout $n \geq 2$. Ainsi, la suite (x_n) est décroissante, d'où $\gamma = \lim x_n \leq x_1 = 1$. Il nous reste à remarquer que, par décroissance de $t \mapsto 1/t$ sur \mathbb{R}_+^* , nous avons :

$$\forall k \in \mathbb{N}^*, \quad \frac{1}{k} \geq \int_k^{k+1} \frac{dt}{t} \Rightarrow \forall n \in \mathbb{N}^*, \quad H_n \geq \int_1^{n+1} \frac{dt}{t} = \ln(n+1).$$

En particulier, la suite (x_n) est à valeurs positives, d'où $\gamma \geq 0$.

On ignore toujours actuellement si la constante d'Euler γ est un nombre rationnel ou non. Une valeur approchée est $\gamma \simeq 0,57722$.

7.2.2 Comparaison à une série géométrique

Proposition 7.2.9 (Règle de D'Alembert) Soit $(u_n) \in (\mathbb{R}_+^*)^{\mathbb{N}}$ telle que $u_{n+1}/u_n \rightarrow l$. Si $l < 1$, alors $\sum u_n$ converge. Si $l > 1$, alors $\sum u_n$ diverge.

Si $l = 1$, on ne peut conclure directement, comme on le constate sur l'exemple de Riemann.

Proposition 7.2.10 (Règle de Cauchy) Soit $(u_n) \in (\mathbb{R}_+^*)^{\mathbb{N}}$ telle que $u_n^{1/n} \rightarrow l$. Si $l < 1$, alors $\sum u_n$ converge. Si $l > 1$, alors $\sum u_n$ diverge.

Ici encore, si $l = 1$, on ne peut conclure directement, comme on le constate sur l'exemple de Riemann : $(n^{-\alpha})^{1/n} = n^{-\alpha/n} = \exp(-(\alpha \ln n)/n) \rightarrow 1$.

Remarque. Comme nous le démontrerons plus loin (cf. Lemme 11.2.5, page 200), nous avons l'implication suivante.

$$\frac{u_{n+1}}{u_n} \rightarrow l \Rightarrow u_n^{1/n} \rightarrow l.$$

Par conséquent, si la règle de Cauchy n'a pas permis de conclure, il est inutile d'essayer d'appliquer la règle de D'Alembert.

Dans le cas « douteux » de la règle de D'Alembert, on peut parfois conclure grâce à la règle suivante.

Proposition 7.2.11 (Règle de Raabe-Duhamel) Soit $(u_n) \in (\mathbb{R}_+^*)^{\mathbb{N}}$ t.q.

$$\exists(\alpha, \beta) \in \mathbb{R} \times]1, +\infty[, \quad \frac{u_{n+1}}{u_n} = 1 - \frac{\alpha}{n} + O\left(\frac{1}{n^\beta}\right).$$

Alors, il existe une constante $K > 0$ telle que $u_n \sim K/n^\alpha$. Par conséquent, $\sum u_n$ converge si, et seulement si, $\alpha > 1$.

Démonstration: On souhaite montrer l'existence d'une constante $K > 0$ telle que $n^\alpha u_n \rightarrow K$. En passant au logarithme, ceci revient à prouver que :

$$\exists k \in \mathbb{R}, \quad a_n := \alpha \ln n + \ln u_n \xrightarrow{n \rightarrow +\infty} k.$$

Cela revient encore à prouver la convergence de la série $\sum v_n$, où l'on a défini la suite $(v_n)_{n \in \mathbb{N}}$ par :

$$\forall n \in \mathbb{N}, \quad v_n := a_{n+1} - a_n = \alpha \ln\left(1 + \frac{1}{n}\right) + \ln \frac{u_{n+1}}{u_n}.$$

Or, au voisinage de 0, nous avons $\ln(1+x) = x + O(x^2)$, d'où :

$$v_n = \alpha \left(\frac{1}{n} + O\left(\frac{1}{n^2}\right) \right) - \frac{\alpha}{n} + O\left(\frac{1}{n^\beta}\right) + O\left(\frac{1}{n^2}\right) = O\left(\frac{1}{n^\gamma}\right),$$

où l'on a posé $\gamma = \min(\beta, 2) > 1$. Nous concluons par comparaison avec l'exemple de Riemann. \square

Application 7.2.12 (Formule de Stirling) Il existe $K > 0$ tel que :

$$n! \sim K e^{-n} n^{n+\frac{1}{2}}.$$

Démonstration: On souhaite montrer l'existence d'une constante $K > 0$ telle que $u_n := n! e^n n^{-n} \sim K n^{1/2}$.

En vue d'appliquer la règle de Raabe-Duhamel, nous calculons :

$$\frac{u_{n+1}}{u_n} = (n+1) e \left(1 + \frac{1}{n}\right)^{-n} (n+1)^{-1} = \exp \left(1 - n \ln \left(1 + \frac{1}{n}\right)\right). \quad (7.3)$$

Or, au voisinage de 0, nous avons le développement limité :

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

Nous en déduisons que :

$$1 - n \ln \left(1 + \frac{1}{n}\right) = 1 - n \left(\frac{1}{n} - \frac{1}{2n^2} + O\left(\frac{1}{n^3}\right)\right) = \frac{1}{2n} + O\left(\frac{1}{n^2}\right),$$

ce qui nous donne, d'après la formule (7.3),

$$\frac{u_{n+1}}{u_n} = \exp \left(\frac{1}{2n} + O\left(\frac{1}{n^2}\right)\right).$$

Nous utilisons à présent le développement limité, toujours au voisinage de 0, $\exp u = 1 + u + O(u^2)$, pour en déduire :

$$\frac{u_{n+1}}{u_n} = 1 + \frac{1}{2n} + O\left(\frac{1}{n^2}\right).$$

Il nous reste à appliquer la règle de Raabe-Duhamel, avec $\alpha = -1/2$, pour obtenir la conclusion désirée. \square

7.2.3 Série et intégrale impropre positives.

Le lecteur retrouvera ici plusieurs résultats qui ont déjà été mentionnés dans le chapitre 2.

La proposition suivante est issue de [MAD 175-176].

Proposition 7.2.13 Soit $a \in \mathbb{R}$ et soit $f \in C_M([a, +\infty[, \mathbb{R}_+)$.

Si l'intégrale impropre $\int_a^{+\infty} f$ converge, alors pour toute suite $(x_n) \in [a, +\infty[^\mathbb{N}$ de limite $+\infty$, la série $\sum \int_{x_n}^{x_{n+1}} f$ converge.

S'il existe une suite croissante $(x_n) \in [a, +\infty[^\mathbb{N}$, de limite $+\infty$ et telle que la série $\sum \int_{x_n}^{x_{n+1}} f$ converge, alors l'intégrale impropre $\int_a^{+\infty} f$ converge.

Application 7.2.14 L'intégrale impropre $\int_0^{+\infty} \left| \frac{\sin t}{t} \right| dt$ diverge.

Proposition 7.2.15 Soit $f \in C_M(\mathbb{R}_+, \mathbb{R}_+)$ une application décroissante et soit $(a_n)_{n \in \mathbb{N}}$ la suite définie par :

$$\forall n \in \mathbb{N}, \quad a_n := \sum_{k=0}^n f(k) - \int_0^n f.$$

Alors, la suite (a_n) converge. En particulier, la série $\sum f(n)$ et l'intégrale impropre $\int_0^{+\infty} f$ sont de même nature.

- Si $\int_0^{+\infty} f$ diverge, alors $S_n := \sum_{k=0}^n f(k) \sim \int_0^n f$;
- si $\int_0^{+\infty} f$ converge et si la convergence vers 0 de la suite des restes intégrales $(\int_n^{+\infty} f)$ est lente, alors $R_n := \sum_{k=n+1}^{+\infty} f \sim \int_n^{+\infty} f$.

Grâce à cette proposition, on retrouve l'existence de la constante d'Euler γ , ainsi que les équivalents suivants dans l'exemple de Riemann :

$$\text{si } \alpha > 1, \quad R_n \sim \frac{1}{(\alpha - 1)n^{\alpha-1}} ; \quad \text{si } 0 < \alpha < 1, \quad S_n \sim \frac{n^{1-\alpha}}{1 - \alpha}.$$

Exemple. (de Bertrand)

$$\sum_{n \geq 2} \frac{1}{n^\alpha (\ln n)^\beta} \text{ converge } \iff \alpha > 1 \text{ ou } (\alpha = 1 \text{ et } \beta > 1).$$

Démonstration: Le seul cas qui ne peut être traité directement par comparaison avec l'exemple de Riemann est celui où $\alpha = 1$. L'idée est alors d'utiliser une comparaison avec une intégrale, en appliquant la proposition précédente. En effet, par simple changement de variable $u = \ln t$, nous constatons que l'intégrale impropre $\int_2^{+\infty} \frac{dt}{t(\ln t)^\beta}$ converge si, et seulement si, $\beta > 1$. Cela nous permet de conclure. \square

7.3 Séries à valeurs dans un espace de Banach

Proposition 7.3.1 Soit $\sum u_n$ une série à valeurs dans un Banach $(E, \|\cdot\|)$. Alors, $\sum u_n$ converge si, et seulement si, elle satisfait le critère de Cauchy :

$$\forall \epsilon > 0, \quad \exists N \in \mathbb{N}, \quad q \geq p \geq N \Rightarrow \left\| \sum_{n=p+1}^q u_n \right\| < \epsilon.$$

Définition 7.3.2 Une série $\sum_{n \in \mathbb{N}} u_n$ à valeurs dans un e.v.n. $(E, \|\cdot\|)$ est dite absolument convergente si la série positive $\sum_{n \in \mathbb{N}} \|u_n\|$ converge.

Remarque. Cette notion est à distinguer de la *convergence normale* que nous définirons plus loin et qui concerne les séries d'applications à valeurs dans un Banach.

Corollaire 7.3.3 *Dans un espace de Banach $(E, \|\cdot\|)$, toute série absolument convergente est convergente.*

Démonstration: On applique le critère de Cauchy pour les séries et l'on conclut grâce à l'inégalité triangulaire :

$$\left\| \sum_{n=p}^q u_n \right\| \leq \sum_{n=p}^q \|u_n\|.$$

Remarque. En fait, cette propriété caractérise les Banach parmi les e.v.n. comme le montre la proposition suivante (hors programme). Le lecteur pourra en trouver la preuve dans BRIANE-PAGES *Théorie de l'intégration* (4ème édition) chez Vuibert, page 162.

Proposition 7.3.4 *Un \mathbb{K} -e.v.n. $(E, \|\cdot\|)$ est un Banach si, et seulement si, toute série absolument convergente dans $(E, \|\cdot\|)$ est convergente.*

Une série convergente mais non absolument convergente est dite *semi-convergente*. Pour mettre en évidence des séries semi-convergentes, on utilise souvent le résultat suivant.

Proposition 7.3.5 (Règle d'Abel) *Soit $(u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans un espace de Banach $(E, \|\cdot\|)$ et $(\alpha_n)_{n \in \mathbb{N}}$ une suite de réels. Pour que la série $\sum \alpha_n u_n$ converge, il suffit que les deux conditions suivantes soient satisfaites :*

- la suite $(S_n)_{n \in \mathbb{N}}$, avec $S_n := \sum_{k=0}^n u_k$, est bornée ;
- la suite $(\alpha_n)_{n \in \mathbb{N}}$ est décroissante et de limite nulle.

Bien que le résultat suivant se démontre directement, nous pouvons le faire apparaître ici comme un cas particulier de la règle d'Abel.

Corollaire 7.3.6 (Critère spécial des séries alternées) *Soit $(\alpha_n)_{n \in \mathbb{N}}$ une suite décroissante de réels qui tend vers 0. Alors, la série $\sum (-1)^n \alpha_n$ converge.*

Nous renvoyons le lecteur à [SK Ana. 47-48] pour la démonstration de la règle d'Abel, ainsi que divers exemples.

7.4 Espaces de Banach usuels de suites et de fonctions

Définition et proposition 7.4.1 On note $l_{\mathbb{K}}^1(\mathbb{N})$, resp. $l_{\mathbb{K}}^2(\mathbb{N})$, le \mathbb{K} -espace vectoriel des suites $(a_n) \in \mathbb{K}^{\mathbb{N}}$ telles que $\sum |a_n|$, resp. $\sum |a_n|^2$, converge. On définit une norme sur $l_{\mathbb{K}}^1(\mathbb{N})$, resp. $l_{\mathbb{K}}^2(\mathbb{N})$, en posant :

$$\|(a_n)\|_1 := \sum_{n=0}^{+\infty} |a_n|, \text{ resp. } \|(a_n)\|_2 := \left(\sum_{n=0}^{+\infty} |a_n|^2 \right)^{\frac{1}{2}}.$$

La démonstration est laissée au lecteur qui remarquera, en anticipant sur le Corollaire et définition 9.2.3, page 170, que $\|\cdot\|_2$ est une *norme euclidienne*.

Proposition 7.4.2 $(l_{\mathbb{K}}^1(\mathbb{N}), \|\cdot\|_1)$ et $(l_{\mathbb{K}}^2(\mathbb{N}), \|\cdot\|_2)$ sont des Banach.

Le lecteur trouvera la démonstration du résultat concernant $l_{\mathbb{K}}^2(\mathbb{N})$ dans [DAN 256-257]. Il adaptera la preuve au cas de $l_{\mathbb{K}}^1(\mathbb{N})$.

Définition et proposition 7.4.3 On note $l_{\mathbb{K}}^{\infty}(\mathbb{N})$ le \mathbb{K} -espace vectoriel des suites $(a_n) \in \mathbb{K}^{\mathbb{N}}$ bornées.

On définit une norme sur $l_{\mathbb{K}}^{\infty}(\mathbb{N})$ en posant $\|(a_n)\|_{\infty} := \sup_{n \in \mathbb{N}} |a_n|$.

Proposition 7.4.4 L'espace $(l_{\mathbb{K}}^{\infty}(\mathbb{N}), \|\cdot\|_{\infty})$ est un Banach.

Démonstration: Considérons une suite de Cauchy $(a^p)_{p \in \mathbb{N}}$ dans l'espace $(l_{\mathbb{K}}^{\infty}(\mathbb{N}), \|\cdot\|_{\infty})$; pour tout $p \in \mathbb{N}$, a^p désigne la suite $(a_n^p)_{n \in \mathbb{N}}$. En appliquant la définition d'une suite de Cauchy, nous avons :

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, q \geq p \geq N \Rightarrow \|a^q - a^p\|_{\infty} < \epsilon, \text{ i.e. } \sup_{n \in \mathbb{N}} |a_n^q - a_n^p| < \epsilon,$$

ce qui nous donne :

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, q \geq p \geq N \Rightarrow \forall n \in \mathbb{N}, |a_n^q - a_n^p| < \epsilon. \quad (7.4)$$

Soit $n \in \mathbb{N}$ arbitraire fixé. L'inégalité précédente prouve que $(a_n^p)_{p \in \mathbb{N}}$ est une suite de Cauchy à valeurs dans \mathbb{K} , qui est complet. Nous sommes donc en droit de définir $a_n := \lim_{p \rightarrow +\infty} a_n^p$. Posons $a := (a_n)_{n \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ et montrons que $a \in l_{\mathbb{K}}^{\infty}(\mathbb{N})$, c'est-à-dire que la suite (a_n) est bornée.

En prenant $\epsilon = 1$ dans l'inégalité (7.4), puis $p = N$, nous obtenons :

$$\forall q \geq N, \forall n \in \mathbb{N}, |a_n^q - a_n^N| < 1.$$

Pour $n \in \mathbb{N}$ arbitraire, par passage à la limite lorsque $q \rightarrow +\infty$, nous en tirons l'inégalité $|a_n - a_n^N| \leq 1$. La suite $(a_n^N)_{n \in \mathbb{N}}$ étant bornée, il est alors facile d'en déduire que la suite $(a_n)_{n \in \mathbb{N}}$ l'est également.

Nous allons maintenant montrer que, lorsque $p \rightarrow +\infty$, $\|a^p - a\|_\infty \rightarrow 0$, ce qui nous donnera la conclusion.

Pour $n \in \mathbb{N}$ arbitraire, par passage à la limite lorsque $q \rightarrow +\infty$ dans l'inégalité (7.4), nous obtenons :

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, p \geq N \Rightarrow |a_n - a_n^p| \leq \epsilon.$$

Comme n était arbitraire, il en résulte que :

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, p \geq N \Rightarrow \|a - a^p\|_\infty = \sup_{n \in \mathbb{N}} |a_n - a_n^p| \leq \epsilon.$$

□

Définition et proposition 7.4.5 Soit X un ensemble non vide et $(E, \|\cdot\|)$ un \mathbb{K} -e.v.n. Nous noterons $\mathcal{B}(X, E)$ le \mathbb{K} -espace vectoriel des applications bornées de X dans E .

On définit une norme sur $\mathcal{B}(X, E)$ en posant :

$$\forall f \in \mathcal{B}(X, E), \quad \|f\|_\infty := \sup_{x \in X} \|f(x)\|.$$

Proposition 7.4.6 Si $(E, \|\cdot\|)$ est un Banach, alors $(\mathcal{B}(X, E), \|\cdot\|_\infty)$ est un Banach.

Le lecteur trouvera dans [DAN 43-44] la démonstration de ce résultat, qui est assez similaire à la démonstration précédente sur $(l_\mathbb{K}^\infty(\mathbb{N}), \|\cdot\|_\infty)$.

On peut dire que $(\mathcal{B}(X, E), \|\cdot\|_\infty)$ « hérite » de la complétude de l'espace d'arrivée E , de même que $(l_\mathbb{K}^\infty(\mathbb{N}), \|\cdot\|_\infty)$ « hérite » de la complétude de \mathbb{K} .

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux \mathbb{K} -e.v.n. Nous noterons $\mathcal{L}(E, F)$, resp. $\mathcal{L}_c(E, F)$ le \mathbb{K} -espace vectoriel des applications linéaires, resp. linéaires continues de E dans F . Rappelons le théorème suivant dont le lecteur pourra retrouver la démonstration dans [DAN 167].

Théorème 7.4.7 Soit $f \in \mathcal{L}(E, F)$. Il y a équivalence entre :

1. f est lipschitzienne sur E ;
2. f est uniformément continue sur E ;
3. f est continue sur E ;
4. f est continue en 0 ;

5. f est bornée sur la boule unité fermée de E ;
6. f est bornée sur la sphère unité de E ;
7. Il existe une constante $M \geq 0$ telle que $\forall x \in E, \quad \|f(x)\|_F \leq M\|x\|_E$.

Définition et proposition 7.4.8 Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux \mathbb{K} -e.v.n. On définit une norme sur $\mathcal{L}_c(E, F)$ en posant :

$$\|u\| = \sup_{x \neq 0} \frac{\|u(x)\|_F}{\|x\|_E} = \sup_{\|x\|_E=1} \|u(x)\|_F.$$

On l'appelle norme subordonnée aux normes $\|\cdot\|_E$ et $\|\cdot\|_F$.

Exemple : Notons $\mathcal{M}_n(\mathbb{K})$ l'ensemble des matrices carrées d'ordre n à coefficients dans \mathbb{K} . Par abus d'écriture, nous identifierons un endomorphisme (nécessairement continu) de \mathbb{K}^n avec sa matrice représentative $M = [m_{i,j}] \in \mathcal{M}_n(\mathbb{K})$ dans la base canonique.

Si nous munissons \mathbb{K}^n de la norme $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$, nous allons montrer que la norme subordonnée de l'endomorphisme M vaut :

$$\|M\| = \max_{1 \leq i \leq n} \sum_{j=1}^n |m_{i,j}|.$$

En effet, pour tout $x \in \mathbb{K}^n$, nous avons :

$$\|Mx\| = \max_{1 \leq i \leq n} \left| \sum_{j=1}^n m_{i,j} x_j \right| \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |m_{i,j}| |x_j| \leq \|x\|_\infty \max_{1 \leq i \leq n} \sum_{j=1}^n |m_{i,j}|,$$

ce qui prouve l'inégalité :

$$\|M\| \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |m_{i,j}|.$$

Pour prouver l'inégalité inverse, considérons un indice $i_0 \in \llbracket 1, n \rrbracket$ tel que

$$\max_{1 \leq i \leq n} \sum_{j=1}^n |m_{i,j}| = \sum_{j=1}^n |m_{i_0,j}|,$$

puis $y \in \mathbb{K}^n$ tel que, pour tout $1 \leq j \leq n$, $|y_j| = 1$ et $m_{i_0,j} y_j = |m_{i_0,j}|$.

Nous concluons alors facilement en écrivant :

$$\|My\| \geq \left| \sum_{j=1}^n m_{i_0,j} y_j \right| = \sum_{j=1}^n |m_{i_0,j}|.$$

Proposition 7.4.9 Si $(F, \|\cdot\|_F)$ est un Banach, alors $(\mathcal{L}_c(E, F), \|\cdot\|)$ est un Banach.

Le lecteur trouvera dans [DAN 172-173] la preuve de ce résultat. Notons que c'est encore de la complétude de l'espace d'arrivée $(F, \|\cdot\|_F)$ qu'hérite $(\mathcal{L}_c(E, F), \|\cdot\|)$.

Pour clore cette section, nous renvoyons le lecteur à la Proposition 1.2.6, page 18, *i.e.* au théorème de prolongement d'une application linéaire continue à valeurs dans un Banach, qui nous a permis notamment de construire l'intégrale de Riemann.

7.5 Suites d'applications à valeurs dans un espace de Banach

Soit X un ensemble non vide et $(E, \|\cdot\|)$ un \mathbb{K} -e.v.n. Dans toute la suite, $(f_n)_{n \in \mathbb{N}}$ et f sont des applications définies sur X et à valeurs dans E . Rappelons d'abord quelques définitions générales, déjà vues en page 13.

Définitions 7.5.1 La suite $(f_n)_{n \in \mathbb{N}}$ converge simplement vers f sur X si :

$$\forall x \in X, \forall \epsilon > 0, \exists N_x \in \mathbb{N}, \forall n \geq N_x, \|f_n(x) - f(x)\| \leq \epsilon.$$

La suite d'applications $(f_n)_{n \in \mathbb{N}}$ converge uniformément vers f sur X si :

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, \forall x \in X, \|f_n(x) - f(x)\| \leq \epsilon.$$

Remarques.

1. La convergence simple de (f_n) vers f sur X signifie donc que pour tout $x \in X$, la suite $(f_n(x))_{n \in \mathbb{N}}$ converge vers $f(x)$ dans E .
2. La convergence uniforme est plus forte que la convergence simple car l'entier N ne dépend pas de x dans le cas de la convergence uniforme. Ainsi, si l'on prend $X =]0, 1[$, $E = \mathbb{R}$, $f_n = \mathbf{1}_{]0, 1/(n+1)[}$ et $f \equiv 0$, la suite $(f_n)_{n \in \mathbb{N}^*}$ converge simplement vers f sur \mathbb{R} mais ne converge pas uniformément vers f sur \mathbb{R} .

La norme $\|\cdot\|_\infty$ sur $\mathcal{B}(X, E)$ est encore appelée *norme de la convergence uniforme* à cause du résultat suivant, laissé au lecteur en exercice.

Proposition 7.5.2 Supposons que (f_n) est une suite dans $\mathcal{B}(X, E)$ qui converge uniformément vers f . Alors, $f \in \mathcal{B}(X, E)$ et $\|f_n - f\|_\infty \rightarrow 0$.

Définition 7.5.3 On dit que la suite d'applications (f_n) satisfait le critère de Cauchy uniforme sur l'ensemble X si :

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n \geq p \geq N, \forall x \in X, \|f_n(x) - f_p(x)\| \leq \epsilon.$$

Proposition 7.5.4 Si $(E, \|\cdot\|)$ est un Banach, alors la suite (f_n) converge uniformément sur X si, et seulement si, elle satisfait le critère de Cauchy uniforme sur X .

Le lecteur trouvera la démonstration de ce résultat dans [DAN 270-271]. Son intérêt est de caractériser la convergence uniforme sans avoir besoin de connaître à l'avance la limite f .

Nous nous plaçons désormais dans le cas où (X, d) est un espace métrique. Le théorème suivant est vrai dans le cas général où $(E, \|\cdot\|)$ est un \mathbb{K} -e.v.n. (il resterait même valable si (E, d') était un espace métrique).

Théorème 7.5.5 (de continuité) Soit $x_0 \in X$. Nous supposons que toutes les applications $f_n, n \in \mathbb{N}$, sont continues au point x_0 et que la suite (f_n) converge uniformément vers f sur X . Alors, f est continue au point x_0 .

Le lecteur trouvera la preuve de ce théorème dans [DAN 271].

Une conséquence immédiate de celui-ci est qu'une limite uniforme d'applications continues sur X est continue sur X .

Corollaire 7.5.6 Supposons que (X, d) est un espace métrique **compact** et que $(E, \|\cdot\|)$ est un Banach. Alors, $(C^0(X, E), \|\cdot\|_\infty)$ est un Banach.

Démonstration: Nous avons l'inclusion $C^0(X, E) \subset \mathcal{B}(X, E)$ puisque l'image du compact X par une application continue $f : X \rightarrow E$ est un compact de E , donc $f(X)$ est un sous-ensemble borné de E , autrement dit l'application f est bornée. Ainsi, la norme $\|\cdot\|_\infty$ est bien définie sur $C^0(X, E)$.

En outre, la proposition 7.4.6 nous apprend que $(\mathcal{B}(X, E), \|\cdot\|_\infty)$ est un Banach et le théorème précédent nous dit que $C^0(X, E)$ en est un sous-espace fermé, d'où la conclusion. \square

Théorème 7.5.7 (de dérivabilité) Nous nous plaçons maintenant dans le cas où $X = I$ est un intervalle réel non trivial (i.e. ni vide ni réduit à un point), $(E, \|\cdot\|)$ est un Banach et $f_n \in C^1(I, E)$ pour tout $n \in \mathbb{N}$.

Supposons que la suite (f'_n) converge uniformément sur I et qu'il existe $x_0 \in I$ tel que $(f_n(x_0))$ converge.

Alors, la suite (f_n) converge uniformément sur toute partie bornée de I vers $f \in C^1(I, E)$ telle que $f' = \lim f'_n$.

Idée de la preuve : Nos hypothèses entraînent :

$$\forall x \in I, \quad f_n(x) = f_n(x_0) + \int_{x_0}^x f'_n(t) \, dt.$$

Notons $l = \lim f_n(x_0)$ et g la limite uniforme de la suite (f'_n) puis définissons l'application f par :

$$\forall x \in I, \quad f(x) := l + \int_{x_0}^x g(t) \, dt.$$

On majore alors $\|f_n(x) - f(x)\|$ pour prouver la convergence uniforme sur tout sous-ensemble borné de I . \square

Remarque. Il suffit de prendre $I = E = \mathbb{R}$ et $f_n \equiv n$ pour constater que l'hypothèse « $\exists x_0 \in I$ tel que $(f_n(x_0))$ converge » est indispensable.

Nous avons aussi un résultat d'intégration relatif à la convergence uniforme, celui-ci résultant de la construction même de l'intégrale de Riemann.

Si $X = [a, b]$, $(E, \|\cdot\|)$ est un Banach et $(f_n), f$ sont toutes des éléments de $C_M([a, b], E)$ tels que (f_n) converge uniformément vers f sur $[a, b]$, alors

$$\int_a^b f_n \xrightarrow{n \rightarrow +\infty} \int_a^b f.$$

En effet, l'application $I : (C_M([a, b], E), \|\cdot\|_\infty) \rightarrow E$, $f \mapsto I(f) = \int_a^b f$ est par construction une forme linéaire continue.

Nous terminons cette section par deux conditions suffisantes d'uniforme convergence pour des suites d'applications à valeurs dans \mathbb{R} . [GOU 228]

Théorème 7.5.8 (premier théorème de Dini) *Soit X un espace métrique compact et $(f_n)_{n \in \mathbb{N}}$ une suite croissante (i.e. $f_n \leq f_{n+1}$, pour tout $n \in \mathbb{N}$) d'applications continues de X dans \mathbb{R} . On suppose que $(f_n)_{n \in \mathbb{N}}$ converge simplement vers f continue.*

Alors, $(f_n)_{n \in \mathbb{N}}$ converge uniformément vers f .

Remarque. On peut notamment utiliser ce théorème avec $X = [a, b]$ pour en déduire que $\lim \int_a^b f_n(x) \, dx = \int_a^b f(x) \, dx$. En ce sens, il est une version élémentaire du théorème de convergence monotone.

Théorème 7.5.9 (second théorème de Dini) *Soit $(f_n)_{n \in \mathbb{N}}$ une suite d'applications croissantes de $[a, b]$ dans \mathbb{R} (i.e. $\forall n \in \mathbb{N}, x \leq y \Rightarrow f_n(x) \leq f_n(y)$). On suppose que $(f_n)_{n \in \mathbb{N}}$ converge simplement vers f continue.*
Alors, $(f_n)_{n \in \mathbb{N}}$ converge uniformément vers f .

Remarque. Une application de ce théorème au calcul des probabilités est le Lemme 18.2.2, page 302.

7.6 Séries d'applications à valeurs dans un espace de Banach

Nous conservons les notations de la section précédente et nous supposons dans toute la suite que $(E, \|\cdot\|)$ est un Banach.

Pour tout $n \in \mathbb{N}$, définissons l'application $S_n : X \rightarrow E$ par $S_n := \sum_{k=0}^n f_k$.

Définition 7.6.1 *Nous dirons que la série d'applications $\sum_{n \in \mathbb{N}} f_n$ converge simplement (resp. uniformément) sur X si la suite d'applications (S_n) converge simplement (resp. uniformément) sur X*

La notion suivante n'avait pas de sens dans le cadre des suites d'applications.

Définition 7.6.2 *Nous dirons que la série d'applications $\sum_{n \in \mathbb{N}} f_n$ converge normalement sur X si la série positive $\sum_{n \in \mathbb{N}} \|f_n\|_\infty$ converge.*

Ceci équivaut à l'existence d'une suite $(a_n) \in (\mathbb{R}_+)^{\mathbb{N}}$ telle que la série $\sum_{n \in \mathbb{N}} a_n$ converge et $\forall x \in X, \|f_n(x)\| \leq a_n$.

Proposition 7.6.3 *Si la série $\sum_{n \in \mathbb{N}} f_n$ converge normalement sur X , alors elle converge uniformément sur X .*

Démonstration: Montrons que la suite (S_n) satisfait le critère de Cauchy uniforme sur X , ce qui suffira d'après la Proposition 7.5.4.

Puisque la série de réels positifs ou nuls $\sum_{n \in \mathbb{N}} \|f_n\|_\infty$ converge, elle satisfait le critère de Cauchy :

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, n \geq p \geq N \Rightarrow \sum_{k=p+1}^n \|f_k\|_\infty < \epsilon.$$

L'inégalité triangulaire nous permet d'en déduire que, pour $\epsilon > 0$ arbitraire :

$$\exists N \in \mathbb{N}, n \geq p \geq N \Rightarrow \forall x \in X, \|S_n(x) - S_p(x)\| \leq \sum_{k=p+1}^n \|f_k(x)\| < \epsilon.$$

□

Dans la section suivante, nous verrons un exemple important de série normalement convergente qui nous permet de définir l'exponentielle d'un endomorphisme continu sur l'espace de Banach E .

Les résultats (continuité, dérivabilité, intégration) établis pour des suites d'applications uniformément convergentes ont pour corollaires des énoncés analogues pour les séries d'applications qui convergent uniformément. Par exemple, le Théorème 7.5.7 nous donne immédiatement :

Théorème 7.6.4 *Soient I est un intervalle réel non trivial, $(E, \|\cdot\|)$ un Banach et $f_n \in C^1(I, E)$ pour tout $n \in \mathbb{N}$.*

Supposons que la série $\sum f'_n$ converge uniformément sur I et qu'il existe $x_0 \in I$ tel que $\sum f_n(x_0)$ converge.

Alors, la série $\sum f_n$ converge uniformément sur toute partie bornée de I vers $S \in C^1(I, E)$ telle que $S' = \sum_{n=0}^{+\infty} f'_n$.

Application. La fonction zêta de Riemann est définie par :

$$\forall s \in]1, +\infty[, \quad \zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}.$$

On démontre que $\zeta \in C^\infty([1, +\infty[, \mathbb{R})$. [SKA exo 6.8]

Enfin, en utilisant une transformation d'Abel, nous obtenons le résultat suivant [DAN 291].

Théorème 7.6.5 (Critère d'Abel uniforme) *Nous nous plaçons ici dans le cas où $X = I$ intervalle réel et, outre la suite (f_n) d'applications de I dans le Banach E , nous considérons une suite (g_n) d'applications de I dans \mathbb{R} . Nous supposons :*

1. $\exists M > 0, \forall n \in \mathbb{N}, \forall x \in I, \|\sum_{k=0}^n f_k(x)\| \leq M$;
2. la suite (g_n) converge uniformément sur I vers $g \equiv 0$ en décroissant.

Alors $\sum_{n \in \mathbb{N}} g_n f_n$ converge uniformément sur I .

Application aux séries trigonométriques : Soit $(c_n) \in \mathbb{R}^\mathbb{N}$ une suite décroissante de limite nulle. Alors la série $\sum_{n \in \mathbb{N}} c_n e^{int}$ d'applications de $[0, 2\pi]$ dans \mathbb{C} converge uniformément sur tout compact de $]0, 2\pi[$.

Si nous disposons d'une structure plus riche que celle d'espace de Banach, à savoir celle d'*algèbre de Banach* qui va être étudiée dans le chapitre suivant, la notion de séries d'applications va nous permettre de définir l'application *exponentielle* dans un cadre beaucoup plus général que celui de l'exponentielle réelle ou complexe.

Chapitre 8

Exponentielle dans une algèbre de Banach

Bibliographie :

FRA Oraux X-ENS Algèbre 2. Serge Francinou *et al.* Cassini. 2^e éd. 2009

SAV Algèbre linéaire (cours et exercices). Jean-Charles Savioz. Vuibert. 2003

8.1 Définitions et premières propriétés

Définition 8.1.1 Soit $(A, +, \times, \cdot)$ une algèbre et soit $\|\cdot\|$ une norme sur l'espace vectoriel $(A, +, \cdot)$. Nous dirons que $\|\cdot\|$ est une norme d'algèbre si :

$$\forall (u, v) \in A^2, \quad \|u \times v\| \leq \|u\| \|v\|. \quad (8.1)$$

Dans ce cas, $(A, +, \times, \cdot, \|\cdot\|)$ est appelée algèbre normée.

Exemple. Le lecteur vérifiera que l'on définit une norme d'algèbre sur $\mathcal{M}_n(\mathbb{K})$ en posant :

$$\forall M \in \mathcal{M}_n(\mathbb{K}) \quad \|M\| = n \max_{1 \leq i, j \leq n} |m_{i,j}|$$

Notons que pour cette norme, une suite $(M^k)_{k \in \mathbb{N}}$ d'éléments de $\mathcal{M}_n(\mathbb{K})$ converge vers $M \in \mathcal{M}_n(\mathbb{K})$ si et seulement si :

$$\forall (i, j) \in \llbracket 1; n \rrbracket^2 \quad m_{i,j}^k \xrightarrow[k \rightarrow +\infty]{} m_{i,j}$$

Comme $\mathcal{M}_n(\mathbb{K})$ est de dimension finie, cette propriété reste vraie pour toute autre norme sur l'espace $(\mathcal{M}_n(\mathbb{K}), +, \cdot)$.

Définition 8.1.2 Soit $(A, +, \times, \cdot, \|\cdot\|)$ une algèbre normée. Si $(A, +, \cdot, \|\cdot\|)$ est un espace de Banach, alors nous dirons que $(A, +, \times, \cdot, \|\cdot\|)$ est une algèbre de Banach.

Si nous revenons à l'exemple précédent, $(\mathcal{M}_n(\mathbb{K}), +, \times, \cdot, \|\cdot\|)$ est une algèbre de Banach puisque tout espace vectoriel normé de dimension finie est un espace de Banach.

Voyons maintenant un exemple en dimension (éventuellement) infinie.

Exemple. Soit $(E, +, \cdot, \|\cdot\|)$ un espace de Banach et $\mathcal{L}_c(E)$ le \mathbb{K} -espace vectoriel des endomorphismes continus de E .

D'après la proposition 7.4.9, $(\mathcal{L}_c(E), +, \cdot, \|\cdot\|)$ est aussi un espace de Banach. Or nous avons l'inégalité $\|u \circ v\| \leq \|u\| \times \|v\|$ puisque, pour tout $x \in E$,

$$\|u \circ v(x)\| = \|u(v(x))\| \leq \|u\| \times \|v(x)\| \leq \|u\| \times \|v\| \times \|x\|.$$

Ainsi, $(\mathcal{L}_c(E), +, \circ, \cdot, \|\cdot\|)$ est une algèbre de Banach.

Nous allons maintenant généraliser une notion déjà rencontrée dans le cadre des séries à valeurs réelles ou complexes.

Définition 8.1.3 Soient $\sum a_n$ et $\sum b_n$ deux séries à valeurs dans une algèbre de Banach $(A, +, \times, \cdot, \|\cdot\|)$. On appelle produit de Cauchy de ces deux séries la série $\sum c_n$ dont le terme général est défini par :

$$\forall n \in \mathbb{N} \quad c_n = \sum_{k=0}^n a_k \times b_{n-k}$$

Proposition 8.1.4 En reprenant les notations de la définition précédente, nous supposons que les séries $\sum a_n$ et $\sum b_n$ sont toutes les deux absolument convergentes. Alors la série produit $\sum c_n$ est absolument convergente et :

$$\sum_{n=0}^{+\infty} c_n = \left(\sum_{n=0}^{+\infty} a_n \right) \times \left(\sum_{n=0}^{+\infty} b_n \right)$$

Démonstration: Pour tout $n \in \mathbb{N}$, nous posons :

$$\Delta_n = \left(\sum_{i=0}^n a_i \right) \times \left(\sum_{j=0}^n b_j \right) - \sum_{k=0}^n c_k$$

Nous avons donc :

$$\Delta_n = \left(\sum_{i=0}^n a_i \right) \times \left(\sum_{j=0}^n b_j \right) - \sum_{k=0}^n \sum_{i+j=k} a_i \times b_j = \sum_{(i,j) \in I_n} a_i \times b_j ,$$

où $I_n = \{(i, j) \in \mathbb{N}^2 \mid 0 \leq i \leq n, 0 \leq j \leq n, i + j \geq n + 1\}$.

L'inégalité triangulaire nous permet d'en déduire :

$$\|\Delta_n\| \leq \sum_{(i,j) \in I_n} \|a_i\| \|b_j\| ,$$

ce qui s'écrit encore :

$$\|\Delta_n\| \leq \left(\sum_{i=0}^n \|a_i\| \right) \left(\sum_{j=0}^n \|b_j\| \right) - \sum_{k=0}^n \sum_{i+j=k} \|a_i\| \|b_j\|$$

Nous sommes alors ramenés au produit de Cauchy de séries à valeurs positives, ce qui nous permet de conclure. \square

Si nous notons I l'unité de notre algèbre de Banach $(A, +, \times, \cdot, \|\cdot\|)$, nous poserons, pour tout $u \in A$ et tout $n \in \mathbb{N}^*$:

$$u^n := \underbrace{u \times u \times \cdots \times u}_{n \text{ facteurs}} \quad \text{et} \quad u^0 = I$$

Enfin, pour tout $n \in \mathbb{N}$, nous définissons $f_n : A \rightarrow A$, $u \mapsto \frac{1}{n!} u^n$.

Nous constatons alors que $\sum_{n \in \mathbb{N}} f_n$ converge normalement sur toute partie bornée de A . Or chacune des applications f_n est continue sur A puisque, de façon générale, la condition (8.1) implique la continuité du produit dans une algèbre normée (Rappelons qu'une application bilinéaire $\phi : E_1 \times E_2 \rightarrow F$ est continue si, et seulement si, il existe $k \in \mathbb{R}_+$ tel que, $\forall (x_1, x_2) \in E_1 \times E_2$, $\|\phi(x_1, x_2)\|_F \leq k \|x_1\|_{E_1} \times \|x_2\|_{E_2}$). Ceci justifie :

Définition et proposition 8.1.5 Soient $(A, +, \times, \cdot, \|\cdot\|)$ une algèbre de Banach et $u \in A$. On définit l'exponentielle de u comme la somme de la série absolument convergente :

$$\exp(u) := \sum_{n=0}^{+\infty} \frac{1}{n!} u^n$$

L'application $\exp : A \rightarrow A$ est continue.

La notion de produit de Cauchy et la proposition 8.1.4 nous donnent :

Proposition 8.1.6 Soit $(u, v) \in A^2$; nous avons l'implication :

$$u \times v = v \times u \implies \exp(u + v) = \exp(u) \times \exp(v)$$

En particulier, pour tout $u \in A$, $\exp(u)$ est un élément inversible d'inverse $\exp(-u)$.

Démonstration: Puisque u et v commutent par hypothèse, nous sommes en droit d'appliquer la formule du binôme de Newton, si bien que :

$$\forall n \in \mathbb{N} \quad \frac{(u + v)^n}{n!} = \sum_{k=0}^n \frac{\binom{n}{k}}{n!} u^k v^{n-k} = \sum_{k=0}^n \frac{u^k}{k!} \frac{v^{n-k}}{(n-k)!}$$

Nous constatons alors que la série définissant $\exp(u + v)$ n'est autre que le produit de Cauchy des séries définissant $\exp(u)$ et $\exp(v)$, ce qui prouve l'égalité annoncée.

Enfin, comme u et $-u$ commutent, nous avons :

$$\exp(u) \exp(-u) = \exp(-u) \exp(u) = \exp(0) = I$$

□

Si u et v ne commutent pas, on peut avoir $\exp(u + v) \neq \exp(u) \times \exp(v)$, comme nous le constaterons dans la section suivante.

8.2 Exponentielle de matrice

Si nous munissons $\mathcal{M}_n(\mathbb{K})$ d'une norme d'algèbre $\|\cdot\|$ telle que celle de l'exemple page 155, ou encore telle que la norme subordonnée de l'exemple page 148, nous pouvons définir l'exponentielle dans l'algèbre de Banach $(\mathcal{M}_n(\mathbb{K}), +, \times, \cdot, \|\cdot\|)$.

Notons que cette définition est indépendante du choix de la norme puisque toutes les normes sont équivalentes, $\mathcal{M}_n(\mathbb{K})$ étant de dimension n^2 . D'après la Proposition 8.1.6, nous avons $\exp(\mathcal{M}_n(\mathbb{K})) \subset GL_n(\mathbb{K})$.

Exemple. Dans $\mathcal{M}_2(\mathbb{R})$, considérons $A = \begin{bmatrix} 0 & 0 \\ \theta & 0 \end{bmatrix}$ et $B = \begin{bmatrix} 0 & -\theta \\ 0 & 0 \end{bmatrix}$, où $\theta \in \mathbb{R}^*$. Comme $A^2 = B^2 = 0$, nous avons $\exp(A) = I_2 + A$ et $\exp(B) = I_2 + B$ d'où :

$$\exp(A) \times \exp(B) = \begin{bmatrix} 1 & -\theta \\ \theta & 1 - \theta^2 \end{bmatrix}.$$

D'autre part, nous calculons :

$$\forall n \in \mathbb{N}, \quad (A + B)^{2n} = \theta^{2n} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^{2n} = \theta^{2n} (-I_2)^n = (-\theta^2)^n I_2,$$

d'où $(A + B)^{2n+1} = \begin{bmatrix} 0 & (-1)^{n+1}\theta^{2n+1} \\ (-1)^n\theta^{2n+1} & 0 \end{bmatrix}$ et finalement :

$$\exp(A + B) = \sum_{n=0}^{+\infty} \frac{(A + B)^{2n}}{(2n)!} + \sum_{n=0}^{+\infty} \frac{(A + B)^{2n+1}}{(2n+1)!} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

Nous constatons que, pour tout $\theta \in \mathbb{R}^*$, $\exp(A + B) \neq \exp(A) \times \exp(B)$.

8.2.1 Calcul explicite

Pour pouvoir utiliser nos connaissances sur la réduction des endomorphismes, nous aurons besoin du lemme suivant :

Lemme 8.2.1 Soient $P \in GL_n(\mathbb{K})$ et $A \in \mathcal{M}_n(\mathbb{K})$; alors on a l'égalité :

$$\exp(P^{-1}AP) = P^{-1}\exp(A)P$$

Démonstration: Pour tout $k \in \mathbb{N}$, nous avons $(P^{-1}AP)^k = P^{-1}A^kP$ donc par linéarité :

$$\forall K \in \mathbb{N} \quad \sum_{k=0}^K \frac{(P^{-1}AP)^k}{k!} = P^{-1} \left(\sum_{k=0}^K \frac{A^k}{k!} \right) P$$

Or $M \mapsto P^{-1}MP$ est un endomorphisme de $\mathcal{M}_n(\mathbb{K})$, espace de dimension finie, ce qui implique sa continuité. Ceci nous permet de conclure en passant à la limite quand $K \rightarrow +\infty$ dans l'égalité précédente.

□

Matrice nilpotente

Soit $N \in \mathcal{M}_n(\mathbb{K})$ une matrice nilpotente d'indice¹ r . Le calcul de l'exponentielle de N se ramène alors à celui d'une somme finie :

$$\exp(N) = \sum_{k=0}^{r-1} \frac{N^k}{k!}$$

1. L'indice de N est le plus petit entier $q \in \mathbb{N}^*$ tel que $N^q = 0$

En particulier, si J est une matrice de Jordan d'ordre r (qui est aussi son indice), nous calculons immédiatement :

$$J = \begin{bmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \dots & 0 & 0 \end{bmatrix} \implies \exp(J) = \begin{bmatrix} 1 & 1 & \frac{1}{2!} & & \frac{1}{(r-1)!} \\ 0 & 1 & 1 & \frac{1}{2!} & \\ \vdots & & \ddots & \ddots & \\ \vdots & & & \ddots & \frac{1}{2!} \\ \vdots & & & & 1 \\ 0 & \dots & \dots & \dots & 0 & 1 \end{bmatrix}$$

Soit $N \in \mathcal{M}_n(\mathbb{C})$ une matrice nilpotente. La *réduction de Jordan* nous apprend qu'il existe une matrice de passage $P \in GL_n(\mathbb{C})$ telle que $P^{-1}NP$ s'écrive comme une matrice diagonale par blocs, ces blocs étant des matrices nilpotentes de Jordan J_1, \dots, J_m d'ordres respectifs n_1, \dots, n_m . On a bien sûr $n_1 + \dots + n_m = n$ et l'on peut noter qu'une matrice nilpotente de Jordan d'ordre 1 est nulle.

Le lemme 8.2.1 nous permet d'en déduire que la matrice $P^{-1} \exp(N) P$ est aussi diagonale par blocs, ceux-ci valant $\exp(J_1), \dots, \exp(J_m)$.

Ceci nous permet de calculer explicitement $\exp(N)$.

Matrice diagonalisable ou trigonalisable

Si $\Delta \in \mathcal{M}_n(\mathbb{K})$ est une matrice diagonale, on vérifie facilement en calculant les sommes partielles associées à la série qui définit l'exponentielle de Δ que :

$$\Delta = \text{Diag}(\lambda_1, \dots, \lambda_n) \implies \exp(\Delta) = \text{Diag}(\exp(\lambda_1), \dots, \exp(\lambda_n)) \quad (8.2)$$

Notons qu'un argument similaire montre que si $T \in \mathcal{M}_n(\mathbb{K})$ est une matrice triangulaire (supérieure, resp. inférieure) dont la diagonale est donnée par $(\lambda_1, \dots, \lambda_n)$, alors $\exp(T)$ est aussi une matrice triangulaire (supérieure, resp. inférieure) dont la diagonale est donnée par $(\exp(\lambda_1), \dots, \exp(\lambda_n))$.

Voici une application du lemme 8.2.1 dans le cas où $\mathbb{K} = \mathbb{C}$.

Proposition 8.2.2 *Soit $A \in M_n(\mathbb{C})$; nous noterons $\text{Sp}(A)$ son spectre, c'est-à-dire l'ensemble de ses valeurs propres, $\det A$ son déterminant et $\text{Tr}(A)$ sa trace, c'est-à-dire la somme de ses coefficients diagonaux. Alors,*

$$\text{Sp}(\exp(A)) = \exp(\text{Sp}(A)) \quad \text{et} \quad \det \exp(A) = \exp(\text{Tr}(A))$$

Démonstration: Comme \mathbb{C} est algébriquement clos, le polynôme caractéristique $\chi_A(X) := \det(XI_n - A)$ est *scindé* :

$$\chi_A(X) = \prod_{i=1}^n (X - \lambda_i) ,$$

avec $\lambda_1, \dots, \lambda_n$ complexes non nécessairement distincts.

Nous pouvons donc trigonaliser la matrice A : il existe une matrice triangulaire $T \in \mathcal{M}_n(\mathbb{C})$ de diagonale $(\lambda_1, \dots, \lambda_n)$ et une matrice inversible $P \in GL_n(\mathbb{C})$ telles que $P^{-1}AP = T$. D'après le lemme précédent, $\exp(A)$ et $\exp(T)$ sont des matrices semblables, cette dernière ayant pour diagonale $(\exp(\lambda_1), \dots, \exp(\lambda_n))$. Il est alors facile de conclure. \square

Si $D \in \mathcal{M}_n(\mathbb{K})$ est une matrice diagonalisable dans \mathbb{K} , il existe une matrice de passage $P \in GL_n(\mathbb{K})$ et une matrice diagonale $\Delta = \text{Diag}(\lambda_1, \dots, \lambda_n)$, où $\lambda_1, \dots, \lambda_n$ sont les racines – non nécessairement distinctes – du polynôme caractéristique χ_D , telles que $P^{-1}DP = \Delta$ donc $D = P\Delta P^{-1}$.

En utilisant l'égalité (8.2) et le lemme 8.2.1 (avec P^{-1}), nous en déduisons :

$$\exp(D) = P \text{Diag}(\exp(\lambda_1), \dots, \exp(\lambda_n)) P^{-1}$$

Le lecteur trouvera un tel calcul explicite dans [SAV 342-343].

Il existe néanmoins une autre méthode permettant d'éviter le calcul parfois fastidieux de la matrice de passage P . Pour en introduire l'idée, commençons par prouver le résultat suivant :

Lemme 8.2.3 *Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, il existe un polynôme $Q_A \in \mathbb{K}[X]$ tel que $\exp(A) = Q_A(A)$.*

Démonstration: Notons $\mathbb{K}[A]$ le sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ constitué par les polynômes en la matrice A .

Ce sous-espace est de dimension finie donc fermé dans $\mathcal{M}_n(\mathbb{K})$. Or $\exp(A)$ est par définition une limite d'éléments de $\mathbb{K}[A]$ donc $\exp(A) \in \mathbb{K}[A]$. \square

Remarque. Il n'existe pas de polynôme $Q \in \mathbb{K}[X]$ tel que $\exp(A) = Q(A)$ pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$.

En effet, si un tel polynôme existait, en prenant $A = \lambda I_n$ avec $\lambda \in \mathbb{R}$, nous obtiendrions $\exp(\lambda)I_n = Q(\lambda)I_n$ d'où $\exp(\lambda) = Q(\lambda)$ pour tout $\lambda \in \mathbb{R}$, ce qui est absurde par croissances comparées.

Supposons que $D \in \mathcal{M}_n(\mathbb{K})$ est une matrice diagonalisable et écrivons-la sous la forme $D = P\Delta P^{-1}$ comme précédemment. D'après le lemme précédent, il existe un polynôme $Q \in \mathbb{K}[X]$ tel que $\exp(\Delta) = Q(\Delta)$, c'est-à-dire :

$$\text{Diag}(\exp(\lambda_1), \dots, \exp(\lambda_n)) = \text{Diag}(Q(\lambda_1), \dots, Q(\lambda_n))$$

Mais dans ce cas particulier d'une matrice diagonale, il est facile d'exhiber un tel polynôme Q : Si $\text{Sp}(D) = \{\mu_1, \dots, \mu_k\}$ avec $k \leq n$, autrement dit si $\{\lambda_1, \dots, \lambda_n\} = \{\mu_1, \dots, \mu_k\}$ où les μ_j , $1 \leq j \leq k$ sont deux à deux distincts, nous pouvons prendre le *polynôme d'interpolation de Lagrange* :

$$Q = \sum_{j=1}^k \exp(\mu_j) l_j \quad \text{où} \quad l_j(X) = \prod_{\substack{1 \leq i \leq k \\ i \neq j}} \frac{X - \mu_i}{\mu_j - \mu_i}$$

Nous avons alors :

$$Q(D) = Q(P\Delta P^{-1}) = PQ(\Delta)P^{-1} = P \exp(\Delta)P^{-1} = \exp(D),$$

ce qui nous permet de calculer explicitement $\exp(D)$.

Exemple Retrouvons par cette méthode la valeur de $\exp M_\theta$, avec :

$$M_\theta = \begin{bmatrix} 0 & -\theta \\ \theta & 0 \end{bmatrix},$$

cette exponentielle ayant déjà été calculée dans l'exemple page 158.

Nous trouvons facilement $\text{Sp}(M_\theta) = \{i\theta, -i\theta\}$ et $Q(X) = \frac{\sin \theta}{\theta} X + \cos \theta$, d'où :

$$\exp(M_\theta) = Q(M_\theta) = \frac{\sin \theta}{\theta} M_\theta + \cos \theta I_2 = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

Cas général

Nous avons traité jusqu'à présent le calcul explicite de l'exponentielle d'une matrice nilpotente et d'une matrice diagonalisable. Rappelons maintenant un important résultat d'algèbre linéaire qui va nous permettre de passer au cas général :

Théorème 8.2.4 (Décomposition de Dunford) *Soit $M \in \mathcal{M}_n(\mathbb{K})$ une matrice dont le polynôme caractéristique est scindé (toujours vrai si $\mathbb{K} = \mathbb{C}$). Alors il existe un unique couple $(D, N) \in (\mathcal{M}_n(\mathbb{K}))^2$ tel que D est diagonalisable, N est nilpotente, $DN = ND$ et $M = D + N$. En outre, D et N sont des polynômes en M .*

Puisque D et N commutent, nous sommes en droit d'écrire $\exp(M) = \exp(D) \exp(N)$ et nous sommes ainsi ramenés aux méthodes précédemment décrites pour faire le calcul explicite.

8.2.2 Quelques applications topologiques

Définition 8.2.5 *Nous dirons qu'une matrice $M \in \mathcal{M}_n(\mathbb{K})$ est unipotente si $M = I_n + N$ avec N matrice nilpotente.*

Proposition 8.2.6 *Notons \mathcal{N} , respectivement \mathcal{U} , l'ensemble des matrices $M \in \mathcal{M}_n(\mathbb{K})$ qui sont nilpotentes, respectivement unipotentes. L'exponentielle induit un homéomorphisme de \mathcal{N} sur \mathcal{U} .*

Nous renvoyons le lecteur à [FRA 245]. Il s'agit de montrer que l'application $\psi : N \mapsto \exp(N) - I_n$ est un homéomorphisme de \mathcal{N} sur \mathcal{N} .

Notons que cette application est en fait polynomiale puisque :

$$\forall N \in \mathcal{N} \quad \psi(N) = \sum_{k=1}^{n-1} \frac{N^k}{k!}$$

L'idée est, par analogie avec le développement en série entière de $\ln(1+x)$, de définir l'application $\varphi : \mathcal{N} \rightarrow \mathcal{N}$ par :

$$\forall N \in \mathcal{N} \quad \varphi(N) = \sum_{k=1}^{n-1} (-1)^{k-1} \frac{N^k}{k}$$

puis de montrer que $\varphi \circ \psi = \psi \circ \varphi = \text{Id}_{\mathcal{N}}$.

Nous pouvons en déduire la proposition suivante :

Proposition 8.2.7 *L'application $\exp : \mathcal{M}_n(\mathbb{C}) \rightarrow GL_n(\mathbb{C})$ est surjective.*

Démonstration: Soit $B \in GL_n(\mathbb{C})$; admet-elle un antécédent dans $\mathcal{M}_n(\mathbb{C})$ par l'application exponentielle ? Par décomposition de Jordan, il suffit de répondre à cette question dans le cas où B est un bloc de Jordan de la forme $B = \lambda(I_n + N)$, avec $\lambda \in \mathbb{C}^*$ et N nilpotente.

D'une part, nous savons que $\exp : \mathbb{C} \rightarrow \mathbb{C}^*$ est surjective, ce qui nous donne l'existence de $\mu \in \mathbb{C}$ tel que $\exp(\mu) = \lambda$.

D'autre part, la proposition précédente nous donne l'existence d'une matrice $M \in \mathcal{M}_n(\mathbb{C})$ telle que $\exp(M) = I_n + N$.

Nous constatons alors que $\exp(\mu I_n + M) = B$.

□

Remarque. En revanche, l'application $\exp : \mathcal{M}_n(\mathbb{R}) \rightarrow GL_n(\mathbb{R})$ n'est pas surjective. En effet, si elle l'était, $GL_n(\mathbb{R})$ serait connexe par continuité de l'exponentielle, l'espace $\mathcal{M}_n(\mathbb{R})$ étant bien sûr connexe. En considérant cette fois l'application continue et surjective $\det : GL_n(\mathbb{R}) \rightarrow \mathbb{R}^*$, on en déduirait que \mathbb{R}^* est connexe, ce qui est absurde.

Proposition 8.2.8 *Notons $\mathcal{S}_n(\mathbb{R})$, resp. $\mathcal{S}_n^{++}(\mathbb{R})$, le sous-ensemble de $\mathcal{M}_n(\mathbb{R})$ constitué par les matrices symétriques, resp. symétriques définies positives. L'exponentielle induit un homéomorphisme de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^{++}(\mathbb{R})$.*

Démonstration: Par continuité de la transposition (qui est un endomorphisme de $\mathcal{M}_n(\mathbb{R})$), nous vérifions facilement que $\exp({}^t M) = {}^t \exp(M)$ donc $\exp(\mathcal{S}_n(\mathbb{R})) \subset \mathcal{S}_n(\mathbb{R})$.

En outre, d'après la proposition 8.2.2 (en plongeant \mathbb{R} dans \mathbb{C}), nous avons $\det(\exp(M)) = \exp(\text{Tr}(M)) > 0$ d'où $\exp(\mathcal{S}_n(\mathbb{R})) \subset \mathcal{S}_n^{++}(\mathbb{R})$.

Si $A \in \mathcal{S}_n^{++}(\mathbb{R})$, nous savons qu'il existe une matrice orthogonale P et une matrice diagonale $\Delta = \text{Diag}(\lambda_1, \dots, \lambda_n)$, avec $(\lambda_1, \dots, \lambda_n) \in (\mathbb{R}_+^*)^n$ telles que $A = P\Delta P^{-1}$. En posant $M = P \text{Diag}(\ln \lambda_1, \dots, \ln \lambda_n) P^{-1} \in \mathcal{S}_n(\mathbb{R})$, nous constatons que $\exp(M) = A$.

Il reste à montrer que M est l'unique antécédent de A par l'application exponentielle pour prouver que celle-ci est une bijection de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^{++}(\mathbb{R})$. Enfin, il reste à montrer que sa bijection réciproque est continue. Nous renvoyons le lecteur à [FRA 243-244]. \square

8.2.3 Application aux systèmes différentiels linéaires

Une application importante de l'exponentielle de matrice est la résolution des systèmes différentiels linéaires du premier ordre à coefficients constants, comme nous allons le voir maintenant.

Proposition 8.2.9 *Soit $M \in \mathcal{M}_n(\mathbb{K})$ et I un intervalle réel non trivial. L'application $f : I \rightarrow \mathcal{M}_n(\mathbb{K})$ définie par $f(t) = \exp(tM)$ est de classe C^1 sur I et vérifie :*

$$\forall t \in I \quad f'(t) = Mf(t) = f(t)M$$

Démonstration: Nous appliquons le théorème 7.6.4 avec $f_n : I \rightarrow \mathcal{M}_n(\mathbb{K})$ définie par $f_n(t) = \frac{t^n}{n!} M^n$. Nous avons bien $f_n \in C^1(I, \mathcal{M}_n(\mathbb{K}))$, $f'_0 \equiv 0$ et :

$$\forall n \in \mathbb{N}^* \quad \forall t \in I \quad f'_n(t) = \frac{t^{n-1}}{(n-1)!} M^n = f_{n-1}(t) M = M f_{n-1}(t)$$

Nous en déduisons facilement que la série $\sum f'_n$ converge normalement (donc uniformément) sur tout compact de I . En outre, nous savons déjà que la série $\sum f_n(t)$ converge en tout point $t \in I$ vers $f(t) = \exp(tM)$. Nous en déduisons (la dérivabilité étant une notion locale) que $f \in C^1(I, \mathcal{M}_n(\mathbb{K}))$ et que :

$$\forall t \in I \quad f'(t) = \sum_{n=1}^{+\infty} f'_n(t) = \sum_{n=0}^{+\infty} f_n(t) M = \sum_{n=0}^{+\infty} M f_n(t)$$

Il nous reste à remarquer que le produit de matrices est bilinéaire continu dans l'algèbre de Banach $(\mathcal{M}_n(\mathbb{K}), +, \times, \cdot, \|\cdot\|)$ pour en déduire les égalités de l'énoncé par passage à la limite quand $N \rightarrow +\infty$ dans :

$$\sum_{n=0}^N f_n(t) M = \left(\sum_{n=0}^N f_n(t) \right) M \quad \text{et} \quad \sum_{n=0}^N M f_n(t) = M \left(\sum_{n=0}^N f_n(t) \right)$$

□

Dans le théorème suivant, nous identifierons par abus d'écriture \mathbb{K}^n et $\mathcal{M}_{n,1}(\mathbb{K})$.

Théorème 8.2.10 Soient I un intervalle réel non trivial, $t_0 \in I$, $A \in \mathcal{M}_n(\mathbb{K})$ et $B \in C^0(I, \mathcal{M}_{n,1}(\mathbb{K}))$. Alors les solutions du système différentiel :

$$Y' = AY + B$$

sont exactement les applications de la forme :

$$\forall t \in I \quad Y(t) = \int_{t_0}^t \exp[(t-u)A] B(u) du + \exp(tA) v \quad ,$$

où $v \in \mathbb{K}^n$ est arbitraire.

Démonstration: En pensant à la méthode de *variation de la constante*, nous posons le changement de variable suivant :

$$Y(t) = \exp(tA)Z(t) \iff Z(t) = \exp(-tA)Y(t)$$

D'après la proposition précédente, nous avons :

$$Y'(t) = A \exp(tA)Z(t) + \exp(tA)Z'(t)$$

si bien que notre système différentiel se réécrit : $\exp(tA)Z'(t) = B(t)$. Les solutions de cette équation différentielle sont de la forme

$$Z(t) = \int_{t_0}^t \exp(-uA)B(u) du + v \quad ,$$

où $v \in \mathbb{K}^n$ est arbitraire, ce qui nous permet de conclure par linéarité² de l'intégrale. \square

Remarque. Soient $A \in \mathcal{M}_n(\mathbb{K})$ une matrice diagonalisable et V_1, \dots, V_n une base de \mathbb{K}^n constituée de vecteurs propres associés aux valeurs propres $\lambda_1, \dots, \lambda_n$ (non nécessairement distinctes).

Il est bien connu que la solution générale du système différentiel homogène $Y' = AY$ est de la forme suivante, avec $(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n$:

$$Y(t) = \alpha_1 \exp(\lambda_1 t) V_1 + \dots + \alpha_n \exp(\lambda_n t) V_n$$

Nous pouvons retrouver ce résultat grâce à l'exponentielle de matrice. En effet, si nous notons P la matrice de passage constituée par les vecteurs colonnes V_1, \dots, V_n , nous avons $P^{-1}AP = \text{diag}(\lambda_1, \dots, \lambda_n)$ et donc, pour tout $t \in I$, $\exp(tA) = P \text{diag}(\exp(\lambda_1 t), \dots, \exp(\lambda_n t)) P^{-1}$.

D'après le théorème précédent, la solution générale du système homogène s'écrit donc $P \text{diag}(\exp(\lambda_1 t), \dots, \exp(\lambda_n t)) P^{-1}v$, avec $v \in \mathbb{K}^n$ quelconque. Or $P \text{diag}(\exp(\lambda_1 t), \dots, \exp(\lambda_n t))$ est la matrice constituée par les colonnes $\exp(\lambda_1 t) V_1, \dots, \exp(\lambda_n t) V_n$; avec le changement de paramètre $\alpha = P^{-1}v$, nous retrouvons donc le résultat annoncé.

2. au sens large : $M \int_a^b N(u) du = \int_a^b MN(u) du$ où $M \in \mathcal{M}_n(\mathbb{K})$ et $N \in C^0([a, b], \mathbb{K}^n)$. Ceci se démontre en passant aux coordonnées dans la base canonique de \mathbb{K}^n .

Chapitre 9

Espaces préhilbertiens

9.1 Produit scalaire sur un \mathbb{K} -espace vectoriel.

Le mot *scalaire* renvoyant à un élément du corps \mathbb{K} , nous allons d'abord considérer le cas $\mathbb{K} = \mathbb{R}$, puis — dans une seconde définition — nous traiterons le cas $\mathbb{K} = \mathbb{C}$.

Définition 9.1.1 *Soit E un \mathbb{R} -espace vectoriel. On appelle produit scalaire sur E toute application $(x, y) \mapsto \langle x, y \rangle$ de $E \times E$ dans \mathbb{R} qui est :*

- symétrique : $\forall (x, y) \in E^2, \quad \langle x, y \rangle = \langle y, x \rangle;$
- linéaire à droite :

$$\forall (x, y_1, y_2) \in E^3, \forall (\lambda_1, \lambda_2) \in \mathbb{R}^2, \quad \langle x, \lambda_1 y_1 + \lambda_2 y_2 \rangle = \lambda_1 \langle x, y_1 \rangle + \lambda_2 \langle x, y_2 \rangle;$$

- définie positive : $\forall x \in E \setminus \{0\}, \quad \langle x, x \rangle > 0.$

Remarque. Les deux premières propriétés impliquant la linéarité à gauche, nous pouvons dire de façon équivalente qu'un produit scalaire sur un \mathbb{R} -espace vectoriel est une forme bilinéaire symétrique définie positive.

Définition 9.1.2 *Soit E un \mathbb{C} -espace vectoriel. On appelle produit scalaire sur E toute application $(x, y) \mapsto \langle x, y \rangle$ de $E \times E$ dans \mathbb{C} qui est :*

- à symétrie hermitienne : $\forall (x, y) \in E^2, \quad \langle x, y \rangle = \overline{\langle y, x \rangle};$
- linéaire à droite :

$$\forall (x, y_1, y_2) \in E^3, \quad \forall (\lambda_1, \lambda_2) \in \mathbb{C}^2, \quad \langle x, \lambda_1 y_1 + \lambda_2 y_2 \rangle = \lambda_1 \langle x, y_1 \rangle + \lambda_2 \langle x, y_2 \rangle;$$

- définie positive : $\forall x \in E \setminus \{0\}, \quad \langle x, x \rangle > 0.$

Remarque. Les 2 premières propriétés impliquant la *semi-linéarité à gauche* :

$$\forall (x_1, x_2, y) \in E^3, \quad \forall (\lambda_1, \lambda_2) \in \mathbb{C}^2, \quad \langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle = \bar{\lambda}_1 \langle x_1, y \rangle + \bar{\lambda}_2 \langle x_2, y \rangle,$$

nous pouvons dire de façon équivalente qu'un produit scalaire sur un \mathbb{C} -espace vectoriel est une forme *sesquilinéaire*¹ hermitienne définie positive.

Définition 9.1.3 On appelle \mathbb{K} -espace préhilbertien tout \mathbb{K} -espace vectoriel muni d'un produit scalaire.

Exemples :

- $\mathbb{K} = \mathbb{C}, \quad E = \mathbb{C}^n, \quad \langle x, y \rangle = \sum_{i=1}^n \bar{x}_i y_i.$
- $\mathbb{K} = \mathbb{C}, \quad E = C^0([0, 1], \mathbb{C}), \quad \langle f, g \rangle = \int_0^1 \overline{f(t)} g(t) dt.$
- $\mathbb{K} = \mathbb{R}, \quad E = l_{\mathbb{R}}^2(\mathbb{N}), \quad \langle (u_n), (v_n) \rangle = \sum_{n=0}^{+\infty} u_n v_n$, la convergence de cette série résultant de l'inégalité $|u_n v_n| \leq (u_n^2 + v_n^2)/2$.
- $\mathbb{K} = \mathbb{R}, \quad E = C^0([a, b], \mathbb{R}), \quad \langle f, g \rangle = \int_{]a, b[} f g \omega$, où $\omega \in C^0(]a, b[, \mathbb{R}_+^*)$ est intégrable sur l'intervalle $]a, b[$.

L'exemple suivant de \mathbb{C} -espace préhilbertien joue un rôle important dans la théorie des séries de Fourier, que nous aborderons dans le chapitre suivant.

Proposition 9.1.4 Notons \mathcal{D} le \mathbb{C} -sous-espace vectoriel de $C_M(\mathbb{R}, \mathbb{C})$ des applications 2π -périodiques vérifiant :

$$\forall x \in \mathbb{R}, \quad f(x) = \frac{f(x-) + f(x+)}{2}.$$

On définit un produit scalaire sur \mathcal{D} en posant :

$$\forall (f, g) \in \mathcal{D}^2, \quad \langle f, g \rangle := \frac{1}{2\pi} \int_0^{2\pi} \overline{f(t)} g(t) dt.$$

Le \mathbb{C} -espace \mathcal{D} muni de ce produit scalaire est appelé espace préhilbertien de Dirichlet.

Démonstration: La seule difficulté est de vérifier que $\langle f, f \rangle = 0 \Rightarrow f \equiv 0$. Si f est continue, c'est immédiat mais dans le cas général, il faut couper l'intégrale le long de la subdivision par la relation de Chasles puis utiliser $f(a_i) = \frac{1}{2}(f(a_i+) + f(a_i-))$ en tout point a_i de la subdivision. Le lecteur pourra consulter [DAN 377-378] pour les détails.

1. du latin *sesqui* qui signifie « un et demi ».

□

Remarque. À toute application $f \in C_M(\mathbb{R}, \mathbb{C})$, 2π -périodique, nous pouvons associer sa « régularisée » $\tilde{f} \in \mathcal{D}$ définie par :

$$\forall x \in \mathbb{R}, \quad \tilde{f}(x) := \frac{f(x-) + f(x+)}{2}. \quad (9.1)$$

Sur une période donnée, par exemple $[0, 2\pi]$, f et \tilde{f} ne diffèrent qu'en un nombre fini de points, inclus dans l'ensemble des discontinuités de f .

9.2 Norme euclidienne, norme hermitienne

Soit E un \mathbb{K} -espace préhilbertien. Notons, pour tout $x \in E$, $\|x\| := \sqrt{\langle x, x \rangle}$. Cette notation sera justifiée ci-dessous en prouvant que $\|\cdot\|$ est une norme sur E .

Proposition 9.2.1 (Inégalité de Cauchy-Schwarz) *Dans le \mathbb{K} -espace préhilbertien E , nous avons l'inégalité :*

$$\forall (x, y) \in E^2, \quad |\langle x, y \rangle| \leq \|x\| \|y\|,$$

avec égalité si, et seulement si, les vecteurs x et y sont liés.

Démonstration: Pour $\mathbb{K} = \mathbb{R}$, nous définissons l'application $P : \mathbb{R} \rightarrow \mathbb{R}_+$ par $P(\lambda) := \|\lambda x + y\|^2 = \langle \lambda x + y, \lambda x + y \rangle = \lambda^2 \|x\|^2 + 2\lambda \langle x, y \rangle + \|y\|^2$. Ainsi, P est une application polynomiale de degré au plus 2 qui ne prend que des valeurs positives ou nulles. Nous en déduisons que son discriminant (réduit) $\Delta' = \langle x, y \rangle^2 - \|x\|^2 \|y\|^2$ est négatif ou nul, ce qui nous permet de conclure.

Pour $\mathbb{K} = \mathbb{C}$, nous considérons $(x, y) \in E^2$ arbitraire fixé et nous choisissons $(\rho, \theta) \in \mathbb{R}_+ \times \mathbb{R}$ tel que $\langle x, y \rangle = \rho e^{i\theta}$. Nous définissons alors $P : \mathbb{R} \rightarrow \mathbb{R}_+$ par $P(\lambda) := \|\lambda e^{i\theta} x + y\|^2 = \lambda^2 \|x\|^2 + 2\lambda \rho + \|y\|^2$.

Nous concluons alors comme dans le cas réel. □

Proposition 9.2.2 (Inégalité de Minkowski) *Dans le \mathbb{K} -espace préhilbertien E , nous avons l'inégalité :*

$$\forall (x, y) \in E^2, \quad \|x + y\| \leq \|x\| + \|y\|.$$

Démonstration: Pour $\mathbb{K} = \mathbb{R}$, l'inégalité de Cauchy-Schwarz nous donne :

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

d'où la conclusion. Pour $\mathbb{K} = \mathbb{C}$, nous utilisons :

$$\|x + y\|^2 = \|x\|^2 + 2\Re\langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2,$$

puis nous concluons par l'inégalité de Cauchy-Schwarz. \square

Corollaire et définition 9.2.3 Soit E un \mathbb{K} -espace préhilbertien.

L'application $\|\cdot\| : E \rightarrow \mathbb{R}_+$ définie par

$$\forall x \in E, \quad \|x\| := \sqrt{\langle x, x \rangle},$$

est une norme sur le \mathbb{K} -espace vectoriel E .

Lorsqu'une norme est ainsi associée à un produit scalaire, on l'appelle norme euclidienne si $\mathbb{K} = \mathbb{R}$, et norme hermitienne si $\mathbb{K} = \mathbb{C}$.

L'inégalité de Minkowski étant établie, le lecteur vérifiera facilement que $\|\cdot\|$ est bien une norme. Un \mathbb{K} -espace préhilbertien est donc un espace vectoriel normé particulier.

Remarque. Un espace préhilbertien complet est appelé *espace de Hilbert*.

Grâce à un simple calcul, le lecteur vérifiera la proposition suivante.

Proposition 9.2.4 (Identité du parallélogramme) Dans le \mathbb{K} -espace préhilbertien E , nous avons l'égalité :

$$\forall (x, y) \in E^2, \quad \|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Remarques.

1. Cette identité nous permet de montrer que, sur $E = \mathbb{R}^n$, les normes $\|\cdot\|_1$ et $\|\cdot\|_\infty$ ne sont pas euclidiennes en prenant $x = (1, 0, \dots, 0)$ et $y = (0, 1, 0, \dots, 0)$.
2. Le *théorème de Fréchet-Von Neumann-Jordan* affirme que l'identité du parallélogramme caractérise les normes euclidiennes pour $\mathbb{K} = \mathbb{R}$ et les normes hermitiennes pour $\mathbb{K} = \mathbb{C}$. Le lecteur intéressé trouvera un plan détaillé de démonstration pour $\mathbb{K} = \mathbb{R}$ dans « Topologie et Analyse » par Georges Skandalis chez Dunod, pages 272 et 318, ou encore dans l'exercice corrigé [SKA 39].

9.3 Orthogonalité et procédé de Gram-Schmidt

Nous allons maintenant généraliser la notion d'orthogonalité bien connue en géométrie classique au cas des espaces préhilbertiens.

Définition 9.3.1 Soit E un \mathbb{K} -espace préhilbertien et $(x, y) \in E^2$. On dit que x et y sont orthogonaux et l'on note $x \perp y$ si $\langle x, y \rangle = 0$.

Théorème 9.3.2 (Pythagore) Dans un \mathbb{K} -espace préhilbertien E , pour tout $(x, y) \in E^2$, nous avons l'implication :

$$x \perp y \Rightarrow \|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Démonstration: Pour $\mathbb{K} = \mathbb{R}$, l'égalité $\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2$ nous donne même l'équivalence.

Pour $\mathbb{K} = \mathbb{C}$, l'égalité $\|x + y\|^2 = \|x\|^2 + 2\Re\langle x, y \rangle + \|y\|^2$ nous donne en fait l'équivalence : $\langle x, y \rangle$ imaginaire pur $\Leftrightarrow \|x + y\|^2 = \|x\|^2 + \|y\|^2$. □

Définition 9.3.3 Une famille $(e_i)_{i \in I} \in E^I$, où I est un ensemble quelconque, est dite orthogonale si, pour tout $(i, j) \in I^2$, $i \neq j \Rightarrow e_i \perp e_j$. Cette famille est dite orthonormée² si, en outre, $\|e_i\| = 1$ pour tout $i \in I$.

Proposition 9.3.4 Une famille orthogonale $(e_i)_{i \in I} \in E^I$ est libre dès que $e_i \neq 0$, pour tout $i \in I$.

Démonstration: Soit J sous-ensemble fini de I et $(\lambda_j)_{j \in J} \in \mathbb{K}^J$ une famille de scalaires telle que $\sum_{j \in J} \lambda_j e_j = 0$. Alors, pour tout $i \in J$, nous avons :

$$\left\langle e_i, \sum_{j \in J} \lambda_j e_j \right\rangle = 0 = \lambda_i \|e_i\|^2,$$

d'où $\lambda_i = 0$, ce qui nous permet de conclure. □

L'exemple suivant nous sera très utile pour étudier les séries de Fourier.

Proposition 9.3.5 Pour tout $n \in \mathbb{Z}$, nous définissons $e_n \in C^0(\mathbb{R}, \mathbb{C})$ par :

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

Alors, dans l'espace préhilbertien de Dirichlet \mathcal{D} , la famille $(e_n)_{n \in \mathbb{Z}}$ est orthonormale.

2. ou orthonormale

Démonstration: Les applications $e_n, n \in \mathbb{Z}$ sont toutes 2π -périodiques et continues donc appartiennent à \mathcal{D} . D'après la définition du produit scalaire dans \mathcal{D} (cf. Proposition 9.1.4), nous avons :

$$\forall (m, n) \in \mathbb{Z}^2, \quad \langle e_m, e_n \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{-imx} e^{inx} dx.$$

Si $m \neq n$, nous obtenons donc :

$$\langle e_m, e_n \rangle = \frac{1}{2\pi} \left[\frac{e^{i(n-m)x}}{i(n-m)} \right]_0^{2\pi} = 0.$$

Si $m = n$, nous calculons immédiatement $\|e_n\|^2 = 1$. □

Théorème 9.3.6 (Procédé d'orthonormalisation de Gram-Schmidt)
Pour toute famille libre (x_1, \dots, x_n) du \mathbb{K} -espace préhilbertien E , il existe une famille orthonormale (e_1, \dots, e_n) telle que :

$$\forall k \in \llbracket 1; n \rrbracket, \quad \text{Vect}(e_1, \dots, e_k) = \text{Vect}(x_1, \dots, x_k).$$

Démonstration: Il suffit de poser $y_1 = x_1$ et $e_1 = y_1/\|y_1\|$, puis de poursuivre la construction par récurrence sur $k \leq n$, en posant :

$$y_k := x_k - \sum_{i=1}^{k-1} \langle x_k, e_i \rangle e_i \quad \text{et} \quad e_k := \frac{y_k}{\|y_k\|}.$$

□

Remarques.

1. Dans le théorème précédent, on a existence et unicité de la famille orthonormale (e_1, \dots, e_k) si l'on ajoute la condition supplémentaire :

$$\forall k \in \llbracket 1; n \rrbracket, \quad \langle e_k, x_k \rangle \geq 0.$$

2. Cette construction par récurrence se généralise au cas d'une famille libre dénombrable $(x_n)_{n \in \mathbb{N}^*}$ de E . On obtient alors une famille orthonormale $(e_n)_{n \in \mathbb{N}^*}$ telle que : $\forall k \in \mathbb{N}^*, \text{Vect}(e_1, \dots, e_k) = \text{Vect}(x_1, \dots, x_k)$.
3. La même idée permet de construire des suites de *polynômes orthogonaux* permettant le calcul de valeurs approchées de certaines intégrales : voir [SKA 33-38].

Corollaire 9.3.7 *Tout sous-espace-vectoriel F de dimension finie du \mathbb{K} -espace préhilbertien E admet une base orthonormée.*

Remarques.

1. D'après la remarque précédente, si F est un sous-espace vectoriel de E admettant une base dénombrable $(f_n)_{n \in \mathbb{N}^*}$, alors F admet une base orthonormée $(e_n)_{n \in \mathbb{N}^*}$.
2. Soit F un sous-espace de dimension finie n de E et (e_1, \dots, e_n) une base orthonormée de F . Nous calculons alors facilement :

$$\forall x \in F, \quad x = \sum_{k=1}^n \langle e_k, x \rangle e_k \quad \text{et} \quad \|x\|^2 = \sum_{k=1}^n |\langle e_k, x \rangle|^2.$$

9.4 Projection orthogonale, meilleure approximation

Définition et proposition 9.4.1 *Pour toute partie A non vide du \mathbb{K} -espace préhilbertien E , on appelle orthogonal de A la partie de E définie par :*

$$A^\perp := \{x \in E, \forall a \in A, \langle x, a \rangle = 0\}.$$

C'est un sous-espace vectoriel de E .

Remarque. Pour toute partie A non vide de E , on vérifie facilement :

$$(\text{Vect} A)^\perp = A^\perp.$$

Proposition 9.4.2 *Si F est un sous-espace vectoriel de E , alors $F \subset F^{\perp\perp}$.*

Démonstration: Soit $x \in F$. On a $\langle x, a \rangle = 0$ pour tout $a \in F^\perp$ par définition de F^\perp , d'où $x \in F^{\perp\perp}$. □

Proposition 9.4.3 *Si F est un sous-espace vectoriel de dimension finie du \mathbb{K} -espace préhilbertien E , alors $E = F \oplus F^\perp$. En outre, si (e_0, e_1, \dots, e_n) est une base orthonormée de F , la projection orthogonale sur F , c'est-à-dire la projection sur F , parallèlement à F^\perp , est donnée par :*

$$\forall x \in E, \quad p(x) = \sum_{k=0}^n \langle e_k, x \rangle e_k. \quad (9.2)$$

Enfin, on a $F = F^{\perp\perp}$.

Démonstration: On sait que F admet une base orthonormale (e_0, \dots, e_n) . Soit $x \in E$ arbitraire fixé. Nous cherchons $(y, z) \in F \times F^\perp$ tel que $x = y + z$; autrement dit, nous cherchons $y \in F$ tel que $x - y \in F^\perp$. Écrivons $y = \sum_{i=0}^n y_i e_i$. Nous avons alors les équivalences :

$$x - y \in F^\perp \Leftrightarrow \forall k \in \llbracket 0, n \rrbracket, \langle e_k, x - y \rangle = 0 \Leftrightarrow \forall k \in \llbracket 0, n \rrbracket, \langle e_k, x \rangle = y_k.$$

Il y a donc une unique solution, ce qui nous donne $E = F \oplus F^\perp$ et nous connaissons les coordonnées de y dans la base (e_0, \dots, e_n) , ce qui nous donne la deuxième phrase de l'énoncé.

Pour terminer, prenons $x \in F^{\perp\perp}$ et écrivons-le sous la forme précédente $x = y + z$, avec $(y, z) \in F \times F^\perp$. Puisque $x \in F^{\perp\perp}$, nous avons $x \perp z$, d'où :

$$0 = \langle x, z \rangle = \langle y + z, z \rangle = \|z\|^2.$$

Nous en déduisons que $z = 0$, et donc $x = y \in F$. Ainsi, $F^{\perp\perp} \subset F$ et l'inclusion inverse est toujours vraie. \square

Remarque. Si F est de dimension infinie, on a toujours $F \cap F^\perp = \{0\}$ mais on peut avoir $F \oplus F^\perp \subsetneq E$, ainsi que $F^{\perp\perp} \subsetneq F$. [DAN 358] construit un contre-exemple en considérant $F = \mathbb{R}[X]$, sous-espace de $E = C^0([0, 1], \mathbb{R})$ muni du produit scalaire défini par :

$$\forall (f, g) \in E^2, \quad \langle f, g \rangle = \int_0^1 f(t)g(t) dt.$$

En notant $\|\cdot\|_2$ la norme associée à ce produit scalaire, nous constatons que $\|f\|_2 \leq \|f\|_\infty$ pour tout $f \in E$ et nous en déduisons, en utilisant le théorème de Weierstrass, que $F = \mathbb{R}[X]$ est dense dans $(E, \|\cdot\|_2)$. Il en résulte que $\mathbb{R}[X]^\perp = \{0\}$, d'où $\mathbb{R}[X] \oplus \mathbb{R}[X]^\perp = \mathbb{R}[X] \subsetneq E$ et $\mathbb{R}[X]^{\perp\perp} = E \subsetneq \mathbb{R}[X]$.

Définition 9.4.4 Soit F un sous-espace d'un \mathbb{K} -e.v.n. E et $x \in E$. On dit que $x_0 \in F$ est une meilleure approximation de x dans F si :

$$\|x - x_0\| = d(x, F) := \inf_{y \in F} \|x - y\|.$$

Dans le cas général, l'existence d'une telle meilleure approximation n'est pas assurée, comme on le constate en reprenant le contre-exemple de la remarque précédente. En effet, par densité de $F = \mathbb{R}[X]$ dans $(E, \|\cdot\|_2)$, nous avons $d(f, F) = 0$ pour toute application $f \in E$. Ainsi, si f n'est pas polynomiale, il n'existe pas de meilleure approximation de f sur F .

L'unicité d'une meilleure approximation n'est pas assurée non plus, comme le lecteur pourra le constater sur le contre-exemple suivant [DAN 359].

Soit \mathcal{A} le \mathbb{R} -espace vectoriel (de dimension 2) des fonctions affines de $[0, 1]$ dans \mathbb{R} , muni de la norme $\|\cdot\|_\infty$. Soit D la droite engendrée par la fonction identité i et $\mathbf{1}$ la fonction constante et égale à 1. Nous avons donc l'égalité :

$$d(\mathbf{1}, D) = \inf_{a \in \mathbb{R}} \|\mathbf{1} - ai\|_\infty.$$

Nous remarquons d'abord que, pour tout $a \in \mathbb{R}$, $\|\mathbf{1} - ai\|_\infty \geq |1 - ai(0)| = 1$, si bien que $d(\mathbf{1}, D) \geq 1$. En outre, pour $a = 0$, $\|\mathbf{1}\|_\infty = 1$ donc $d(\mathbf{1}, D) = 1$ et la fonction nulle est une meilleure approximation de $\mathbf{1}$ sur D . Mais ce n'est pas la seule car, pour tout $a \in [0, 1]$, en traçant le tableau de variation de $x \mapsto 1 - ax$ sur $[0, 1]$, nous constatons que $\|\mathbf{1} - ai\|_\infty = 1$, si bien que ai est une meilleure approximation de $\mathbf{1}$ sur D .

Proposition 9.4.5 *Soit F un sous-espace de dimension finie d'un \mathbb{K} -espace préhilbertien E , et p la projection orthogonale sur F . Alors, pour tout $x \in E$, $p(x)$ est l'unique meilleure approximation de x dans F et l'on a :*

$$d(x, F)^2 = \|x - p(x)\|^2 = \|x\|^2 - \|p(x)\|^2.$$

Démonstration: Soit $x \in E$ arbitraire fixé.

Nous avons $p(x) \in F$ et $x - p(x) \in F^\perp$ donc, pour tout $y \in F$,

$$x - p(x) \perp p(x) - y.$$

Le théorème de Pythagore nous donne alors :

$$\|x - y\|^2 = \|x - p(x)\|^2 + \|p(x) - y\|^2, \quad (9.3)$$

d'où $\|x - y\| \geq \|x - p(x)\|$ avec égalité si, et seulement si, $y = p(x)$.

En prenant $y = 0$ dans (9.3), nous obtenons la dernière égalité de l'énoncé. \square

Remarque. Si (e_0, \dots, e_n) est une base orthonormée de F , la Proposition 9.4.3 nous donne $p(x) = \sum_{k=0}^n \langle e_k, x \rangle e_k$, d'où

$$\|p(x)\|^2 = \sum_{k=0}^n |\langle e_k, x \rangle|^2 \quad (9.4)$$

et

$$d(x, F)^2 = \|x\|^2 - \sum_{k=0}^n |\langle e_k, x \rangle|^2. \quad (9.5)$$

9.5 Inégalité de Bessel et égalité de Parseval

Dans tout ce paragraphe, sauf mention explicite, nous supposons que E est un \mathbb{K} -espace préhilbertien de dimension infinie. En procédant par récurrence, nous pouvons construire une famille libre dénombrable $(x_n)_{n \in \mathbb{N}}$ de E ; nous en déduisons l'existence d'une famille orthonormée $(e_n)_{n \in \mathbb{N}}$ de E grâce au procédé d'orthonormalisation de Gram-Schmidt.

Théorème 9.5.1 (Inégalité de Bessel) *Pour tout $x \in E$, $\sum_{n \in \mathbb{N}} |\langle e_n, x \rangle|^2$ est une série convergente et sa somme est majorée par $\|x\|^2$:*

$$\sum_{n=0}^{+\infty} |\langle e_n, x \rangle|^2 \leq \|x\|^2.$$

Démonstration: Pour tout $n \in \mathbb{N}$, posons $F_n := \text{Vect}(e_0, e_1, \dots, e_n)$. D'après l'égalité (9.5), nous avons :

$$0 \leq d(x, F_n)^2 = \|x\|^2 - \sum_{k=0}^n |\langle e_k, x \rangle|^2,$$

d'où $\sum_{k=0}^n |\langle e_k, x \rangle|^2 \leq \|x\|^2$. Comme l'entier $n \in \mathbb{N}$ est arbitraire, nous obtenons la convergence de la série de l'énoncé et la majoration de sa somme. \square

Corollaire 9.5.2 (Lemme de Riemann-Lebesgue) *Pour tout $x \in E$,*

$$\lim_{n \rightarrow +\infty} \langle e_n, x \rangle = 0.$$

Définition 9.5.3 *On dit qu'une famille $(x_i)_{i \in I}$ est totale dans un espace vectoriel normé E si $\text{Vect}(x_i)_{i \in I}$ est dense dans E .*

Exemple. D'après le théorème de Weierstrass, la suite des fonctions monômes $(x \mapsto x^n)_{n \in \mathbb{N}}$ est totale dans $(C^0([0, 1], \mathbb{K}), \|\cdot\|_\infty)$.

Proposition 9.5.4 *Supposons qu'il existe une suite $(e_n)_{n \in \mathbb{N}}$ orthonormée et totale dans le \mathbb{K} -espace préhilbertien E . Pour tout $n \in \mathbb{N}$, notons p_n la projection orthogonale sur le sous-espace $F_n := \text{Vect}(e_0, e_1, \dots, e_n)$. Alors,*

$$\forall x \in E, \quad p_n(x) \xrightarrow[n \rightarrow +\infty]{} x.$$

Démonstration: Soit $x \in E$ et $\epsilon > 0$ arbitraires fixés. Puisque la suite $(e_n)_{n \in \mathbb{N}}$ est totale,

$$\exists n_0 \in \mathbb{N}, \quad \exists (\lambda_0, \dots, \lambda_{n_0}) \in \mathbb{K}^{n_0+1}, \quad \left\| x - \sum_{k=0}^{n_0} \lambda_k e_k \right\| < \epsilon.$$

Comme $p_n(x)$ est la meilleure approximation de x sur F_n , nous avons pour tout $n \geq n_0$:

$$\|x - p_n(x)\| \leq \underbrace{\left\| x - \sum_{k=0}^{n_0} \lambda_k e_k \right\|}_{\in F_n} < \epsilon,$$

d'où la conclusion. □

Remarque. Puisque (e_0, e_1, \dots, e_n) est une base orthonormale de F_n , nous avons d'après (9.2) : $p_n(x) = \sum_{k=0}^n \langle e_k, x \rangle e_k \rightarrow x$ lorsque $n \rightarrow +\infty$. Ainsi, la série $\sum_{k \in \mathbb{N}} \langle e_k, x \rangle e_k$ est convergente dans E et nous pouvons écrire :

$$\sum_{k=0}^{+\infty} \langle e_k, x \rangle e_k = x.$$

Théorème 9.5.5 (Égalité de Parseval) *Supposons qu'il existe une suite $(e_n)_{n \in \mathbb{N}}$ orthonormée et totale dans le \mathbb{K} -espace préhilbertien E . Alors, pour tout $x \in E$, la série positive $\sum_{n \in \mathbb{N}} |\langle e_n, x \rangle|^2$ converge et sa somme vaut :*

$$\sum_{n=0}^{+\infty} |\langle e_n, x \rangle|^2 = \|x\|^2.$$

Démonstration: L'application $\|\cdot\|$ étant continue (car 1-lipschitzienne) de E dans \mathbb{R} , nous déduisons de la proposition précédente :

$$\forall x \in E, \quad \|p_n(x)\|^2 \xrightarrow{n \rightarrow +\infty} \|x\|^2.$$

L'égalité (9.4) nous permet alors de conclure. □

Chapitre 10

Séries de Fourier

Historiquement, c'est pour résoudre *l'équation de la chaleur* que Fourier¹ a introduit en 1822 les séries qui portent son nom. Cette équation régit l'évolution de la température dans une barre de longueur finie, maintenue à une température constante à ses extrémités.

Pour fixer les idées, nous supposons la barre de longueur π et la température aux extrémités égale à 0. L'équation de la chaleur s'écrit alors, pour une application $u \in C^0([0, \pi] \times \mathbb{R}_+, \mathbb{R})$ telle que les dérivées partielles $\partial u / \partial t$, $\partial u / \partial x$ et $\partial^2 u / \partial x^2$ existent et sont continues sur $[0, \pi] \times \mathbb{R}_+$,

$$\left\{ \begin{array}{l} \forall (x, t) \in [0, \pi] \times \mathbb{R}_+, \quad \frac{\partial u}{\partial t}(x, t) = \frac{\partial^2 u}{\partial x^2}(x, t); \\ \forall t \in \mathbb{R}_+, \quad u(0, t) = u(\pi, t) = 0; \\ \forall x \in [0, \pi], \quad u(x, 0) = f(x), \end{array} \right.$$

où $f \in C^1([0, \pi], \mathbb{R})$ est la condition initiale (température le long de la barre à l'instant 0). Après l'étude de ce chapitre, le lecteur pourra se reporter au Problème 7.3 de [CMP 429] pour résoudre cette équation.

10.1 Polynôme et série trigonométriques

Pour tout $n \in \mathbb{Z}$, nous définissons comme suit l'application $e_n : \mathbb{R} \rightarrow \mathbb{C}$.

$$\forall x \in \mathbb{R}, \quad e_n(x) = e^{inx}. \quad (10.1)$$

1. Joseph Fourier (1768-1830), mathématicien et physicien français.

Définition 10.1.1 On appelle polynôme trigonométrique toute application $P : \mathbb{R} \rightarrow \mathbb{C}$ de la forme :

$$P = \sum_{k=-n}^n c_k e_k, \text{ avec } n \in \mathbb{N} \text{ et } (c_k)_{-n \leq k \leq n} \in \mathbb{C}^{2n+1}.$$

Un polynôme trigonométrique est donc une application continue et 2π -périodique.

Définition 10.1.2 On appelle série trigonométrique toute série d'applications de \mathbb{R} dans \mathbb{C} de la forme :

$$c_0 + \sum_{n \in \mathbb{N}^*} (c_n e_n + c_{-n} e_{-n}), \text{ avec } (c_n)_{n \in \mathbb{Z}} \in \mathbb{C}^{\mathbb{Z}}. \quad (10.2)$$

Nous la noterons² $\sum_{n \in \mathbb{Z}} c_n e_n$.

Attention, la convergence de la série $\sum_{n \in \mathbb{Z}} c_n e_n$ au point $x \in \mathbb{R}$ peut donc avoir lieu sans que la limite suivante existe :

$$\lim_{\substack{p \rightarrow -\infty \\ q \rightarrow +\infty}} \sum_{n=p}^q c_n e^{inx}.$$

Néanmoins, si la série $\sum_{n \in \mathbb{Z}} c_n e_n$ converge au point $x \in \mathbb{R}$, la « convention de Cauchy » veut que l'on note sa somme comme suit :

$$\sum_{n=-\infty}^{+\infty} c_n e^{inx}.$$

Remarque. En utilisant la formule de Moivre $e^{inx} = \cos nx + i \sin nx$, nous constatons que la série trigonométrique (10.2) s'écrit encore

$$\frac{a_0}{2} + \sum_{n \in \mathbb{N}^*} (a_n \cos nx + b_n \sin nx), \quad (10.3)$$

où l'on a posé :

$$\forall n \in \mathbb{N}, \quad a_n = c_n + c_{-n} \quad ; \quad \forall n \in \mathbb{N}^*, \quad b_n = i(c_n - c_{-n}). \quad (10.4)$$

Proposition 10.1.3 Si les séries positives $\sum_{n \in \mathbb{N}} |c_n|$ et $\sum_{n \in \mathbb{N}} |c_{-n}|$ convergent, alors la série trigonométrique $\sum_{n \in \mathbb{Z}} c_n e_n$ converge normalement sur \mathbb{R} .

2. On utilise aussi, par abus d'écriture, la notation $\sum_{n \in \mathbb{Z}} c_n e^{inx}$.

Démonstration: L'inégalité triangulaire nous permet d'écrire :

$$\forall n \in \mathbb{N}^*, \quad \forall x \in \mathbb{R} \quad |c_n e^{inx} + c_{-n} e^{-inx}| \leq |c_n| + |c_{-n}|,$$

et ce majorant est le terme général d'une série convergente par somme de séries convergentes. \square

Proposition 10.1.4 *Si les deux suites $(c_n)_{n \in \mathbb{N}}$ et $(c_{-n})_{n \in \mathbb{N}}$ sont réelles et tendent vers 0 en décroissant, alors la série trigonométrique $\sum_{n \in \mathbb{Z}} c_n e_n$ converge simplement sur $\mathbb{R} - (2\pi\mathbb{Z})$ et uniformément sur toute intervalle de la forme $2k\pi + [\alpha, 2\pi - \alpha]$, $k \in \mathbb{Z}$, $0 < \alpha < \pi$.*

Démonstration: Nous allons établir dans un premier temps la convergence uniforme de la série d'applications $\sum_{n \in \mathbb{N}} c_n e_n$ sur $2k\pi + [\alpha, 2\pi - \alpha]$. Il suffit d'appliquer le *critère d'Abel uniforme* (cf. Théorème 7.6.5, page 153) avec $f_n = e_n$ et $g_n \equiv c_n$ pour tout $n \in \mathbb{N}$. En effet, l'égalité

$$\sum_{k=0}^n f_k(x) = \frac{1 - e^{i(n+1)x}}{1 - e^{ix}} = \frac{1 - e^{i(n+1)x}}{e^{-ix/2} - e^{ix/2}} e^{-ix/2}$$

entraîne la majoration suivante, pour tout $x \in 2k\pi + [\alpha, 2\pi - \alpha]$:

$$\left| \sum_{k=0}^n f_k(x) \right| \leq \frac{2}{|e^{-ix/2} - e^{ix/2}|} = \frac{1}{|\sin \frac{x}{2}|} \leq \frac{1}{|\sin \frac{\alpha}{2}|},$$

ce qui nous permet de conclure.

Nous reprenons alors la même démonstration avec $f_n = e_{-n}$ et $g_n \equiv c_{-n}$ pour obtenir un résultat similaire avec la série $\sum_{n \in \mathbb{N}} c_{-n} e_{-n}$, ce qui permet de conclure à la convergence uniforme annoncée.

Enfin, la convergence simple de la série trigonométrique $\sum_{n \in \mathbb{Z}} c_n e_n$ sur l'intervalle $[\alpha, 2\pi - \alpha]$, pour tout $0 < \alpha < \pi$, nous donne sa convergence simple sur l'intervalle $]0, 2\pi[$ puis, par 2π -périodicité, sur $\mathbb{R} \setminus (2\pi\mathbb{Z})$. \square

Nous terminons cette section en citant sans démonstration un théorème (hors programme) qui établit une condition nécessaire à la convergence d'une série trigonométrique.

Théorème 10.1.5 (Cantor-Lebesgue) *Si la série trigonométrique $\sum_{n \in \mathbb{Z}} c_n e^{inx}$ converge simplement sur \mathbb{R} , alors :*

$$c_n \xrightarrow{n \rightarrow \pm\infty} 0.$$

10.2 Coefficients et série de Fourier

Définition 10.2.1 Soit $f \in C_M(\mathbb{R}, \mathbb{C})$ une application 2π -périodique. On appelle coefficients de Fourier de f les nombres complexes

$$c_n(f) := \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt, \quad n \in \mathbb{Z}.$$

On appelle série de Fourier associée à f la série trigonométrique $\sum_{n \in \mathbb{Z}} c_n(f) e_n$, où les applications $e_n : \mathbb{R} \rightarrow \mathbb{C}$, $n \in \mathbb{Z}$, sont définies par (10.1).

Cette définition des coefficients et de la série de Fourier est naturelle à cause du résultat suivant.

Lemme 10.2.2 Une série trigonométrique qui converge uniformément sur \mathbb{R} est égale à sa série de Fourier.

Démonstration: Notons $S(x) := \sum_{n=-\infty}^{+\infty} c_n e^{inx}$. Grâce à la Proposition 1.5.1, page 40, nous pouvons écrire, pour tout $p \in \mathbb{Z}$:

$$\begin{aligned} c_p(S) &= \frac{1}{2\pi} \int_0^{2\pi} S(t) e^{-ipt} dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} c_0 e^{-ipt} dt + \sum_{n=1}^{+\infty} \frac{1}{2\pi} \int_0^{2\pi} (c_n e^{int} + c_{-n} e^{-int}) e^{-ipt} dt \\ &= c_p, \end{aligned}$$

puisque, selon la Proposition 9.3.5, page 171, la famille $(e_n)_{n \in \mathbb{Z}}$ est orthonormale dans l'espace de Dirichlet \mathcal{D} . \square

Remarque. La série de Fourier associée à f s'écrit encore :

$$\frac{a_0(f)}{2} + \sum_{n \in \mathbb{N}^*} (a_n(f) \cos nx + b_n(f) \sin nx),$$

avec, selon (10.4),

$$\forall n \in \mathbb{N}, \quad a_n(f) := \frac{1}{\pi} \int_0^{2\pi} f(t) \cos nt dt; \quad \forall n \in \mathbb{N}^*, \quad b_n(f) := \frac{1}{\pi} \int_0^{2\pi} f(t) \sin nt dt.$$

Tous les résultats énoncés dans la suite de ce chapitre avec les coefficients de Fourier $(c_n(f))_{n \in \mathbb{Z}}$ ont donc des analogues avec les coefficients $(a_n(f))_{n \in \mathbb{N}}$ et $(b_n(f))_{n \in \mathbb{N}^*}$, que nous donnerons parfois brièvement. Nous renvoyons le lecteur à [GOU 256 sqq.] pour les détails.

Remarques.

1. Si la série trigonométrique converge simplement mais pas uniformément sur \mathbb{R} , il n'est pas garanti que sa somme S soit continue par morceaux donc que l'on puisse lui associer une série de Fourier.
2. Le critère d'Abel nous permet de prouver la convergence simple sur \mathbb{R} de la série trigonométrique $\sum_{n \in \mathbb{N}^*} \sin(nx)/\sqrt{n}$ vers une somme S impaire. Si cette convergence était uniforme, nous aurions $S \in C_M(\mathbb{R}, \mathbb{C})$ et $b_n(S) = 1/\sqrt{n}$ pour tout $n \in \mathbb{N}^*$, ce qui impliquerait, d'après l'inégalité de Bessel (que nous prouverons ci-dessous) :

$$\sum_{n=1}^{+\infty} \frac{1}{n} = \sum_{n=1}^{+\infty} b_n^2(S) \leq \frac{1}{\pi} \int_0^{2\pi} S^2 < +\infty.$$

Ceci est absurde puisque la série harmonique diverge. Nous avons donc démontré par l'absurde que $\sum_{n \in \mathbb{N}^*} \sin(nx)/\sqrt{n}$ ne converge pas uniformément sur \mathbb{R} .

10.3 Le théorème de convergence de Dirichlet

Définition et proposition 10.3.1 On appelle noyau de Dirichlet la suite de polynômes trigonométriques $(D_n)_{n \in \mathbb{N}}$ définie par :

$$\forall n \in \mathbb{N}, \quad \forall x \in \mathbb{R}, \quad D_n(x) := \sum_{k=-n}^n e^{ikx}.$$

Pour tout $n \in \mathbb{N}$, le polynôme trigonométrique D_n est une fonction paire, à valeurs réelles, et telle que $\frac{1}{2\pi} \int_0^{2\pi} D_n = 1$. En outre, nous avons l'égalité :

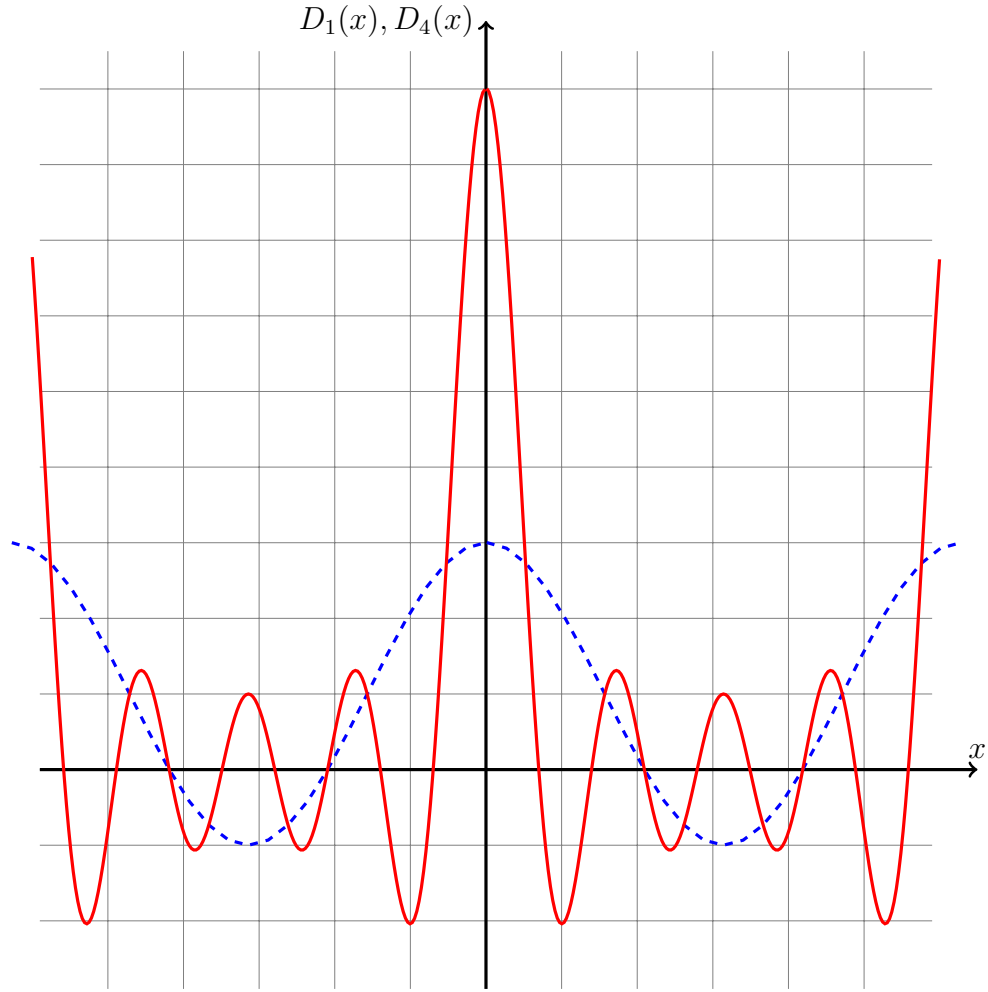
$$\forall n \in \mathbb{N}, \quad \forall x \in \mathbb{R} \setminus (2\pi\mathbb{Z}), \quad D_n(x) = \frac{\sin(2n+1)\frac{x}{2}}{\sin \frac{x}{2}}.$$

Démonstration: La première phrase est immédiate.

Pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R} - (2\pi\mathbb{Z})$, nous calculons, en utilisant la somme d'une série géométrique :

$$D_n(x) = e^{-inx} \frac{e^{i(2n+1)x} - 1}{e^{ix} - 1} = e^{-inx} \frac{e^{i(2n+1)x/2}}{e^{ix/2}} \frac{e^{i(2n+1)x/2} - e^{-i(2n+1)x/2}}{e^{ix/2} - e^{-ix/2}}$$

d'où l'égalité annoncée. □



Théorème 10.3.2 (Jordan-Dirichlet) Soit $f \in C_M^1(\mathbb{R}, \mathbb{C})$ une application 2π -périodique. Alors, la série de Fourier associée à f converge simplement sur \mathbb{R} vers la régularisée de f :

$$\forall x \in \mathbb{R}, \quad \sum_{n=-\infty}^{+\infty} c_n(f) e^{inx} = \tilde{f}(x) := \frac{f(x-) + f(x+)}{2}.$$

Démonstration: Si nous remplaçons l'application f par sa régularisée $\tilde{f} \in \mathcal{D}$, cela ne la modifie sur $[0, 2\pi]$ qu'en un nombre fini de points et donc cela laisse les coefficients de Fourier inchangés. Par conséquent, il suffit de prouver que si $f \in \mathcal{D} \cap C_M^1(\mathbb{R}, \mathbb{C})$, alors la série de Fourier associée à f converge simplement vers f sur \mathbb{R} .

En considérant la fonction $f_{x_0} = f(x_0 + \cdot)$, nous calculons, à l'aide du chan-

gement de variable $u = x_0 + t$,

$$\begin{aligned}\forall n \in \mathbb{Z}, \quad c_n(f_{x_0}) &= \frac{1}{2\pi} \int_0^{2\pi} f(x_0 + t) e^{-int} dt \\ &= \frac{1}{2\pi} \int_{x_0}^{x_0+2\pi} f(u) e^{-in(u-x_0)} du \\ &= \frac{1}{2\pi} e^{inx_0} \int_{x_0}^{x_0+2\pi} f(u) e^{-inu} du.\end{aligned}$$

Ce dernier intégrande étant 2π -périodique, nous en déduisons l'égalité :

$$\forall n \in \mathbb{Z}, \quad c_n(f_{x_0}) = c_n(f) e^{inx_0}.$$

Par conséquent, si nous établissons l'égalité annoncée au point $x = 0$, *i.e.*

$$\forall f \in \mathcal{D} \cap C_M^1(\mathbb{R}, \mathbb{C}), \quad \sum_{n=-\infty}^{+\infty} c_n(f) = f(0),$$

alors cela suffit pour l'obtenir en un point $x_0 \in \mathbb{R}$ arbitraire.

Nous posons donc, pour tout $n \in \mathbb{N}$, $s_n := \sum_{k=-n}^n c_k(f)$, $u_n := s_n - f(0)$ et nous cherchons à montrer que $u_n \rightarrow 0$.

Un calcul simple utilisant les propriétés de D_n énoncées précédemment nous donne $2\pi u_n = \int_0^\pi (f(t) + f(-t) - 2f(0)) D_n(t) dt$, d'où, avec un changement de variable élémentaire :

$$2\pi u_n = 2 \int_0^{\pi/2} \frac{f(2t) + f(-2t) - f(0+) - f(0-)}{\sin t} \sin(2n+1)t dt.$$

Comme $f \in C_M^1(\mathbb{R}, \mathbb{C})$, en utilisant l'équivalent $\sin t \sim t$ au voisinage de 0, le lecteur établira facilement la convergence suivante :

$$\frac{f(2t) + f(-2t) - f(0+) - f(0-)}{\sin t} \xrightarrow[t \rightarrow 0+]{\quad} 2(f'(0+) - f'(0-)).$$

Avec un instant de réflexion, nous en déduisons qu'il est possible de définir une application $g \in C_M(\mathbb{R}, \mathbb{C})$, 2π -périodique, vérifiant :

$$\forall t \in]0, 2\pi[, \quad g(t) = \frac{f(2t) + f(-2t) - f(0+) - f(0-)}{\sin t} \mathbf{1}_{]0, \frac{\pi}{2}[}(t).$$

Le lemme de Riemann-Lebesgue nous donne alors $\lim_{n \rightarrow \pm\infty} c_n(g) = 0$, ce qui nous permet de conclure puisque :

$$u_n = \frac{1}{\pi} \int_0^{2\pi} g(t) \sin(2n+1)t dt = i (c_{2n+1}(g) - c_{-(2n+1)}(g)).$$

□

Sous les hypothèses du théorème de Jordan-Dirichlet, si f est en outre continue sur \mathbb{R} , alors $\tilde{f} = f$ et la série de Fourier associée à f converge simplement sur \mathbb{R} vers f . En fait, nous avons dans ce cas un résultat bien meilleur.

Théorème 10.3.3 *Soit $f \in C^0(\mathbb{R}, \mathbb{C}) \cap C_M^1(\mathbb{R}, \mathbb{C})$ une application 2π -périodique. Alors la série de Fourier associée à f converge normalement vers f sur \mathbb{R} .*

Pour démontrer ce théorème, nous établissons d'abord un lemme.

Lemme 10.3.4 *Soit $f \in C_M^1(\mathbb{R}, \mathbb{C})$ une application 2π -périodique. Nous définissons l'application $\varphi : \mathbb{R} \rightarrow \mathbb{C}$ par*

$$\varphi(t) := \begin{cases} f'(t) & \text{si } f \text{ est dérivable au point } t, \\ \frac{f'(t-) + f'(t+)}{2} & \text{sinon.} \end{cases}$$

Alors $\varphi \in \mathcal{D}$ et, pour tout $n \in \mathbb{Z}$, $c_n(\varphi) = inc_n(f)$.

Démonstration: Il est facile de vérifier que $\varphi \in \mathcal{D}$. Si $f \in C^1(\mathbb{R}, \mathbb{C})$, une intégration par parties nous donne :

$$c_n(f') = \frac{1}{2\pi} \int_0^{2\pi} f'(t) e^{-int} dt = \left[\frac{f(t) e^{-int}}{2\pi} \right]_0^{2\pi} + \frac{in}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt,$$

d'où $c_n(\varphi) = inc_n(f)$ pour tout $n \in \mathbb{Z}$.

Dans le cas général, on applique la relation de Chasles le long d'une subdivision de $[0, 2\pi]$ adaptée à $f \in C_M^1(\mathbb{R}, \mathbb{C})$; le reste de la preuve est analogue. Le lecteur trouvera les détails dans [GOU 261]. □

Démonstration du théorème : En reprenant les notations du lemme, nous avons :

$$\forall n \in \mathbb{Z}^*, \quad |c_n(f)| = \frac{1}{n} |c_n(\varphi)| \leq \frac{1}{2} \left(\frac{1}{n^2} + |c_n(\varphi)|^2 \right).$$

Puisque $\varphi \in \mathcal{D}$, l'égalité de Parseval nous permet d'en déduire que $\sum_{n \in \mathbb{Z}} |c_n(f)|$ converge. Nous concluons grâce à la Proposition 10.1.3. □

Sous les hypothèses du théorème de Jordan-Dirichlet, si f n'est pas continue sur \mathbb{R} , il est impossible que la série de Fourier associée à f converge uniformément vers la régularisée \tilde{f} , puisque celle-ci n'est pas non plus continue sur \mathbb{R} . Ce défaut de convergence uniforme est flagrant quand on observe le

phénomène de Gibbs (cf. [GOU 266-267]). Considérons par exemple le *signal carré impair* f défini sur $[-\pi, \pi[$ par

$$f(t) = \begin{cases} -\frac{\pi}{4} & \text{si } -\pi < t < 0, \\ 0 & \text{si } t = -\pi \text{ ou } t = 0, \\ \frac{\pi}{4} & \text{si } 0 < t < \pi, \end{cases}$$

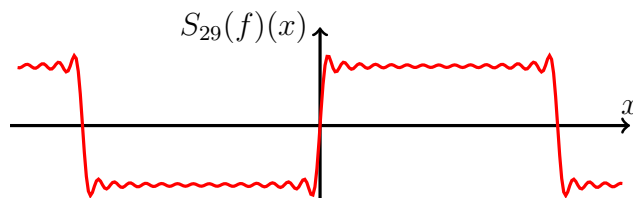
et prolongé à \mathbb{R} par 2π -périodicité. Il est facile de calculer :

$$\forall n \in \mathbb{N}, \quad b_{2n+1}(f) = \frac{1}{2n+1} \quad \text{et} \quad \forall n \in \mathbb{N}^*, \quad b_{2n}(f) = 0.$$

Ainsi, pour tout $n \in \mathbb{N}$, la somme partielle d'ordre $2n+1$ de la série de Fourier associée à f est donnée par :

$$\forall x \in \mathbb{R}, \quad S_{2n+1}(f)(x) = \sum_{k=0}^n \frac{\sin(2k+1)x}{2k+1}.$$

On peut montrer qu'il se produit au voisinage de 0 une « oscillation » de la fonction $S_{2n+1}(f)$ d'amplitude incompressible, aussi grande que soit la valeur de l'entier n .



Remarque. L'hypothèse $f \in C_M^1(\mathbb{R}, \mathbb{C})$ est importante pour obtenir la convergence simple de la série de Fourier associée à f vers la régularisée \tilde{f} . Le mathématicien allemand Paul Du Bois-Reymond a construit en 1873 une application 2π -périodique $f \in C^0(\mathbb{R}, \mathbb{R})$ dont la série de Fourier diverge en 0. Le lecteur en trouvera un autre exemple, construit ultérieurement par le mathématicien hongrois Lipot Fejér, dans [GOU 264-265].

Exercice : On considère la suite de polynômes à coefficients réels $(P_k)_{k \geq 1}$ caractérisés par les relations $P_1 = \pi - X$ et, pour tout $k \geq 1$, $P'_{k+1} = P_k$ et $\int_0^{2\pi} P_k(t) dt = 0$.

1. Montrer que :

$$\forall k \geq 1, \quad \forall t \in \mathbb{R}, \quad P_k(2\pi - t) = (-1)^k P_k(t).$$

2. Montrer que :

$$\forall k \geq 1, \quad \forall n \in \mathbb{Z}^*, \quad (2\pi)^{-1} \int_0^{2\pi} P_k(t) e^{-int} dt = (in)^{-k}.$$

3. Montrer que, pour tout $k \in \mathbb{N}$, on a $P_{2k+1}(\pi) = 0$ et, pour $k \geq 1$,

$$\frac{1}{2\pi} \int_0^{2\pi} P_k(t)^2 dt = 2 \sum_{n=1}^{+\infty} n^{-2k} = (-1)^k P_{2k}(0).$$

4. En déduire les égalités $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$ et $\sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$.

Solution (abrégée) :

1. Remarquons d'abord que la relation $P'_{k+1} = P_k$ et $\int_0^{2\pi} P_{k+1}(t) dt = 0$ déterminent entièrement P_{k+1} .

Si P_k vérifie $P_k(2\pi - t) = (-1)^k P_k(t)$, alors le polynôme Q défini par $Q(t) = (-1)^{k+1} P_{k+1}(2\pi - t)$ vérifie $Q' = P_k$ et $\int_0^{2\pi} Q(t) dt = 0$, donc $Q = P_{k+1}$.

2. Par intégration par parties, on trouve :

$$\int_0^{2\pi} P_k(t) e^{-int} dt = \left[P_k(t) \frac{e^{-int}}{-in} \right]_0^{2\pi} - \int_0^{2\pi} P'_k(t) \frac{e^{-int}}{-in} dt$$

Il vient $(2\pi)^{-1} \int_0^{2\pi} P_1(t) e^{-int} dt = (in)^{-1}$, puis, par récurrence, l'égalité demandée.

3. D'après la première question, nous avons $(-1)^{2k+1} P_{2k+1}(\pi) = P_{2k+1}(\pi)$, d'où $P_{2k+1}(\pi) = 0$.

D'après l'identité de Parseval, on a $\frac{1}{2\pi} \int_0^{2\pi} P_k(t)^2 dt = \sum_{n=-\infty}^{+\infty} |c_n(P_k)|^2$. Or $c_0(P_k) = 0$ et $|c_n(P_k)|^2 = n^{-2k}$ pour $n \neq 0$.

D'après le théorème de Dirichlet,

$$P_{2k}(0) = \sum_{n \in \mathbb{Z}} c_n(P_{2k}) = \sum_{n \neq 0} (in)^{-2k} = (-1)^k 2 \sum_{n=1}^{+\infty} n^{-2k}.$$

On a $P_2(t) = -\frac{(t-\pi)^2}{2} + c$; or $\int_0^{2\pi} \left(-\frac{(t-\pi)^2}{2} + c\right) dt = 2\pi c - \left[\frac{(t-\pi)^3}{6}\right]_0^{2\pi} = 2\pi c - \frac{\pi^3}{3}$. Il vient $P_2(t) = \frac{\pi^2}{6} - \frac{(t-\pi)^2}{2}$.

On en déduit $\sum_{n=1}^{+\infty} n^{-2} = -\frac{P_2(0)}{2} = \frac{\pi^2}{6}$ et

$$\begin{aligned} \sum_{n=1}^{+\infty} n^{-4} &= \frac{1}{4\pi} \int_0^{2\pi} P_2(t)^2 dt = \frac{1}{4\pi} \int_0^{2\pi} \left(\frac{\pi^4}{36} - \frac{\pi^2(t-\pi)^2}{6} + \frac{(t-\pi)^4}{4} \right) dt \\ &= \frac{\pi^4}{72} - \left[\frac{\pi}{4} \frac{(t-\pi)^3}{18} \right]_0^{2\pi} + \left[\frac{1}{4\pi} \frac{(t-\pi)^5}{20} \right]_0^{2\pi} \\ &= \frac{\pi^4}{72} - \frac{\pi^4}{36} + \frac{\pi^4}{40} = \frac{\pi^4}{90}. \end{aligned}$$

Le lecteur trouvera un exercice analogue dans [DAN 392-395].

Exercice : Inégalité de Wirtinger Soit $f \in C^1(\mathbb{R}, \mathbb{C})$ une application 2π -périodique telle que $\int_0^{2\pi} f = 0$. Établir l'inégalité :

$$\int_0^{2\pi} |f(t)|^2 dt \leq \int_0^{2\pi} |f'(t)|^2 dt.$$

Solution : [GOU 264].

Le même genre d'idée est utilisée dans la preuve de *l'inégalité isopérimétrique*.

Exercice : Développement du sinus en produit infini [GOU 262].

Soit $\alpha \in \mathbb{R} - \mathbb{Z}$; nous définissons l'application 2π -périodique $f_\alpha : \mathbb{R} \rightarrow \mathbb{R}$ par $f_\alpha(t) = \cos(\alpha t)$ pour tout $t \in]-\pi, \pi]$.

En développant f_α en série de Fourier, montrer l'égalité :

$$\forall t \in \mathbb{R} \setminus (\pi\mathbb{Z}), \quad \cotant = \frac{1}{t} + 2t \sum_{n=1}^{+\infty} \frac{1}{t^2 - n^2\pi^2}.$$

En intégrant l'égalité précédente entre 0 et $x \in]0, \pi[$, établir la formule :

$$\forall t \in]-\pi, \pi[, \quad \sin t = t \prod_{n=1}^{+\infty} \left(1 - \frac{t^2}{n^2\pi^2} \right).$$

Ce développement du sinus en produit infini et la formule d'Euler-Gauss (Proposition 3.3.3, page 83) nous permettent d'établir le résultat suivant.

Proposition 10.3.5 (Formule des compléments) La fonction Γ d'Euler vérifie l'égalité suivante :

$$\forall x \in]0, 1[, \quad \frac{1}{\Gamma(x)\Gamma(1-x)} = \frac{\sin(\pi x)}{\pi}.$$

10.4 Le théorème de Fejér

Nous reprenons ici la notation $(D_n)_{n \in \mathbb{N}}$ pour le noyau de Dirichlet introduit dans la section précédente.

Définition et proposition 10.4.1 *On appelle noyau de Fejér la suite de polynômes trigonométriques $(F_n)_{n \in \mathbb{N}^*}$ définie par :*

$$\forall n \in \mathbb{N}^*, \quad \forall x \in \mathbb{R}, \quad F_n(x) := \frac{1}{n} \sum_{k=0}^{n-1} D_k(x).$$

Pour tout $n \in \mathbb{N}^*$, F_n est une fonction paire telle que $\frac{1}{2\pi} \int_0^{2\pi} F_n = 1$.
En outre, nous avons :

$$\forall n \in \mathbb{N}^*, \quad \forall x \in \mathbb{R} \setminus (2\pi\mathbb{Z}), \quad F_n(x) = \frac{1}{n} \left(\frac{\sin \frac{nx}{2}}{\sin \frac{x}{2}} \right)^2 \geq 0.$$

Démonstration: La première phrase se vérifie immédiatement.

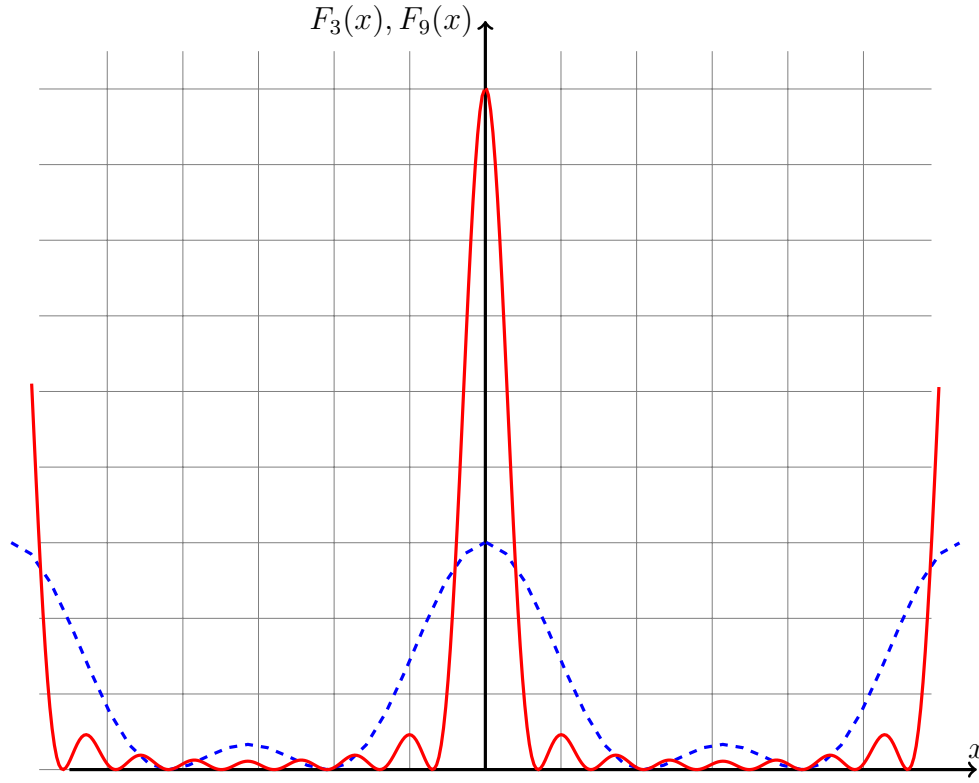
En utilisant la Proposition 10.3.1, nous calculons, pour tout $n \in \mathbb{N}^*$ et tout $x \in \mathbb{R} \setminus (2\pi\mathbb{Z})$,

$$F_n(x) = \frac{1}{n} \sum_{k=0}^{n-1} \frac{\sin(2k+1)\frac{x}{2}}{\sin \frac{x}{2}} = \frac{1}{n \sin \frac{x}{2}} \Im \left(\sum_{k=0}^{n-1} e^{i(2k+1)\frac{x}{2}} \right). \quad (10.5)$$

Or la somme entre parenthèses s'écrit encore :

$$e^{i\frac{x}{2}} \sum_{k=0}^{n-1} e^{ikx} = e^{i\frac{x}{2}} \frac{e^{inx} - 1}{e^{ix} - 1} = e^{i\frac{x}{2}} \frac{e^{i\frac{nx}{2}} e^{i\frac{nx}{2}} - e^{-i\frac{nx}{2}}}{e^{i\frac{x}{2}} e^{i\frac{x}{2}} - e^{-i\frac{x}{2}}} = e^{i\frac{nx}{2}} \frac{\sin(\frac{nx}{2})}{\sin(\frac{x}{2})}.$$

En passant à la partie imaginaire puis en revenant à (10.5), nous obtenons le résultat annoncé. □



Soit $f \in C_M(\mathbb{R}, \mathbb{C})$ une application 2π -périodique. Nous reprenons la notation $(S_n(f))_{n \in \mathbb{N}}$ pour la suite des sommes partielles associées à la série de Fourier de f , i.e.

$$\forall n \in \mathbb{N}, \quad \forall x \in \mathbb{R}, \quad S_n(f)(x) := \sum_{k=-n}^n c_k(f) e^{ikx}.$$

Définition 10.4.2 Soit $f \in C_M(\mathbb{R}, \mathbb{C})$ une application 2π -périodique. Pour tout $n \in \mathbb{N}^*$, on appelle somme de Fejér d'ordre n associée à f le polynôme trigonométrique $\sigma_n(f) : \mathbb{R} \rightarrow \mathbb{C}$ défini par :

$$\sigma_n(f) := \frac{1}{n} \sum_{k=0}^{n-1} S_k(f).$$

La somme de Fejér d'ordre n est donc la *moyenne de Cesàro* des n premières sommes partielles de Fourier.

Théorème 10.4.3 (Fejér) Soit $f \in C^0(\mathbb{R}, \mathbb{C})$ une application 2π -périodique. Alors, la suite des sommes de Fejér $(\sigma_n(f))_{n \in \mathbb{N}^*}$ converge uniformément vers f sur \mathbb{R} .

Démonstration: Pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R}$, nous avons :

$$\begin{aligned} S_n(f)(x) &= \sum_{k=-n}^n \frac{1}{2\pi} \left(\int_0^{2\pi} f(t) e^{-ikt} dt \right) e^{ikx} \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(t) \sum_{k=-n}^n e^{ik(x-t)} dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(t) D_n(x-t) dt. \end{aligned}$$

Nous en déduisons que, pour tout $n \in \mathbb{N}^*$ et tout $x \in \mathbb{R}$,

$$\sigma_n(f)(x) = \frac{1}{2\pi} \int_0^{2\pi} f(t) \frac{1}{n} \sum_{k=0}^{n-1} D_k(x-t) dt = \frac{1}{2\pi} \int_0^{2\pi} f(t) F_n(x-t) dt.$$

Grâce au changement de variable $s = x - t$ et à la 2π -périodicité de f et F_n , nous en déduisons :

$$\sigma_n(f)(x) = \frac{1}{2\pi} \int_{x-2\pi}^x f(x-t) F_n(t) dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) F_n(t) dt.$$

Comme $\frac{1}{2\pi} \int_0^{2\pi} F_n = 1$, nous avons donc, pour tout $n \in \mathbb{N}^*$ et tout $x \in \mathbb{R}$,

$$\sigma_n(f)(x) - f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(x-t) - f(x)] F_n(t) dt.$$

Comme f est continue et 2π -périodique, on montre facilement grâce au théorème de Heine que f est uniformément continue sur \mathbb{R} . Soit $\epsilon > 0$ arbitraire ; il existe donc $\delta \in]0, \pi[$ tel que $|x - y| \leq \delta \Rightarrow |f(x) - f(y)| < \frac{\epsilon}{2}$. De l'égalité précédente, nous déduisons :

$$2\pi |\sigma_n(f)(x) - f(x)| \leq \int_{-\pi}^{-\delta} |f(x-t) - f(x)| \underbrace{F_n(t)}_{\geq 0} dt + \frac{\epsilon}{2} \underbrace{\int_{-\delta}^{\delta} F_n}_{\leq \int_{-\pi}^{\pi} F_n = 2\pi} + \int_{\delta}^{\pi} \dots$$

Nous allons majorer cette dernière intégrale, la 1^{re} se traitant de même.

$$\int_{\delta}^{\pi} |f(x-t) - f(x)| F_n(t) dt \leq 2\|f\|_{\infty} \int_{\delta}^{\pi} \frac{1}{n} \left(\frac{\sin \frac{nt}{2}}{\sin \frac{t}{2}} \right)^2 dt \leq \frac{2\pi \|f\|_{\infty}}{n \sin^2 \frac{\delta}{2}}.$$

Revenant à l'inégalité précédente, nous en déduisons :

$$\forall n \in \mathbb{N}^*, \quad \forall x \in \mathbb{R}, \quad |\sigma_n(f)(x) - f(x)| \leq \frac{\epsilon}{2} + \frac{2\|f\|_{\infty}}{n \sin^2 \frac{\delta}{2}}.$$

Finalement, $\exists N \in \mathbb{N}$, $\forall n \geq N$, $\forall x \in \mathbb{R}$, $|\sigma_n(f)(x) - f(x)| \leq \epsilon$, ce qui nous permet de conclure puisque $\epsilon > 0$ était arbitraire. \square

Remarque. On peut montrer plus généralement que si $f \in \mathcal{D}$, alors la suite d'applications $(\sigma_n(f))_{n \in \mathbb{N}^*}$ converge simplement sur \mathbb{R} vers f , et uniformément sur tout intervalle sur lequel f est continue [DAN 386].

Nous déduisons du théorème de Fejér le résultat suivant, qui est un analogue du théorème de Weierstrass dans le cas des fonctions continues périodiques.

Corollaire 10.4.4 *Toute application 2π -périodique $f \in C^0(\mathbb{R}, \mathbb{C})$ est limite uniforme d'une suite de polynômes trigonométriques.*

Application : Valeur moyenne d'une fonction périodique continue

Soit $\alpha \in \mathbb{R}$ tel que $\frac{\alpha}{\pi} \notin \mathbb{Q}$ et $f \in C^0(\mathbb{R}, \mathbb{C})$ une application 2π -périodique. Alors, on a la convergence suivante :

$$\frac{1}{n} \sum_{k=0}^{n-1} f(k\alpha) \xrightarrow{n \rightarrow +\infty} \frac{1}{2\pi} \int_0^{2\pi} f.$$

L'idée est de montrer cette propriété quand f est un monôme trigonométrique, puis un polynôme trigonométrique. On passe ensuite au cas général en utilisant le corollaire précédent [DAN 395-396].

Une conséquence importante du théorème de Fejér est la suivante.

Corollaire 10.4.5 *Dans le \mathbb{C} -espace préhilbertien de Dirichlet \mathcal{D} , la suite orthonormée $(e_n)_{n \in \mathbb{Z}}$ définie par $e_n : \mathbb{R} \rightarrow \mathbb{C}, x \mapsto e^{inx}$, est totale.*

Démonstration: Soit $f \in \mathcal{D}$. Si f est continue, le théorème de Fejér nous dit qu'il existe une suite $(g_n)_{n \in \mathbb{N}}$ de polynômes trigonométriques telle que $\|g_n - f\|_\infty \rightarrow 0$. Comme nous avons l'inégalité $\|\cdot\|_2 \leq \|\cdot\|_\infty$, nous en déduisons que $\|g_n - f\|_2 \rightarrow 0$.

Si f n'est pas continue, pour tout $\epsilon > 0$, on peut construire (cf. [DAN 380]) une application \tilde{f} continue telle que $\|f - \tilde{f}\|_2 < \epsilon/2$. Or, nous venons de voir qu'il existe un polynôme trigonométrique g tel que $\|g - \tilde{f}\|_2 < \epsilon/2$. Grâce à l'inégalité triangulaire, nous en déduisons $\|g - f\|_2 < \epsilon$.

Nous avons ainsi prouvé la densité du sous-espace $\text{Vect}(e_n)_{n \in \mathbb{Z}}$ des polynômes trigonométriques dans \mathcal{D} . \square

10.5 Approximation en moyenne quadratique

Nous allons maintenant appliquer les résultats généraux que nous avons rappelés dans le chapitre précédent au cas particulier du \mathbb{C} -espace préhilbertien \mathcal{D} . Nous noterons $\|\cdot\|_2$ la norme hermitienne associée au produit scalaire

sur \mathcal{D} , donnée par :

$$\forall f \in \mathcal{D}, \quad \|f\|_2 = \sqrt{\frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt},$$

et $(e_n)_{n \in \mathbb{Z}}$ la famille orthonormale dans \mathcal{D} définie par $e_n : \mathbb{R} \rightarrow \mathbb{C}, x \mapsto e^{inx}$.
Considérons la suite $(S_n(f))_{n \in \mathbb{N}}$ des sommes partielles associées à la série de Fourier de f , c'est-à-dire

$$\forall n \in \mathbb{N}, \quad \forall x \in \mathbb{R}, \quad S_n(f)(x) := \sum_{k=-n}^n c_k(f) e^{ikx}.$$

Nous remarquons que $S_n(f) = \sum_{k=-n}^n c_k(f) e_k$, avec :

$$c_k(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt = \langle e_k, f \rangle.$$

Pour tout $n \in \mathbb{N}$, considérons le sous-espace de \mathcal{D} défini par $F_n := \text{Vect}(e_k)_{-n \leq k \leq n}$.
Par analogie avec la notion de degré d'un polynôme dans $\mathbb{C}[X]$, un élément de F_n est appelé polynôme trigonométrique de degré $\leq n$.

Proposition 10.5.1 *Pour tout $n \in \mathbb{N}$, nous avons $F_n \oplus F_n^\perp = \mathcal{D}$ et la projection orthogonale p_n sur F_n vérifie :*

$$\forall f \in \mathcal{D}, \quad p_n(f) = \sum_{k=-n}^n c_k(f) e_k = S_n(f).$$

Ainsi, pour tout $f \in \mathcal{D}$, $S_n(f)$ est l'unique meilleure approximation de f dans le sous-espace F_n , et on a les égalités :

$$\inf_{g \in F_n} \|f - g\|_2^2 = \|f - S_n(f)\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |f|^2 - \sum_{k=-n}^n |c_k(f)|^2.$$

Proposition 10.5.2 (Inégalité de Bessel) *Pour toute application $f \in C_M(\mathbb{R}, \mathbb{C})$, 2π -périodique, la série $\sum_{n \in \mathbb{Z}} |c_n(f)|^2$ converge et :*

$$\sum_{n=-\infty}^{+\infty} |c_n(f)|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} |f|^2$$

Démonstration: Considérons la régularisée $\tilde{f} \in \mathcal{D}$ associée à f selon la formule (9.1) page 169. D'après ce que nous avons établi précédemment, l'inégalité de Bessel est vérifiée pour \tilde{f} . Or, f et \tilde{f} ne diffèrent sur $[0, 2\pi]$ qu'en un nombre fini de points, nous avons $c_n(\tilde{f}) = c_n(f)$ pour tout $n \in \mathbb{Z}$, ainsi que :

$$\int_0^{2\pi} |\tilde{f}|^2 = \int_0^{2\pi} |f|^2,$$

ce qui nous permet de conclure. \square

Corollaire 10.5.3 (Lemme de Riemann-Lebesgue) *Pour toute application $f \in C_M(\mathbb{R}, \mathbb{C})$, 2π -périodique,*

$$\lim_{n \rightarrow \pm\infty} c_n(f) = 0.$$

Grâce au *théorème de Fejér*, nous avons établi le Corollaire 10.4.5 énonçant que la suite orthonormale $(e_n)_{n \in \mathbb{Z}}$ est totale dans le \mathbb{C} -espace préhilbertien de Dirichlet \mathcal{D} . Nous en déduisons :

Théorème 10.5.4 (Égalité de Parseval) *Soit $f \in C_M(\mathbb{R}, \mathbb{C})$ une application 2π -périodique. Alors, nous avons la convergence :*

$$\|f - S_n(f)\|_2 \rightarrow 0.$$

Ceci implique notamment que la série $\sum_{n \in \mathbb{Z}} |c_n(f)|^2$ converge et a pour somme

$$\sum_{n=-\infty}^{+\infty} |c_n(f)|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f|^2. \quad (10.6)$$

Remarque. L'égalité de Parseval s'écrit encore :

$$\frac{|a_0|^2}{4} + \frac{1}{2} \sum_{n=1}^{+\infty} |a_n(f)|^2 + \frac{1}{2} \sum_{n=1}^{+\infty} |b_n(f)|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f|^2. \quad (10.7)$$

Exercice : Soit $a \in \mathbb{C}$. Notons f la fonction périodique de période 2π telle que, pour tout $x \in [0, 2\pi[$, on ait $f(x) = e^{ax}$.

1. Notons $b = \Re a$. Montrer que si $b = 0$, alors on a $\sum_{k=-\infty}^{+\infty} |c_k(f)|^2 = 1$ et

$$\text{que si } b \neq 0, \text{ alors on a } \sum_{k=-\infty}^{+\infty} |c_k(f)|^2 = \frac{e^{4\pi b} - 1}{4\pi b}.$$

2. Calculer les coefficients de Fourier de f .
3. Montrer que pour tout nombre réel non nul a on a

$$\sum_{n=-\infty}^{+\infty} \frac{1}{n^2 + a^2} = \left(\frac{\pi}{a}\right) \left(\frac{e^{2a\pi} + 1}{e^{2a\pi} - 1}\right).$$

4. Montrer que pour tout nombre réel c non entier on a :

$$\sum_{n=-\infty}^{+\infty} \frac{1}{(n - c)^2} = \left(\frac{\pi}{\sin \pi c}\right)^2.$$

Solution :

1. La fonction f est périodique de période 2π et continue par morceaux. L'identité de Parseval donne $\frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt = \sum_{k=-\infty}^{+\infty} |c_k(f)|^2$, d'où le résultat puisque $|f(t)| = e^{bt}$.
2. On a $c_k(f) = \frac{1}{2\pi} \int_0^{2\pi} e^{(a-ik)t} dt$ donc $c_k(f) = 1$ si $a = ik$ et $c_k(f) = \frac{e^{2\pi(a-ik)} - 1}{2\pi(a-ik)} = \frac{e^{2\pi a} - 1}{2\pi(a-ik)}$ sinon.
3. Pour a réel non nul, il vient

$$\sum_{n=-\infty}^{+\infty} \frac{(e^{2\pi a} - 1)^2}{4\pi^2(a^2 + n^2)} = \frac{e^{4\pi a} - 1}{4\pi a}.$$

Écrivant $e^{4\pi a} - 1 = (e^{2\pi a} - 1)(e^{2\pi a} + 1)$ et simplifiant on trouve le résultat escompté.

4. Posant $a = ic$, il vient $c_n(f) = \frac{e^{i\pi c} \sin \pi c}{\pi(c-n)}$, donc $1 = \frac{\sin^2 \pi c}{\pi^2} \sum_{n \in \mathbb{Z}} \frac{1}{(c-n)^2}$.

Chapitre 11

Séries entières

Bibliographie :

- DANTZER chapitre 18 ;
- GOURDON Analyse (2ème édition) chapitre 4, paragraphe 4 ;
- RAMIS-DESCHAMPS-ODOUX tome 4, chapitre 3.

Nous allons nous placer de préférence sur $\mathbb{K} = \mathbb{C}$. Pour traiter le cas $\mathbb{K} = \mathbb{R}$, nous raisonnerons simplement par restriction.

11.1 L'algèbre des séries entières

Rappelons qu'une série $\sum f_n$ d'applications de \mathbb{C} dans \mathbb{C} est définie comme la suite d'applications $(\sum_{k=0}^n f_k)_{n \in \mathbb{N}}$ sans préjuger de la convergence de la série $\sum f_n(z)$ en un point $z \in \mathbb{C}$. Dans cette section, les définitions que nous allons donner sont purement algébriques. Les questions de convergence seront abordées dans la section suivante.

Définition 11.1.1 *Une série entière est une série d'applications de la forme $\sum a_n z^n$, où $(a_n) \in \mathbb{C}^{\mathbb{N}}$.*

Remarque. Avec un abus d'écriture couramment employé, nous avons noté ci-dessus $a_n z^n$ l'application $z \mapsto a_n z^n$ de \mathbb{C} dans \mathbb{C} .

Définition 11.1.2 *Nous définissons les trois opérations suivantes sur les séries entières :*

- Somme $\sum a_n z^n + \sum b_n z^n := \sum (a_n + b_n) z^n$;
- Produit par un scalaire *Pour tout $\lambda \in \mathbb{C}$, $\lambda \sum a_n z^n = \sum (\lambda a_n) z^n$;*

— Produit de Cauchy $(\sum a_n z^n)(\sum b_n z^n) = \sum c_n z^n$, où $c_n := \sum_{k=0}^n a_k b_{n-k}$.

Théorème 11.1.3 *Muni de ces trois opérations, l'ensemble des séries entières est une \mathbb{C} -algèbre commutative.*

Nous pouvons de même, d'une façon purement algébrique, définir la dérivée d'une série entière, sans aucunement préjuger de sa convergence.

Définition 11.1.4 *On appelle série entière dérivée de $\sum a_n z^n$ et l'on note $(\sum a_n z^n)'$ la série entière $\sum (n+1)a_{n+1} z^n$.*

Remarque. En itérant le processus, pour tout $p \in \mathbb{N}^*$, nous pouvons définir la série entière dérivée p -ième de $\sum a_n z^n$ par :

$$\left(\sum a_n z^n\right)^{(p)} = \sum (n+p)(n+p-1) \cdots (n+1) a_{n+p} z^n = \sum \frac{(n+p)!}{n!} a_{n+p} z^n.$$

Exercice. Montrer l'égalité :

$$\left[\left(\sum a_n z^n\right)\left(\sum b_n z^n\right)\right]' = \left(\sum a_n z^n\right)' \left(\sum b_n z^n\right) + \left(\sum a_n z^n\right) \left(\sum b_n z^n\right)'$$

11.2 Rayon de convergence

Définition 11.2.1 *On appelle rayon de convergence de la série entière $\sum a_n z^n$ l'élément de \mathbb{R}_+ suivant :*

$$R = \sup \left\{ r \geq 0, \left(|a_n| r^n \right)_{n \in \mathbb{N}} \text{ est bornée} \right\}.$$

Pour $r \geq 0$, nous noterons $D(r) := \{z \in \mathbb{C}, |z| \leq r\}$ le disque fermé, $\overset{\circ}{D}(r) := \{z \in \mathbb{C}, |z| < r\}$ le disque ouvert, et enfin $\mathcal{C}(r) := \{z \in \mathbb{C}, |z| = r\}$ le cercle, tous de centre 0 et de rayon r .

Nous conviendrons que $D(+\infty) = \overset{\circ}{D}(+\infty) = \mathbb{C}$.

Proposition 11.2.2 (Lemme d'Abel) *Soit $\sum a_n z^n$ une série entière de rayon de convergence R .*

- Si $|z| < R$, alors $\sum a_n z^n$ converge absolument.
- Si $0 < r < R$, alors $\sum a_n z^n$ converge normalement sur $D(r)$.
- Si $|z| > R$, alors $\sum a_n z^n$ diverge.

Démonstration: Comparaison avec une série géométrique. \square

Le disque ouvert $\overset{\circ}{D}(R)$ est appelé *disque de convergence* de la série entière $\sum a_n z^n$. Lorsque $R < +\infty$, le cercle $\mathcal{C}(R)$ est appelé *cercle d'incertitude* car, comme nous le verrons sur des exemples, différents comportements de la série entière y sont possibles.

Passons maintenant au calcul pratique de R . Dans la formule suivante, nous adopterons les conventions $(+\infty)^{-1} = 0$ et $0^{-1} = +\infty$.

Proposition 11.2.3 (Formule d'Hadamard) *Le rayon de convergence de la série entière $\sum a_n z^n$ est donné par l'égalité :*

$$R = \left(\overline{\lim} |a_n|^{1/n} \right)^{-1}.$$

Démonstration: Rappelons la *règle de Cauchy* pour une série numérique $\sum u_n$: posons $\lambda = \overline{\lim} |u_n|^{1/n}$; si $\lambda < 1$ alors $\sum u_n$ est absolument convergente, si $\lambda > 1$, alors $\sum u_n$ est divergente.

En considérant $u_n = a_n z^n$, on en déduit facilement le résultat annoncé. \square

Bien sûr, si la suite $(|a_n|^{1/n})_{n \in \mathbb{N}^*}$ converge, nous avons $R = \left(\lim |a_n|^{1/n} \right)^{-1}$; c'est la *règle de Cauchy* pour les séries entières.

Exercice. Que vaut le rayon de convergence de $\sum 2^n z^{2^n}$?

Solution : Ici, $a_{2p} = 2^p$ et $a_{2p+1} = 0$ pour tout $p \in \mathbb{N}$ donc :

$$\overline{\lim}_n |a_n|^{1/n} = \overline{\lim}_p (2^p)^{\frac{1}{2p}} = \sqrt{2},$$

d'où $R = 1/\sqrt{2}$.

Exercice. Soient $\sum a_n z^n$ et $\sum b_n z^n$ des séries entières, de rayons respectifs R et R' . Que peut-on dire du rayon R'' de la série entière $\sum a_n b_n z^n$?

Solution : Nous avons l'inégalité :

$$\overline{\lim} |a_n b_n|^{1/n} \leq \overline{\lim} |a_n|^{1/n} \overline{\lim} |b_n|^{1/n},$$

d'où, par Hadamard, $R'' \geq RR'$.

Cette dernière inégalité peut être large (prendre $a_n \equiv 1$ et $b_n \equiv 1$) ou stricte, comme nous le constatons sur l'exemple suivant.

$\sum z^{2^n}$ et $\sum z^{2^{n+1}}$ ont pour rayon de convergence $R = R' = 1$ mais nous avons ici $a_n b_n = 0$ pour tout $n \in \mathbb{N}$, d'où $R'' = +\infty$.

Proposition 11.2.4 (Règle de d'Alembert) Soit $\sum a_n z^n$ une série entière telle que :

$$\lim \frac{|a_{n+1}|}{|a_n|} = l \in \bar{\mathbb{R}}_+.$$

Alors, son rayon de convergence vaut $R = l^{-1}$.

Démonstration: Règle de d'Alembert pour les séries numériques. □

Si la règle de Cauchy ne permet pas de conclure pour une certaine série entière, il est inutile d'essayer la règle de d'Alembert, comme le montre le résultat suivant.

Lemme 11.2.5 Nous avons l'implication :

$$\lim \frac{|a_{n+1}|}{|a_n|} = l \in \bar{\mathbb{R}}_+ \Rightarrow \lim |a_n|^{1/n} = l.$$

Démonstration: Si $l_1 > l$, d'après la proposition précédente, $\sum |a_n|/l_1^n$ converge donc son terme général tend vers 0. En particulier, il existe un rang $N_1 \in \mathbb{N}$ tel que, pour tout $n \geq N_1$, $|a_n| \leq l_1^n$.

Si $0 < l_0 < l$, $\sum l_0^n/|a_n|$ converge et donc il existe un rang $N_0 \in \mathbb{N}$ tel que, pour tout $n \geq N_0$, $l_0^n \leq |a_n|$.

Si $l_0 < l < l_1$, il existe donc un rang $N := \max(N_0, N_1) \in \mathbb{N}$ tel que, pour tout $n \geq N$, $l_0^n \leq |a_n| \leq l_1^n$, ce qui s'écrit encore $l_0 \leq |a_n|^{1/n} \leq l_1$, d'où :

$$l_0 \leq \underline{\lim} |a_n|^{1/n} \leq \overline{\lim} |a_n|^{1/n} \leq l_1.$$

Nous laissons au lecteur les détails de la conclusion (traiter à part les cas $l = 0$ et $l = +\infty$). □

Exemples :

- La série entière $\sum z^n/n!$ a pour rayon $R = +\infty$.
- La série entière $\sum n! z^n$ a un rayon de convergence nul.
- Pour tout $\alpha \in \mathbb{R}$, la série entière $\sum \frac{z^n}{n^\alpha}$ a pour rayon $R = 1$

Attardons-nous sur ce dernier exemple pour étudier le comportement de la série entière sur le cercle d'incertitude $\mathcal{C}(1)$ en fonction des valeurs de α .

Si $\alpha \leq 0$, nous avons $|\frac{z^n}{n^\alpha}| = n^{-\alpha} \geq 1$ pour tout $n \geq 1$ donc la série entière diverge grossièrement sur $\mathcal{C}(1)$.

Si $\alpha > 1$, la série entière converge absolument sur $\mathcal{C}(1)$ comme le montre

l'exemple de Riemann $\sum 1/n^\alpha$.

Si $0 < \alpha \leq 1$, en $z = 1$ la série diverge, toujours d'après l'exemple de Riemann. En $z \neq 1$, *i.e.* en $z = \exp(i\theta)$ avec $\theta \in]0, 2\pi[$, la série entière converge comme le montre une application de la règle d'Abel pour les séries.

Bien que la formule d'Hadamard soit universelle, elle ne permet pas toujours un calcul explicite du rayon de convergence. Il arrive qu'il soit plus simple de raisonner directement, comme le lecteur le constatera dans les deux exercices suivants [DAN 304-305].

Exercice. Que vaut le rayon de convergence de $\sum a_n z^n$, où a_n est la n -ième décimale du nombre e ?

Solution : Comme la suite (a_n) est bornée, nous avons $R \geq 1$. D'autre part, comme e n'est pas un nombre décimal, nous avons :

$$\forall n_0 \in \mathbb{N}, \quad \exists n \geq n_0, \quad a_n \geq 1.$$

En particulier, la suite (a_n) ne tend pas vers 0 donc $\sum a_n$ diverge et $R \leq 1$. Finalement, $R = 1$.

Exercice. Même question avec $a_n = \#\{k \leq n, k \text{ nombre premier}\}$.

Solution : Pour tout $n \geq 2$, on a $a_n \geq 1$ donc $\sum a_n$ diverge et $R \leq 1$. Par ailleurs, nous avons $0 \leq a_n \leq n$ pour tout $n \in \mathbb{N}$ d'où, pour tout $r \in [0, 1[$, $0 \leq a_n r^n \leq n r^n$ ce qui implique : $a_n r^n \rightarrow 0$. En particulier, $(a_n r^n)$ est bornée et donc $R \geq r$. Comme $r \in [0, 1[$ était arbitraire, nous en déduisons $R \geq 1$. Finalement, $R = 1$.

Proposition 11.2.6 Soient $\sum a_n z^n$, $\sum b_n z^n$ des séries entières de rayons resp. R_a et R_b . Nous noterons R_{a+b} le rayon de convergence de la série somme $\sum (a_n + b_n) z^n$. Alors,

- Si $R_a \neq R_b$, $R_{a+b} = \min(R_a, R_b)$;
- Si $R_a = R_b$, $R_{a+b} \geq R_a$.

Dans ce dernier cas, l'inégalité peut être stricte, comme le lecteur le constatera en prenant $a_n = 1+2^n$ et $b_n = 1-2^n$: on a $R_a = R_b = \frac{1}{2}$ et $R_{a+b} = 1 > \frac{1}{2}$.

Proposition 11.2.7 Soient $\sum a_n z^n$, $\sum b_n z^n$ des séries entières de rayons resp. R_a et R_b . Nous noterons $\sum c_n z^n$ leur produit de Cauchy et R_c son rayon de convergence. Alors $R_c \geq \min(R_a, R_b)$ et nous avons :

$$\forall z \in \mathbb{C}, \quad |z| < \min(R_a, R_b) \Rightarrow \sum_{n=0}^{+\infty} c_n z^n = \left(\sum_{n=0}^{+\infty} a_n z^n \right) \left(\sum_{n=0}^{+\infty} b_n z^n \right).$$

Démonstration: Il s'agit du produit de Cauchy de deux séries numériques absolument convergentes. □

On peut avoir $R_c > \min(R_a, R_b)$ comme le montre l'exemple suivant : $a_n \equiv 1$, $b_0 = 1$, $b_1 = -1$ et $b_n = 0$ pour tout $n \geq 2$. Le lecteur vérifiera que $R_a = 1$, $R_b = +\infty$ et $R_c = +\infty$.

Nous allons maintenant établir un résultat qui nous sera très utile dans la section suivante.

Proposition 11.2.8 Une série entière et sa série dérivée ont même rayon de convergence.

Démonstration: Notons R et R' les rayons de convergence des séries entières $\sum a_n z^n$ et $\sum (n+1)a_{n+1} z^n$.

Si $r > R$, pour tout entier $n \geq r-1$ nous avons $(n+1)|a_{n+1}|r^n \geq |a_{n+1}|r^{n+1}$, cette dernière suite étant non bornée par définition de R . Nous en déduisons :

$$\{r \geq 0, ((n+1)|a_{n+1}|r^n)_{n \in \mathbb{N}} \text{ est bornée} \} \subset [0, R], \text{ d'où } R' \leq R.$$

Si $0 \leq r < R$, choisissons $\rho \in]0, R-r[$ de sorte que $\sum |a_n|(r+\rho)^n$ converge. D'après la formule du binôme, nous avons l'inégalité $(r+\rho)^{n+1} \geq (n+1)r^n \rho$ pour tout $n \in \mathbb{N}$, si bien que :

$$\forall n \in \mathbb{N}, \quad (n+1)|a_{n+1}|r^n \leq \frac{1}{\rho} |a_{n+1}|(r+\rho)^{n+1},$$

ce qui implique la convergence de $\sum (n+1)a_{n+1} r^n$, d'où $R' \geq R$. □

11.3 Continuité et dérivabilité

Nous noterons $S(z) = \sum_{n=0}^{+\infty} a_n z^n$ la somme de la série entière en tout point z où elle converge. Dans toute la suite, nous supposons $R > 0$, si bien que l'application S est définie au moins sur le disque ouvert $\overset{\circ}{D}(R)$.

Proposition 11.3.1 *La somme S est continue sur $\overset{\circ}{D}(R)$.*

Ceci résulte de la convergence normale donc uniforme de la série entière sur $D(r)$, pour tout $0 < r < R$.

Proposition 11.3.2 (Théorème d'Abel radial) *On suppose qu'il existe $z_0 \in \mathbb{C}$ tel que $|z_0| = R$ et $\sum a_n z_0^n$ converge. Alors la série entière $\sum a_n z^n$ converge uniformément sur le segment $\{tz_0, t \in [0, 1]\}$ et l'on a en particulier la propriété de continuité suivante :*

$$\lim_{t \rightarrow 1-} S(tz_0) = \sum_{n=0}^{+\infty} a_n z_0^n.$$

Démonstration: Posons, pour tout $n \in \mathbb{N}$, $b_n = a_n z_0^n$. La série entière $\sum b_n z^n$ admet 1 pour rayon de convergence et, en notant S_b sa somme, nous avons l'équivalence :

$$\lim_{t \rightarrow 1-} S_b(t) = \sum_{n=0}^{+\infty} b_n \iff \lim_{t \rightarrow 1-} S(tz_0) = \sum_{n=0}^{+\infty} a_n z_0^n.$$

Nous pouvons donc supposer, sans perte de généralité, $z_0 = R = 1$. Le lecteur trouvera la suite de la démonstration dans [SK 59].

□

Revenons maintenant au cas général d'une série entière de rayon de convergence $R > 0$ et de somme S . Si nous raisonnions sur $\mathbb{K} = \mathbb{R}$, nous pourrions montrer que l'application $S :]-R, R[\rightarrow \mathbb{C}$ est dérivable et même de classe C^∞ . En fait, nous allons prouver un résultat beaucoup plus fort en nous plaçant sur $\mathbb{K} = \mathbb{C}$.

Définition 11.3.3 *Soit U un ouvert de \mathbb{C} , $z_0 \in U$ et $f : U \rightarrow \mathbb{C}$. On dit que f est holomorphe (ou encore dérivable au sens complexe) au point z_0 si l'application $z \mapsto \frac{f(z) - f(z_0)}{z - z_0}$ de $U \setminus \{z_0\}$ dans \mathbb{C} admet une limite lorsque $z \rightarrow z_0$. Dans ce cas, on dit que cette limite est la dérivée de f au point z_0 et l'on note :*

$$f'(z_0) = \lim_{\substack{z \rightarrow z_0 \\ z \neq z_0}} \frac{f(z) - f(z_0)}{z - z_0}.$$

Enfin, on dit que f est holomorphe sur U si f est holomorphe en tout point $z_0 \in U$.

Théorème 11.3.4 *L'application somme S est holomorphe sur $\overset{\circ}{D}(R)$ et*

$$\forall z \in \overset{\circ}{D}(R), \quad S'(z) = \sum_{n=0}^{+\infty} (n+1)a_{n+1} z^n.$$

Démonstration: Soit $z_0 \in \overset{\circ}{D}(R)$; nous choisissons $r \in]|z_0|, R[$.

Nous pouvons alors écrire, pour tout $z \in \overset{\circ}{D}(r) - \{z_0\}$:

$$\frac{S(z) - S(z_0)}{z - z_0} = \sum_{n=0}^{+\infty} a_n \frac{z^n - z_0^n}{z - z_0} = \sum_{n=1}^{+\infty} a_n (z^{n-1} + z_0 z^{n-2} + \cdots + z_0^{n-2} z + z_0^{n-1})$$

Considérons $f_n : \overset{\circ}{D}(r) \rightarrow \mathbb{C}, z \mapsto a_n (z^{n-1} + z_0 z^{n-2} + \cdots + z_0^{n-2} z + z_0^{n-1})$; pour tout $n \in \mathbb{N}^*$, f_n est manifestement une application continue.

La série $\sum f_n$ est normalement convergente sur $\overset{\circ}{D}(r)$ car

$$\sup_{|z| < r} |f_n(z)| \leq n |a_n| r^{n-1}$$

et la série

$$\sum_{n \geq 1} n |a_n| r^{n-1} = \sum_{n \in \mathbb{N}} (n+1) |a_{n+1}| r^n$$

converge puisque $r < R$ et la série entière dérivée $\sum (n+1)a_{n+1} z^n$ a pour rayon de convergence R .

Ainsi, $\sum_{n=1}^{+\infty} f_n$ est continue sur $\overset{\circ}{D}(r)$ et, en particulier,

$$\lim_{\substack{z \rightarrow z_0 \\ z \neq z_0}} \sum_{n=1}^{+\infty} f_n(z) = \sum_{n=1}^{+\infty} f_n(z_0),$$

ce qui s'écrit encore :

$$\lim_{\substack{z \rightarrow z_0 \\ z \neq z_0}} \frac{S(z) - S(z_0)}{z - z_0} = \sum_{n=1}^{+\infty} n a_n z_0^{n-1} = \sum_{n=0}^{+\infty} (n+1) a_{n+1} z_0^n.$$

□

Corollaire 11.3.5 *L'application somme S est infiniment dérivable au sens complexe sur $\overset{\circ}{D}(R)$ et ses dérivées successives sont données par :*

$$\forall k \in \mathbb{N}^*, \quad \forall z \in \overset{\circ}{D}(R), \quad S^{(k)}(z) = \sum_{n=0}^{+\infty} \frac{(n+k)!}{n!} a_{n+k} z^n.$$

En particulier, pour tout $k \in \mathbb{N}$, $S^{(k)}(0) = k! a_k$.

Par restriction à \mathbb{R} , ce corollaire entraîne évidemment que $S :]-R, R[\rightarrow \mathbb{C}$ est de classe C^∞ et que $a_k = \frac{S^{(k)}(0)}{k!}$, pour tout $k \in \mathbb{N}$.

Remarque. En reprenant les notations de la Définition 11.3.3, et en identifiant \mathbb{C} à \mathbb{R}^2 (par isomorphisme canonique), nous pouvons considérer que l'application f est définie sur un ouvert U de \mathbb{R}^2 . Le lecteur pourra vérifier que f est holomorphe au point $z_0 = x_0 + iy_0 \simeq (x_0, y_0)$ si, et seulement si, elle est différentiable en ce point, de différentielle df_{z_0} nulle ou égale à une similitude directe (composée d'une rotation par une homothétie).

Indication : Écrire $f(z) = f(a) + f'(a)(z-a) + (z-a)\epsilon(z-a)$, avec $\lim_0 \epsilon = 0$. Si $c = a + ib$, l'application $z \mapsto cz$ de \mathbb{C} dans \mathbb{C} peut être identifiée à la similitude de \mathbb{R}^2 définie par la matrice :

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

puisque $(a + ib)(x + iy) = (ax - by) + i(bx + ay)$.

11.4 Principe des zéros isolés

Théorème 11.4.1 *Supposons qu'il existe une suite $(z_p)_{p \in \mathbb{N}}$ à valeurs dans $\overset{\circ}{D}(R) \setminus \{0\}$ telle que $z_p \rightarrow 0$ et $S(z_p) = 0$ pour tout $p \in \mathbb{N}$. Alors, $a_n = 0$ pour tout $n \in \mathbb{N}$ et donc $S \equiv 0$.*

Démonstration: On raisonne par l'absurde en supposant que la suite (a_n) n'est pas identiquement nulle, ce qui nous permet de définir :

$$q = \min\{n \in \mathbb{N}, a_n \neq 0\}.$$

Nous avons donc, pour tout $z \in \overset{\circ}{D}(R)$,

$$S(z) = \sum_{n=q}^{+\infty} a_n z^n = \sum_{n=0}^{+\infty} a_{n+q} z^{n+q} = z^q \sum_{n=0}^{+\infty} a_{n+q} z^n.$$

Appelons $T(z)$ la somme de la série $\sum a_{n+q} z^n$, en remarquant que celle-ci a même rayon de convergence $R > 0$ que $\sum a_n z^n$ puisque ces deux séries convergent ou divergent en même temps.

Notre hypothèse implique que, pour tout $p \in \mathbb{N}$, $T(z_p) = 0$ et donc, par continuité de T en 0, $T(0) = a_q = 0$, d'où une contradiction.

Nous en déduisons que (a_n) est identiquement nulle. \square

Corollaire 11.4.2 Soient $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières de rayons de convergence respectifs $R_a > 0, R_b > 0$ et de sommes respectives S_a, S_b . Nous posons $R = \inf(R_a, R_b)$ et nous supposons qu'il existe une suite $(z_p)_{p \in \mathbb{N}}$ à valeurs dans $\overset{\circ}{D}(R) \setminus \{0\}$ telle que $z_p \rightarrow 0$ et $\forall p \in \mathbb{N}, S_a(z_p) = S_b(z_p)$. Alors, $a_n = b_n$ pour tout $n \in \mathbb{N}$ et donc $S_a = S_b$.

Application 11.4.3 La transformée de Laplace de la densité de Gauss est donnée par l'égalité :

$$\forall z \in \mathbb{C}, \quad \int_{-\infty}^{+\infty} e^{zx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = e^{\frac{z^2}{2}}. \quad (11.1)$$

Démonstration: La densité de Gauss est l'application $g : \mathbb{R} \rightarrow \mathbb{R}_+$, définie par :

$$\forall x \in \mathbb{R}, \quad g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Nous noterons $\mathcal{L}g : \mathbb{C} \rightarrow \mathbb{C}$ sa transformée de Laplace, définie par :

$$\forall z \in \mathbb{C}, \quad \mathcal{L}g(z) = \int_{-\infty}^{+\infty} e^{zx} g(x) dx = \int_{-\infty}^{+\infty} e^{zx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

En appliquant les théorèmes de comparaison, on vérifie que $\mathcal{L}g$ est bien définie sur \mathbb{C} tout entier.

Dans un premier temps, nous considérons le cas z réel. Par simple changement de variable, nous faisons le calcul suivant :

$$\forall z \in \mathbb{R}, \quad \mathcal{L}g(z) = \int_{\mathbb{R}} e^{zx} g(x) dx = e^{\frac{z^2}{2}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-z)^2}{2}} dx = e^{\frac{z^2}{2}}. \quad (11.2)$$

Prenant maintenant $z \in \mathbb{C}$, nous avons :

$$\int_{\mathbb{R}} e^{zx} g(x) dx = \int_{\mathbb{R}} \lim_{N \rightarrow \infty} \sum_{n=0}^N \frac{z^n x^n}{n!} g(x) dx.$$

Or, pour tout $N \in \mathbb{N}$ et tout $x \in \mathbb{R}$, nous avons :

$$\left| \sum_{n=0}^N \frac{z^n x^n}{n!} g(x) \right| \leq e^{|z||x|} g(x)$$

Le membre de droite de cette inégalité constituant un majorant intégrable, le théorème de convergence dominée nous donne :

$$\mathcal{L}g(z) = \int_{\mathbb{R}} e^{zx} g(x) dx = \sum_{n=0}^{+\infty} \frac{z^n}{n!} \int_{\mathbb{R}} x^n g(x) dx. \quad (11.3)$$

Ceci prouve que la série entière qui constitue le membre de droite de (11.3) est de rayon de convergence infini et que sa somme S n'est autre que $\mathcal{L}g$. Remarquons maintenant que le membre de droite de (11.1) est développable en série entière sur \mathbb{C} , puisque :

$$\forall z \in \mathbb{C}, \quad \exp\left(\frac{z^2}{2}\right) = \sum_{p=0}^{+\infty} \frac{z^{2p}}{2^p p!}. \quad (11.4)$$

La série entière qui apparaît dans le membre de droite de (11.4) est de rayon de convergence infini et, si nous notons sa somme T , nous avons donc, pour tout $z \in \mathbb{C}$, $T(z) = \exp(z^2/2)$.

Revenant alors à (11.2), nous constatons que $S(z) = T(z)$ pour tout $z \in \mathbb{R}$. Il suffit de prendre une suite de réels non nuls $(z_p)_{p \in \mathbb{N}}$, par exemple donnée par $z_p = 1/(p+1)$, pour déduire du corollaire précédent l'égalité de S et T sur \mathbb{C} tout entier, d'où notre conclusion. \square

Remarque : En identifiant les coefficients des séries entières qui constituent les membres de droite de (11.3) et (11.4), nous obtenons les *moments* de la densité de Gauss :

$$\forall p \in \mathbb{N}, \quad \int_{-\infty}^{+\infty} x^{2p} g(x) dx = \frac{(2p)!}{2^p p!} \quad \text{et} \quad \int_{-\infty}^{+\infty} x^{2p+1} g(x) dx = 0.$$

11.5 Formule de Cauchy, égalité de Parseval

Théorème 11.5.1 *En reprenant les notations précédentes, on a l'égalité :*

$$\forall r \in]0, R[, \quad \forall n \in \mathbb{N}, \quad 2\pi a_n r^n = \int_0^{2\pi} S(re^{i\theta}) e^{-ni\theta} d\theta.$$

Démonstration: Par convergence normale, le membre de droite vaut :

$$\sum_{p=0}^{+\infty} a_p r^p \int_0^{2\pi} e^{i(p-n)\theta} d\theta.$$

Or, cette dernière intégrale est nulle pour tout $p \neq n$, et vaut 2π pour $p = n$, ce qui nous permet de conclure. \square

Application 11.5.2 (Théorème de Liouville) *On suppose $R = +\infty$. Si S est bornée sur \mathbb{C} , alors S est constante.*

Démonstration: Posons $M = \sup_{z \in \mathbb{C}} |S(z)| < +\infty$.

La formule de Cauchy et l'inégalité triangulaire dans une intégrale nous donnent, pour tout $r > 0$ et tout $n \in \mathbb{N}$, $2\pi|a_n|r^n \leq 2\pi M$.

En particulier, pour tout $n \in \mathbb{N}^*$, $|a_n| \leq \frac{M}{r^n}$ et il suffit de faire tendre r vers l'infini pour conclure. \square

Exercice. [GOU 248] S'il existe $m \in \mathbb{N}^*$ et $C > 0$ tels que $|S(z)| \leq C|z|^m$ pour tout $z \in \mathbb{C}$, alors S est un polynôme de degré inférieur ou égal à m . L'idée est d'introduire $g(z) = \sum_{n=m}^{+\infty} a_n z^{n-m}$ de sorte que :

$$\forall z \in \mathbb{C}^* \quad g(z) = \frac{S(z) - a_0 - a_1 z - \cdots - a_{m-1} z^{m-1}}{z^m}$$

On en déduit que g est entière et bornée donc constante, ce qui signifie que $a_n = 0$ dès que $n > m$.

Terminons cette section par une application de l'égalité de Parseval dans le cadre des séries entières.

Théorème 11.5.3 (Égalité de Parseval) *En reprenant les notations précédentes, nous avons :*

$$\forall r \in]0, R[, \quad \sum_{n=0}^{+\infty} |a_n|^2 r^{2n} = \frac{1}{2\pi} \int_0^{2\pi} |S(re^{i\theta})|^2 d\theta.$$

Démonstration: Pour $r \in]0, R[$ fixé, définissons $f : \mathbb{R} \rightarrow \mathbb{C}$ par :

$$\forall \theta \in \mathbb{R}, \quad f(\theta) = S(re^{i\theta}) = \sum_{n=0}^{+\infty} a_n r^n e^{in\theta},$$

de sorte que f est clairement continue et 2π -périodique.

Grâce à la convergence de la série $\sum_{n=0}^{+\infty} |a_n| r^n$, on calcule facilement les coefficients de Fourier de l'application f :

$$\forall n \in \mathbb{N}, \quad c_n(f) = a_n r^n \quad \text{et} \quad \forall n \in \mathbb{Z}_-^*, \quad c_n(f) = 0.$$

On conclut alors grâce au Théorème 10.5.4, page 195. \square

Remarque. On peut aussi démontrer le théorème de Liouville à partir de ce dernier théorème.

11.6 Fonctions développables en série entière

Voir [DAN 313 ss.] [GOU 240ss.]

Les développements en séries entières des fonctions usuelles sont rappelés dans [GOU 242].

Probabilités : bibliographie

Ouvrages de référence pour l'agrégation interne

- ESC** Probabilités et Statistiques pour le CAPES et l'Agrégation interne. Escoffier. Éditions Ellipses. 2006
- MEY** Probabilités, cours et exercices corrigés. Agrég. interne, Classes prépa., Licence, Capes. Tome 1. Meyre. Éd. Calvage & Mounet. 2016
- OUV¹** Probabilités 1, Capes et Agrégation. Ouvrard. Éditions Cassini.

Ouvrages plus difficiles

- BAR** Probabilité. Barbe, Ledoux. EDP Sciences. 2007.
- COT** Exercices de Probabilités. Cottrell et al. Éditions Cassini. 2005.
- FOA** Calcul des Probabilités. Foata, Fuchs. Éditions Dunod. 2ème édition. 1998

Autres ouvrages pour les corrigés des exercices

- BIL** Probability and Measure. Billingsley. Wiley.
- DAC¹ cours** Probabilités et Statistiques 1. Problèmes à temps fixe. Dacunha-Castelle, Duflo. 2e édition. Éditions Masson.
- DAC¹ exo** Exercices de Probabilités et Statistiques 1. Problèmes à temps fixe. Dacunha-Castelle, Duflo. 3e tirage corrigé. Éditions Masson.
- FEL¹** An Introduction to Probability Theory and its Applications. Feller. Wiley 3rd edition. Volume I
- FEL²** An Introduction to Probability Theory and its Applications. Feller. Wiley 2nd edition. Volume II
- OUV²** Probabilités 2, Maîtrise et Agrégation. Ouvrard. Éditions Cassini.
- REV** Probabilités. Revuz. Editions Hermann. Collection Méthodes.
- ROS** Initiation aux Probabilités. Ross. Presses polytechniques et universitaires romandes. Troisième édition.
- RUD** Analyse réelle et complexe. Rudin. Editions Masson.

Chapitre 12

Espaces mesurés. Espaces probabilisés.

Dans les deux premières sections de ce chapitre, nous allons définir la structure d'espace mesuré. Bien que la théorie abstraite de l'intégration de Lebesgue repose sur cette structure, nous ne la développerons pas ici car elle ne figure pas au programme officiel. En revanche, la structure d'espace mesuré va nous mener au modèle fondamental utilisé en calcul des probabilités, que l'on appelle *espace probabilisé*.

12.1 Clans et tribus.

12.1.1 Définitions.

Définition 12.1.1 Soit E un ensemble. On appelle clan sur E un ensemble \mathcal{C} de parties de E tel que les trois conditions suivantes soient satisfaites :

1. $E \in \mathcal{C}$
2. $A \in \mathcal{C} \implies A^c = E - A \in \mathcal{C}$
3. $(A, B) \in \mathcal{C}^2 \implies A \cup B \in \mathcal{C}$

Remarque. Un clan sur E est encore appelé algèbre (de Boole) de parties de E .

De cette définition nous déduisons les propriétés immédiates suivantes :

1. $\emptyset \in \mathcal{C}$
2. $(A, B) \in \mathcal{C}^2 \implies A \cap B \in \mathcal{C}$
3. $(A_1, \dots, A_n) \in \mathcal{C}^n \implies \bigcup_{i=1}^n A_i \in \mathcal{C}$ et $\bigcap_{i=1}^n A_i \in \mathcal{C}$

4. Si A et B sont des ensembles de \mathcal{C} , alors leur différence $A - B = A \cap B^c$ est encore un ensemble de \mathcal{C} . Un cas particulier est celui où $B \subset A$: on parle alors de différence propre.
5. Si A et B sont des ensembles de \mathcal{C} , alors leur différence symétrique $A \Delta B = (A \cup B) - (A \cap B) = (A - B) \cup (B - A)$ est encore un ensemble de \mathcal{C} .

Définition 12.1.2 Soit E un ensemble. On appelle tribu sur E un ensemble \mathcal{T} de parties de E tel que les trois conditions suivantes soient satisfaites :

1. $E \in \mathcal{T}$
2. $A \in \mathcal{T} \implies A^c \in \mathcal{T}$
3. $(A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}} \implies \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{T}$

Remarque. Une tribu sur E est encore appelée σ -algèbre (de Boole) de parties de E .

Comme $\emptyset = E^c \in \mathcal{T}$, la stabilité par union dénombrable nous donne aussi la stabilité de \mathcal{T} par union finie. Une tribu est donc un clan stable par passage à l'union dénombrable.

Une tribu est aussi stable par passage à l'intersection dénombrable puisque :

$$\bigcap_{n \in \mathbb{N}} A_n = \left[\bigcup_{n \in \mathbb{N}} A_n^c \right]^c$$

Exemples.

1. L'ensemble $\mathcal{P}(E)$ de toutes les parties de E est la plus grande tribu sur E (au sens de l'inclusion). On l'appelle tribu totale sur E .
2. L'ensemble $\{\emptyset, E\}$ est la plus petite tribu sur E . On l'appelle tribu grossière sur E .
3. Sur \mathbb{R} , l'ensemble \mathcal{C} des unions finies d'intervalles de la forme

$$]-\infty, a[\quad , \quad [a, b[\quad \text{ou} \quad [b, +\infty[$$

est un clan et non une tribu.

Le lecteur pourra montrer d'abord la stabilité de \mathcal{C} par intersection finie avant de prouver la stabilité par passage au complémentaire. Il pourra remarquer aussi que tout élément de \mathcal{C} a un nombre fini de composantes connexes et donc que :

$$\bigcup_{n \in \mathbb{N}} \left[n, n + \frac{1}{2} \right] \notin \mathcal{C}$$

12.1.2 Clans et tribus engendrés.

Proposition 12.1.3 *L'intersection d'une famille non vide $(\mathcal{C}_i)_{i \in I}$ de clans sur E est encore un clan sur E .*

L'intersection d'une famille non vide $(\mathcal{T}_i)_{i \in I}$ de tribus sur E est encore une tribu sur E .

Corollaire 12.1.4 *Soit \mathcal{E} un ensemble quelconque de parties de E . Il existe un plus petit clan sur E (au sens de l'inclusion) qui contienne \mathcal{E} ; on l'appelle clan engendré par \mathcal{E} et on le note $\mathcal{C}(\mathcal{E})$.*

Il existe une plus petite tribu sur E (au sens de l'inclusion) qui contienne \mathcal{E} ; on l'appelle tribu engendrée par \mathcal{E} et on la note $\sigma(\mathcal{E})$.

Démonstration: La famille des clans sur E contenant \mathcal{E} n'est pas vide car elle contient $\mathcal{P}(E)$. L'intersection de tous les clans sur E contenant \mathcal{E} est donc encore un clan, il contient bien sûr \mathcal{E} et est manifestement le plus petit au sens de l'inclusion ayant cette propriété. Un raisonnement similaire s'applique aux tribus sur E . □

Définition 12.1.5 *Soient E et F des ensembles et $f : E \rightarrow F$ une application. Pour tout $A \subset F$, on appelle image réciproque de A par f et l'on note $f^{-1}(A)$ le sous-ensemble de E défini par :*

$$f^{-1}(A) = \{x \in E, f(x) \in A\}$$

Remarques.

1. Nous utiliserons également la notation $\{f \in A\}$ pour l'image réciproque de A par f .
2. Si \mathcal{F} est un ensemble de parties de F , nous noterons $f^{-1}(\mathcal{F})$ l'ensemble des parties de E de la forme $f^{-1}(A)$, avec $A \in \mathcal{F}$.

Proposition 12.1.6 *Soient E et F des ensembles et $f : E \rightarrow F$ une application.*

Si \mathcal{U} est une tribu sur F , alors $f^{-1}(\mathcal{U})$ est une tribu sur E appelée tribu image réciproque de \mathcal{U} par f .

Démonstration: Nous constatons d'abord que $E = f^{-1}(F) \in f^{-1}(\mathcal{U})$. Pour tout $U \in \mathcal{U}$, nous avons l'égalité

$$[f^{-1}(U)]^c = f^{-1}(U^c)$$

d'où la stabilité de $f^{-1}(\mathcal{U})$ par passage au complémentaire.
Enfin, si $(U_n) \in \mathcal{U}^{\mathbb{N}}$, l'égalité

$$\bigcup_{n \in \mathbb{N}} f^{-1}(U_n) = f^{-1} \left(\bigcup_{n \in \mathbb{N}} U_n \right)$$

nous donne la stabilité de $f^{-1}(\mathcal{U})$ par union dénombrable. □

12.2 Mesures positives.

12.2.1 Définitions.

Définition 12.2.1 On appelle *espace mesurable* un couple (E, \mathcal{T}) , où E est un ensemble et \mathcal{T} une tribu sur E . Les éléments de la tribu \mathcal{T} sont appelés *ensembles mesurables*.

Définition 12.2.2 Soit (E, \mathcal{T}) un espace mesurable. On appelle *mesure positive* sur (E, \mathcal{T}) une application $\mu : \mathcal{T} \rightarrow [0, +\infty]$ qui satisfait les deux conditions suivantes :

1. $\mu(\emptyset) = 0$,
2. pour toute suite $(A_n)_{n \in \mathbb{N}^*}$ d'ensembles de \mathcal{T} deux à deux disjoints, on a l'égalité :

$$\mu \left(\bigcup_{n \in \mathbb{N}^*} A_n \right) = \sum_{n \in \mathbb{N}^*} \mu(A_n)$$

Remarques.

1. La 2^e condition intervenant dans cette définition s'appelle *σ -additivité*.
2. Pour tout ensemble mesurable $A \in \mathcal{T}$, on appelle $\mu(A)$ la mesure de A .
3. Il est fréquent de dire simplement mesure au lieu de mesure positive.

Notation. Pour une suite $(A_n)_{n \in \mathbb{N}^*}$ d'ensembles dont nous aurons préalablement vérifié qu'ils sont 2 à 2 disjoints, nous emploierons la notation :

$$\bigsqcup_{n \in \mathbb{N}^*} A_n = \bigcup_{n \in \mathbb{N}^*} A_n$$

De la sorte, la propriété de σ -additivité d'une mesure se réécrit :

$$\mu \left(\bigsqcup_{n \in \mathbb{N}^*} A_n \right) = \sum_{n \in \mathbb{N}^*} \mu(A_n)$$

Définition 12.2.3 On appelle *espace mesuré* un triplet (E, \mathcal{T}, μ) , où μ est une mesure sur l'espace mesurable (E, \mathcal{T}) .

Exemples.

1. Soit E un ensemble. Sur l'espace mesurable $(E, \mathcal{P}(E))$, nous définissons une mesure μ , dite de comptage, en convenant que pour toute partie A de E , $\mu(A)$ vaut le cardinal de A (noté dans la suite $\#A$). Le lecteur vérifiera la σ -additivité en distinguant les deux cas où le sous-ensemble $\sum_{n \in \mathbb{N}^*} A_n$ est fini ou infini.
2. Soit E un ensemble non vide et $a \in E$. Sur l'espace mesurable $(E, \mathcal{P}(E))$, nous définissons la mesure de Dirac au point a , notée δ_a , comme suit :

$$\forall A \in \mathcal{P}(E) \quad \delta_a(A) = 1 \text{ si } a \in A \quad ; \quad \delta_a(A) = 0 \text{ si } a \notin A.$$

Le lecteur vérifiera facilement la propriété de σ -additivité en remarquant que, si les $(A_n)_{n \in \mathbb{N}^*}$ sont deux à deux disjoints, alors a est élément au plus d'un d'entre eux.

12.2.2 Propriétés d'une mesure.

Additivité.

Une conséquence immédiate de la définition est la propriété d'*additivité* d'une mesure : si A_1, \dots, A_n sont des ensembles mesurables 2 à 2 disjoints, alors

$$\mu(A_1 \cup \dots \cup A_n) = \mu(A_1) + \dots + \mu(A_n).$$

Il suffit en effet d'appliquer la définition 12.2.2 en prenant $A_k = \emptyset$ pour $k \geq n+1$.

Cette propriété d'additivité nous donne l'égalité suivante :

$$\forall (A, B) \in \mathcal{T}^2 \quad \mu(A) = \mu(A - B) + \mu(A \cap B) \quad (12.1)$$

En effet, l'ensemble A est égal à la réunion des deux ensembles disjoints $A - B$ et $A \cap B$.

En particulier, si $B \subset A$, nous obtenons $\mu(A) = \mu(A - B) + \mu(B)$, d'où la propriété suivante concernant une différence propre :

$$B \subset A \text{ et } \mu(B) < +\infty \implies \mu(A - B) = \mu(A) - \mu(B)$$

Comme l'ensemble $A \cup B$ est réunion des ensembles disjoints B et $A - B$, nous obtenons en utilisant (12.1) :

$$\forall (A, B) \in \mathcal{T}^2 \quad \mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B).$$

Si $\mu(A \cap B) < +\infty$, nous pouvons également écrire :

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B).$$

Croissance.

$$\forall (A, B) \in \mathcal{T}^2 \quad A \subset B \implies \mu(A) \leq \mu(B)$$

En effet, on a l'égalité : $\mu(B) = \mu(A) + \mu(B - A)$.

Sous- σ -additivité.

$$\forall (A_n) \in \mathcal{T}^{\mathbb{N}^*} \quad \mu \left(\bigcup_{n \in \mathbb{N}^*} A_n \right) \leq \sum_{n \in \mathbb{N}^*} \mu(A_n) \quad (12.2)$$

Démonstration: Posons $U_0 = \emptyset$ et, pour $n \in \mathbb{N}^*$, $U_n = A_1 \cup \dots \cup A_n$ et $B_n = U_n - U_{n-1} = A_n - U_{n-1}$. Par construction, les ensembles B_n sont deux à deux disjoints, on a $B_n \subset A_n$ pour tout $n \in \mathbb{N}^*$ et :

$$\bigcup_{n \in \mathbb{N}^*} B_n = \bigcup_{n \in \mathbb{N}^*} A_n.$$

Nous en déduisons :

$$\mu \left(\bigcup_{n \in \mathbb{N}^*} A_n \right) = \mu \left(\bigcup_{n \in \mathbb{N}^*} B_n \right) = \sum_{n \in \mathbb{N}^*} \mu(B_n) \leq \sum_{n \in \mathbb{N}^*} \mu(A_n)$$

□

Passage à la limite croissante.

Si $(A_n) \in \mathcal{T}^{\mathbb{N}^*}$ est une suite croissante au sens de l'inclusion, c'est-à-dire telle que $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}^*$, nous convenons d'adopter la notation suivante :

$$\bigcup_{n \in \mathbb{N}^*} A_n = \lim \uparrow A_n$$

Pour toute suite $(A_n) \in \mathcal{T}^{\mathbb{N}^*}$ croissante au sens de l'inclusion, nous avons alors :

$$\mu(\lim \uparrow A_n) = \lim \uparrow \mu(A_n)$$

Démonstration: Si nous introduisons les mêmes ensembles que dans la démonstration précédente, la croissance de la suite $(A_n) \in \mathcal{T}^{\mathbb{N}^*}$ nous donne ici : $U_n = A_n = B_1 \cup \dots \cup B_n$, d'où $\mu(A_n) = \mu(B_1) + \dots + \mu(B_n)$. Toujours en procédant comme dans la démonstration précédente, nous en déduisons :

$$\mu\left(\bigcup_{n \in \mathbb{N}^*} A_n\right) = \mu\left(\bigcup_{n \in \mathbb{N}^*} B_n\right) = \sum_{n \in \mathbb{N}^*} \mu(B_n) = \lim \uparrow \mu(A_n).$$

□

Passage à la limite décroissante.

Si $(A_n) \in \mathcal{T}^{\mathbb{N}^*}$ est une suite décroissante au sens de l'inclusion, c'est-à-dire telle que $A_{n+1} \subset A_n$ pour tout $n \in \mathbb{N}^*$, nous convenons d'adopter la notation suivante :

$$\bigcap_{n \in \mathbb{N}^*} A_n = \lim \downarrow A_n$$

Pour toute suite $(A_n) \in \mathcal{T}^{\mathbb{N}^*}$ décroissante au sens de l'inclusion, nous avons alors l'implication suivante :

$$\mu(A_1) < +\infty \implies \mu(\lim \downarrow A_n) = \lim \downarrow \mu(A_n)$$

Démonstration: Pour tout $n \in \mathbb{N}^*$, posons $C_n = A_1 - A_n$; puisque $A_n \subset A_1$, nous avons :

$$\mu(A_1) = \mu(C_n) + \mu(A_n).$$

Puisque $\mu(A_1) < +\infty$, nous en déduisons que $\mu(C_n) < +\infty$ et $\mu(A_n) < +\infty$, ainsi que l'égalité :

$$\mu(C_n) = \mu(A_1) - \mu(A_n) \tag{12.3}$$

La suite $(C_n)_{n \in \mathbb{N}^*}$ est croissante et nous avons l'égalité ensembliste :

$$\bigcup_{n \in \mathbb{N}^*} C_n = A_1 - \lim \downarrow A_n$$

Par passage à la limite croissante, puis en utilisant (12.3) , nous obtenons donc :

$$\mu(A_1 - \lim \downarrow A_n) = \lim \uparrow \mu(C_n) = \mu(A_1) - \lim \downarrow \mu(A_n).$$

Il nous reste à écrire, en utilisant de nouveau $\mu(A_1) < +\infty$:

$$\mu(\lim \downarrow A_n) = \mu(A_1) - \mu(A_1 - \lim \downarrow A_n)$$

pour obtenir la conclusion voulue.

□

Remarques.

1. Il est facile de constater que la condition $\mu(A_1) < +\infty$ peut être remplacée dans l'énoncé précédent par :

$$\exists n \in \mathbb{N}^* \quad \mu(A_n) < +\infty$$

2. En revanche, si tous les A_n sont de mesure infinie, la propriété précédente peut tomber en défaut, comme le montre le contre-exemple suivant.

Si nous notons μ la mesure de comptage sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ et si nous prenons $A_n = \{k \in \mathbb{N}, k \geq n\}$, nous avons $\lim \downarrow A_n = \emptyset$ d'où $\mu(\lim \downarrow A_n) = 0$ alors que $\lim \downarrow \mu(A_n) = +\infty$.

12.3 Application à la modélisation du hasard.

L'objet de la théorie des Probabilités est de construire un modèle permettant d'étudier les phénomènes dépendant du hasard. Nous allons procéder en deux étapes en présentant d'abord un modèle simple, mais dont le champ d'application sera limité, puis en le raffinant pour aboutir à notre modèle définitif.

12.3.1 Un modèle simple : le modèle additif.

L'ensemble des résultats possibles.

Qu'est-ce que le hasard ? Comme nous enseignons les mathématiques et non la philosophie, nous allons adopter une définition empirique, c'est-à-dire basée sur l'expérience.

Nous dirons qu'une expérience (scientifique ou autre, de la vie courante par exemple) dépend du hasard et nous l'appellerons *expérience aléatoire* – du mot latin *alea* qui signifie jet de dés – si elle produit un résultat que l'on ne connaît pas à l'avance mais qui en revanche appartient à un ensemble connu à l'avance. Cet ensemble, appelé *ensemble des résultats possibles* est noté traditionnellement Ω . Il peut être plus ou moins grand suivant l'expérience considérée.

Exemples :

1. Jeu de pile ou face. $\Omega = \{p, f\}$.
2. Jet de dé à 6 faces. $\Omega = \{1, 2, \dots, 6\}$.

3. On joue à pile ou face de façon répétée et l'on s'intéresse au rang d'apparition du premier pile. $\Omega = \mathbb{N}^*$.
4. Durée de vie d'une batterie de voiture. $\Omega = \mathbb{R}_+$ ou $\Omega = [0, T]$ avec T suffisamment grand.
5. Courbe d'évolution de la température sur une journée dans une station météo. $\Omega = C^0([0, T], \mathbb{R})$.

Dans les exemples 1 à 3, Ω est fini ou dénombrable : c'est le cas *discret*. Dans les exemples 4 et 5, Ω est infini non dénombrable : c'est le cas *continu*.

L'ensemble Ω est le premier ingrédient du modèle mathématique que nous allons construire pour représenter le mieux possible notre expérience aléatoire. Nous allons mettre en oeuvre deux ingrédients supplémentaires pour construire notre modèle additif.

L'ensemble des événements.

Une fois faite notre expérience aléatoire, nous nous trouvons devant un certain résultat $\omega \in \Omega$. À part peut-être le cas où Ω est très petit, il arrive très souvent que ce ne soit pas le résultat précis ω qui nous intéresse mais que l'on cherche plutôt à répondre à la question : "est-ce que ω appartient à tel sous-ensemble donné de Ω ?"

Exemple. Durée de vie d'un ordinateur.

Si le constructeur vous garantit la machine pendant trois ans, il a bien sûr d'abord étudié la question : « quelles sont les chances pour que $\omega \leq 3$? » De la réponse à cette question dépendra le prix auquel il va vous facturer la garantie. Dans cet exemple, la question intéressante s'écrit donc : « $\omega \in A$? », avec $A = [3, +\infty[\subset \Omega = \mathbb{R}_+$.

Nous appellerons *événement* un sous-ensemble A de Ω tel qu'un observateur de l'expérience est capable de répondre à la question " $\omega \in A$?"

Exemples. Ω est l'événement certain. \emptyset est l'événement impossible.

Dès que nous allons manipuler des événements, certains opérateurs ensemblistes vont apparaître :

- L'événement contraire de A est son complémentaire $A^c = \Omega - A$.
- L'événement "le résultat est dans A_1 ou A_2 " s'écrit $A_1 \cup A_2$
- L'événement "le résultat est dans A_1 et A_2 " s'écrit $A_1 \cap A_2$

Nous travaillerons donc sur des intersections ou des unions finies d'événements, ainsi que sur des complémentaires d'événements.

Un observateur de l'expérience étant donné, il est donc naturel de supposer que l'ensemble des parties de Ω dont il peut discerner la réalisation ou la non-réalisation après l'expérience a une structure de clan. Nous noterons dans la suite \mathcal{C} le clan des événements discernables par notre observateur.

Remarque. Plusieurs observateurs distincts peuvent être associés à une même expérience. Par exemple, si le résultat de l'expérience aléatoire qui nous intéresse est le temps (mesuré en secondes) mis par une athlète pour courir un 100 mètres (prenons $\Omega = [0; 15]$) et si un premier arbitre n'est muni que d'une montre à trotteuse tandis que le second se réfère à un dispositif électronique, nous avons $\mathcal{C}_1 \subsetneq \mathcal{C}_2$. Ainsi, l'événement $A = [0; 9, 73]$ qui se traduit en langage courant par "elle a battu le record du monde féminin du 9 septembre 2007" vérifie $A \in \mathcal{C}_2$ mais $A \notin \mathcal{C}_1$.

Probabilités.

Nous avons tous l'intuition que certains événements ont plus de chances de se produire que d'autres. Par exemple, dans une loterie, il est plus probable de tirer un billet perdant que de tirer le gros lot.

Pour préciser cette intuition, nous souhaitons associer à un événement donné un nombre réel qui mesure les chances qu'il se produise; par exemple, si l'événement se produit avec une chance sur deux (pile ou face avec une pièce équilibrée), nous lui associerons le nombre $1/2$.

Nous appellerons donc probabilité une application $P : \mathcal{C} \rightarrow [0, 1]$ vérifiant certaines conditions que nous allons préciser maintenant.

Tout d'abord, Ω étant l'événement certain, nous demanderons $P(\Omega) = 1$ (en langage courant, il y a "100 chances sur 100" qu'il se produise).

Ensuite, il est naturel d'exiger la propriété d'*additivité* :

$$A \in \mathcal{C}, B \in \mathcal{C}, A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B).$$

Notons que la condition $A \cap B = \emptyset$ – les deux événements A et B sont *disjoints* – est indispensable pour ne pas compter deux fois les mêmes résultats possibles.

Une des raisons pour lesquelles il est naturel de faire cette hypothèse d'additivité d'une probabilité s'appelle :

La loi empirique des grands nombres

Si nous pouvons répéter une même expérience un grand nombre de fois dans des conditions identiques et que nous notions $\varphi_n(A)$ le nombre de fois où le résultat a appartenu à l'événement A au cours des n premières répétitions de l'expérience, nous nous attendons à ce que le rapport $\varphi_n(A)/n$ (fréquence de réalisation de l'événement A au cours des n premières répétitions de l'expérience) tende vers une certaine limite lorsque n augmente.

Intuitivement, c'est cette limite que nous aurions envie d'appeler "probabilité que l'événement A se réalise". Si nous admettons qu'une telle limite existe pour tout événement A et que nous la notions $P(A)$, il est facile de constater que nous avons ainsi défini une application $P : \mathcal{C} \rightarrow [0, 1]$, additive et telle que $P(\Omega) = 1$.

Le modèle additif

Nous pouvons donc modéliser une expérience aléatoire par un triplet (Ω, \mathcal{C}, P) , où Ω est un ensemble, \mathcal{C} un clan sur Ω et P une application additive de \mathcal{C} dans $[0, 1]$ telle que $P(\Omega) = 1$.

Ce modèle s'avère en général suffisant lorsque l'expérience aléatoire considérée ne mène qu'à un nombre fini de résultats possibles, c'est-à-dire lorsque le cardinal de Ω vérifie : $\#\Omega < +\infty$.

Lorsque Ω est fini, nous prendrons pratiquement toujours $\mathcal{C} = \mathcal{P}(\Omega)$, qui est le seul clan sur Ω contenant tous les singletons : ce clan d'événements correspond à un observateur qui discerne le résultat précis ω de l'expérience.

Nous allons maintenant considérer un cas particulier qui a de nombreuses applications pratiques.

Exemple. Probabilité uniforme sur un ensemble Ω fini.

Nous nous plaçons dans le cas où $\#\Omega < +\infty$ et nous prenons pour ensemble des événements $\mathcal{C} = \mathcal{P}(\Omega)$.

Il existe une unique probabilité sur (Ω, \mathcal{C}) qui attribue les mêmes chances à tous les résultats possibles (on dit alors qu'on est en situation d'équiprobabilité). Elle est appelée *probabilité uniforme* sur Ω et est définie par :

$$\forall \omega \in \Omega, \quad P(\{\omega\}) = \frac{1}{\#\Omega}.$$

Pour tout événement $A \subset \Omega$, on a alors :

$$P(A) = \frac{\#A}{\#\Omega} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}.$$

Nous terminons ce sous-paragraphe par une proposition caractérisant les probabilités sur $(\Omega, \mathcal{P}(\Omega))$ lorsque l'ensemble des résultats possibles Ω est fini.

Proposition 12.3.1 *Soit $n \in \mathbb{N}^*$ et $\Omega = \{\omega_1, \dots, \omega_n\}$ un ensemble fini. Il existe une bijection entre l'ensemble des probabilités P sur $(\Omega, \mathcal{P}(\Omega))$ et l'ensemble des n -uplets $(p_k)_{1 \leq k \leq n} \in (\mathbb{R}_+)^n$ tels que :*

$$\sum_{1 \leq k \leq n} p_k = 1.$$

Cette bijection est donnée par :

$$\forall k \in \{1, \dots, n\} \quad p_k = P(\{\omega_k\}).$$

Démonstration: Dans le sens direct, il est évident que si P est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$, alors le n -uplet $(p_k)_{1 \leq k \leq n}$ ainsi défini a les propriétés voulues.

Réciproquement, un tel n -uplet étant donné, s'il existe une probabilité P qui convienne, alors elle satisfait nécessairement la relation suivante par additivité :

$$\forall A \in \mathcal{P}(\Omega) \quad P(A) = \sum_{k/\omega_k \in A} p_k.$$

Il est facile de conclure en vérifiant que l'on a bien défini ainsi une probabilité sur $(\Omega, \mathcal{P}(\Omega))$. \square

Exemples.

1. Le cas le plus simple d'espace de probabilité fini non trivial est donné par $\#\Omega = 2$ et correspond à la modélisation du jeu de pile ou face (avec une pièce éventuellement biaisée). Il y a alors bijection entre les probabilités sur $(\Omega, \mathcal{P}(\Omega))$ et le segment $[0, 1]$: celle-ci est donnée par la probabilité $p \in [0, 1]$ que la pièce tombe sur pile. Formellement, on écrit en général $\Omega = \{0, 1\}$ (pile étant représenté par 1 et face par 0) et la probabilité définie par $P(\{1\}) = p$ et donc $P(\{0\}) = 1 - p$ est appelée probabilité de Bernoulli de paramètre $p \in [0, 1]$.

2. Pour tout $n \in \mathbb{N}^*$ et tout $p \in [0, 1]$, on définit sur $\Omega = \{0, 1, \dots, n\}$ (qui est donc de cardinal $n + 1$) la probabilité binomiale de paramètres n et p , notée $P = B_{n,p}$, par

$$\forall k \in \{0, 1, \dots, n\} \quad B_{n,p}(\{k\}) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Nous verrons plus loin dans quel type de situation la probabilité binomiale apparaît.

12.3.2 Modèle définitif : les espaces probabilisés.

Le modèle développé dans le paragraphe précédent s'applique bien au cas Ω fini, qui couvre un certain nombre de situations courantes : jeux de hasard, sondages d'opinion, contrôles de qualité etc. Néanmoins, même avec un jeu aussi simple que le pile ou face, on ne peut pas toujours se restreindre au cas $\#\Omega < +\infty$: si nous étudions par exemple le rang d'apparition du premier pile lors de jets successifs, on est amené à prendre $\Omega = \mathbb{N}^*$ puisqu'il est impossible de borner à l'avance ce rang par un entier fixé.

Lorsque le temps intervient, nous voyons apparaître naturellement des espaces encore plus "gros" (infinis non dénombrables) :

- durée de vie d'un composant électronique : $\Omega = \mathbb{R}_+$
- évolution du cours d'un actif financier : $\Omega = C^0([0, T], \mathbb{R})$
- déplacement d'une particule : $\Omega = C^0([0, T], \mathbb{R}^3)$

Le modèle du paragraphe précédent présente d'autres insuffisances : la structure de clan pour l'ensemble \mathcal{C} des événements et l'additivité d'une probabilité sont des hypothèses trop faibles pour nous permettre de mener effectivement des calculs lorsque Ω est infini, comme nous allons le voir sur le simple exemple suivant.

Si nous revenons au rang du premier pile lors de jets successifs d'une pièce équilibrée, une première question qui se pose est de savoir si ce rang est fini : peut-on obtenir face indéfiniment ?

Intuitivement, cela paraît extrêmement improbable et l'on s'attend à ce que la probabilité en soit nulle. Plus précisément, il est naturel de faire le raisonnement suivant : l'événement $A_n = \{ \text{on a obtenu face lors des } n \text{ premiers jets} \}$ est de probabilité $1/2^n$ et donc l'événement $A = \{ \text{on obtient face indéfiniment} \}$ a pour probabilité $\lim 1/2^n = 0$.

Pour rendre ce raisonnement rigoureux, il faudrait déjà que $A = \cap_n A_n$ soit effectivement un événement, ce qui n'est pas assuré sous la seule hypothèse : l'ensemble des événements est un clan. Nous avons besoin de supposer que

l'ensemble des événements est une tribu.

En outre, nous voudrions avoir : $A = \lim \downarrow A_n \Rightarrow P(A) = \lim \downarrow P(A_n)$. En passant aux événements complémentaires, cela revient à demander la propriété $P(\lim \uparrow B_n) = \lim \uparrow P(B_n)$. Or, pour une application additive comme P , cette propriété est équivalente à la σ -additivité : pour le voir, il suffit d'écrire, pour une suite (A_n) d'événements deux à deux disjoints

$$\bigcup_{n \in \mathbb{N}^*} A_n = \lim \uparrow \bigcup_{k=1}^n A_k.$$

Pour pouvoir mener nos calculs, nous allons donc imposer des contraintes supplémentaires sur les éléments constitutifs de notre modèle : l'ensemble des événements doit être une tribu (ou σ -algèbre), que nous noterons désormais plutôt \mathcal{A} , et P une application σ -additive. Nous en arrivons donc à la définition suivante, qui décrit le modèle que nous utiliserons désormais.

Définition 12.3.2 *Une probabilité sur un espace mesurable (Ω, \mathcal{A}) est une mesure positive sur cet espace telle que $P(\Omega) = 1$.*

Le triplet (Ω, \mathcal{A}, P) est alors appelé espace probabilisé (ou espace de probabilité).

En d'autres termes, un espace probabilisé est un espace mesuré de masse totale égale à 1.

Notons que dans le cas $\#\Omega < +\infty$, cette définition est équivalente à celle du paragraphe précédent. En effet, dans ce cas, $\mathcal{A} \subset \mathcal{P}(\Omega)$ est fini si bien que \mathcal{A} algèbre et \mathcal{A} σ -algèbre sont équivalents, de même que P additive et P σ -additive. Comme nous n'utiliserons le modèle additif que dans le cas $\#\Omega < +\infty$, nous le considérerons comme un cas particulier de notre modèle général.

Contrairement à ce qui se passait dans le cas $\#\Omega < +\infty$, dans le cas général, nous serons amenés à considérer des tribus d'événements $\mathcal{A} \subsetneq \mathcal{P}(\Omega)$. En effet, plus la tribu \mathcal{A} est grande, plus il est difficile de construire une mesure de probabilité sur celle-ci. Lorsque Ω est infini non dénombrable, $\mathcal{P}(\Omega)$ est en général trop grand pour que l'on puisse construire sur cette tribu une mesure de probabilité qui mène à un modèle satisfaisant. En fait, on a souvent recours à une politique minimaliste en prenant pour \mathcal{A} la plus petite tribu (c'est-à-dire la tribu engendrée) qui contienne les événements que nous voulons pouvoir observer.

Dans le cas où Ω est infini dénombrable, nous avons la caractérisation suivante d'une probabilité sur $(\Omega, \mathcal{P}(\Omega))$, que nous admettrons.

Proposition 12.3.3 Soit $\Omega = \{\omega_n, n \in \mathbb{N}\}$ un ensemble infini dénombrable. Il existe une bijection entre l'ensemble des probabilités P sur $(\Omega, \mathcal{P}(\Omega))$ et l'ensemble des suites $(p_n)_{n \in \mathbb{N}} \in (\mathbb{R}_+)^{\mathbb{N}}$ telles que :

$$\sum_{n \in \mathbb{N}} p_n = 1.$$

Cette bijection est donnée par :

$$\forall n \in \mathbb{N} \quad p_n = P(\{\omega_n\}).$$

Par σ -additivité, nous avons alors la relation :

$$\forall A \in \mathcal{P}(\Omega) \quad P(A) = \sum_{k/\omega_k \in A} p_k.$$

Exemples. Sur l'espace mesurable $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, on définit :

1. pour tout $p \in]0, 1[$, la probabilité géométrique de paramètre p par

$$\forall n \in \mathbb{N} \quad P(\{n\}) = (1 - p) p^n.$$

2. pour tout $\lambda > 0$, la probabilité de Poisson de paramètre λ par

$$\forall n \in \mathbb{N} \quad P(\{n\}) = e^{-\lambda} \frac{\lambda^n}{n!}.$$

□

Le choix d'une probabilité P sur (Ω, \mathcal{A}) qui “colle au plus près” à la réalité d'une expérience aléatoire est une question suffisamment vaste pour faire l'objet d'une branche entière des mathématiques : la statistique.

Le fait qu'un même espace probabilisable (Ω, \mathcal{A}) puisse être muni de probabilités différentes correspondant à des “hasards différents” est illustré par le *paradoxe de Bertrand*, que vous trouverez en ligne par exemple sur Wikipédia.

12.4 Probabilités conditionnelles. Indépendance d'événements.

12.4.1 Définition des probabilités conditionnelles.

La notion de probabilité conditionnelle apparaît naturellement lorsqu'on possède une information partielle sur le résultat d'une expérience aléatoire.

Exemple : À la Foire du Trône, un jeu de loterie est constitué d'une urne contenant 80 billets perdants et 20 billets gagnants, dont 2 seulement donnent droit au gros lot.

Je tire un billet « au hasard » (cette expression sous-entend que l'on est en situation d'équiprobabilité) ; avec quelle probabilité vais-je gagner le gros lot ? Nous pouvons représenter cette expérience aléatoire par le modèle suivant :

- $\Omega = \{b_1, \dots, b_{100}\}$ où b_1 et b_2 représentent les billets qui donnent droit au gros lot, b_3, \dots, b_{20} représentent des billets gagnants ordinaires et b_{21}, \dots, b_{100} des billets perdants ;
- $\mathcal{A} = \mathcal{P}(\Omega)$;
- la probabilité P est uniforme : $\forall i = 1, 2, \dots, 100 \quad P(\{b_i\}) = 1/100$.

L'événement qui nous intéresse est $A = \{b_1, b_2\}$ et nous obtenons donc immédiatement : $P(A) = 2/100$.

Imaginons maintenant que juste après mon tirage et avant que je ne découvre mon billet, le forain — qui a un signe de reconnaissance secret — m'affirme : « Vous avez tiré un billet gagnant ! ». Sachant cela, je ne vais plus calculer ma probabilité de la même façon.

Muni de ce renseignement, je sais que le résultat de l'expérience appartient à l'ensemble $E = \{b_1, \dots, b_{20}\}$. Comme il y a symétrie entre tous les billets, ils sont encore équiprobables, d'où le nouveau calcul de ma probabilité :

$$\frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{2}{20} = \frac{1}{10}.$$

Nous constatons que la valeur de la probabilité n'est plus la même car, disposant d'une information supplémentaire, nous avons en fait modifié le calcul comme suit :

$$P'(A) = \frac{\#(A \cap E)}{\#E} = \frac{P(A \cap E)}{P(E)}.$$

P' est appelée probabilité conditionnelle sachant E (i.e. sachant mon information partielle). De façon plus générale, nous avons la proposition suivante :

Proposition 12.4.1 Soit (Ω, \mathcal{A}, P) un espace de probabilité et $E \in \mathcal{A}$ un événement tel que $P(E) > 0$. L'application $P^E : \mathcal{A} \rightarrow [0, 1]$ définie par :

$$\forall A \in \mathcal{A}, \quad P^E(A) = \frac{P(A \cap E)}{P(E)},$$

est une probabilité. On l'appelle probabilité conditionnelle à l'événement E ou probabilité sachant E . On note sa valeur sur un événement A comme suit :

$$P(A|E) = P^E(A) = \frac{P(A \cap E)}{P(E)}.$$

Démonstration : C'est bien une application de \mathcal{A} dans $[0, 1]$ puisqu'on a :

$$A \cap E \subset E \Rightarrow 0 \leq P(A \cap E) \leq P(E).$$

De plus, on vérifie immédiatement : $P^E(\emptyset) = 0$ et $P^E(\Omega) = 1$.

Il nous reste à montrer que l'application P^E est σ -additive, ce qui résulte du calcul suivant, valable pour toute suite $(A_n) \in \mathcal{A}^{\mathbb{N}}$ constituée d'événements 2 à 2 disjoints :

$$\frac{P\left(\left(\biguplus_n A_n\right) \cap E\right)}{P(E)} = \frac{P\left(\biguplus_n (A_n \cap E)\right)}{P(E)} = \frac{\sum_n P(A_n \cap E)}{P(E)} = \sum_n \frac{P(A_n \cap E)}{P(E)} \quad \square$$

Remarque : Dans cette définition de la probabilité conditionnelle, l'idée est que, étant donnée l'information partielle E , l'événement intéressant n'est plus l'événement A de départ mais plutôt $A \cap E$. L'application $A \mapsto P(A \cap E)$ est bien positive et σ -additive mais n'a pas la bonne masse totale (i.e. la valeur 1 sur Ω). Pour en faire une probabilité, il suffit de la renormaliser en divisant par $P(E)$ mais ceci n'est possible que si $P(E) > 0$.

Exemple : On change les règles de la loterie : il y a toujours 20 billets gagnants sur 100 mais pour obtenir le gros lot, il faut d'abord tirer un billet gagnant puis il faut faire tourner une roue divisée en 5 parties égales dont une seule donne droit au gros lot. Quelle est la probabilité d'obtenir celui-ci ? Écrire le modèle serait un peu plus long : $\Omega = \{b_1, \dots, b_{100}\} \times \{r_1, \dots, r_5\}$ etc. Pour gagner du temps, passons directement au calcul, en écrivant les événements informellement :

$$A = \{\text{le joueur gagne le gros lot}\}; E = \{\text{le joueur tire un billet gagnant}\}.$$

Nous pouvons alors mener le calcul de la façon suivante :

$$P(A) = P(A \cap E) = P(A|E)P(E) = \frac{1}{5} \times \frac{20}{100} = \frac{4}{100}.$$

Nous venons d'utiliser la formule $P(A \cap E) = P(A|E)P(E)$; plus généralement nous avons la proposition suivante.

Proposition 12.4.2 (Formule des probabilités composées) *Considérons $(A_1, \dots, A_n) \in \mathcal{A}^n$ avec $P(A_1 \cap \dots \cap A_{n-1}) > 0$. Alors, on a l'égalité :*

$$P(A_1 \cap \dots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \dots \cap A_{n-1})$$

Démonstration: Le membre de droite s'écrit par définition d'une probabilité conditionnelle :

$$P(A_1) \times \frac{P(A_1 \cap A_2)}{P(A_1)} \times \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \times \cdots \times \frac{P(A_1 \cap \cdots \cap A_n)}{P(A_1 \cap \cdots \cap A_{n-1})},$$

et l'on conclut par simplification. \square

12.4.2 Formule de Bayes.

Proposition 12.4.3 (Formule des probabilités totales) *Si E_1, \dots, E_n sont des événements tels que $P(E_i) > 0$ pour tout $1 \leq i \leq n$ et formant une partition de l'ensemble Ω (i.e. $\biguplus_{1 \leq i \leq n} E_i = \Omega$), alors on a pour tout événement $A \in \mathcal{A}$ l'égalité suivante :*

$$P(A) = \sum_{i=1}^n P(A|E_i)P(E_i).$$

Démonstration: Ceci résulte des égalités successives :

$$\begin{aligned} P(A) &= P\left(A \cap \left(\biguplus_{1 \leq i \leq n} E_i\right)\right) = P\left(\biguplus_{1 \leq i \leq n} (A \cap E_i)\right) \\ &= \sum_{1 \leq i \leq n} P(A \cap E_i) = \sum_{1 \leq i \leq n} P(A|E_i)P(E_i) \end{aligned}$$

\square

Notons en particulier que la formule des probabilités totales se réécrit dans le cas $n = 2$:

Soient deux événements E et A avec $0 < P(E) < 1$. Alors, on a l'égalité :

$$P(A) = P(A|E)P(E) + P(A|E^c)P(E^c)$$

Nous déduisons de la formule des probabilités totales le théorème suivant :

Théorème 12.4.4 (Formule de Bayes) *Si E_1, \dots, E_n sont des événements de probabilités strictement positives, formant une partition de l'ensemble Ω et si A est un événement tel que $P(A) > 0$, alors on a l'égalité :*

$$P(E_1|A) = \frac{P(A|E_1)P(E_1)}{\sum_{i=1}^n P(A|E_i)P(E_i)}.$$

Démonstration : Il suffit d'écrire les égalités :

$$P(E_1|A) = \frac{P(A \cap E_1)}{P(A)} = \frac{P(A|E_1)P(E_1)}{P(A)}$$

puis d'appliquer la formule des probabilités totales au dénominateur. \square

Remarque : Revenons au cas particulier $n = 2$. La formule de Bayes s'écrit alors :

$$P(E|A) = \frac{P(A|E)P(E)}{P(A|E)P(E) + P(A|E^c)P(E^c)}$$

Notons qu'il n'est donc pas possible de calculer $P(E|A)$ à partir de la seule donnée de $P(A|E)$; il faut connaître d'autres quantités pour faire ce calcul.

Exercice : Une maladie rare touche 1 personne sur 10000 dans la population française. Quand cette maladie est présente, un test sanguin permet de la détecter dans 99% des cas. En revanche, ce test produit des faux positifs dans 1 cas sur 1000.

Le test d'une personnes est positif. Quelle est la probabilité qu'elle soit vraiment atteinte de la maladie ? Que pensez-vous de la qualité de ce test sanguin ?

Solution : Définissons les événements :

$$E = \{\text{la personne est atteinte de la maladie}\}; A = \{\text{le test est positif}\}$$

L'énoncé nous fournit les données suivantes :

$$P(E) = 10^{-4}, P(A|E) = 0,99, P(A|E^c) = 10^{-3}$$

La formule de Bayes nous permet alors de calculer :

$$P(E|A) = \frac{0,99 \times 10^{-4}}{0,99 \times 10^{-4} + 10^{-3} \times 0,9999} \sim \frac{1}{11}$$

Conclusion : À moins de vouloir provoquer beaucoup de panique inutile, le test est bon pour la poubelle !

12.4.3 Indépendance d'événements.

Définissons les événements $A = \{\text{au moins un ascenseur du bâtiment dans lequel nous nous trouvons sera en panne demain}\}$ et $B = \{\text{l'indice CAC 40 va s'effondrer demain}\}$. Intuitivement, la réalisation (ou non-réalisation) de l'événement A n'a aucune influence sur les chances que l'événement B se produise, ce qui s'écrit $P(A|B) = P(A)$ ou encore $P(A \cap B) = P(A)P(B)$. De tels événements seront dits *indépendants*.

Définition 12.4.5 Deux événements A et B sont dits indépendants s'ils vérifient l'égalité :

$$P(A \cap B) = P(A)P(B)$$

La notion d'indépendance est essentielle en calcul des probabilités ; elle fera souvent partie de nos hypothèses en modélisation. Par exemple, pour étudier la durée de vie d'un circuit électronique constitué de plusieurs composants élémentaires, nous supposons souvent que ces différents composants tombent en panne indépendamment les uns des autres.

Remarque. Il est facile de vérifier que si le couple (A, B) est constitué d'événements indépendants, alors il en est de même pour les couples suivants :

$$(A, B^c), (A^c, B), (A^c, B^c).$$

Exemple. On tire une carte au hasard dans un jeu de 52 et l'on considère les événements $A = \{\text{c'est un roi}\}$ et $B = \{\text{c'est un coeur}\}$. Nous calculons facilement :

$$P(A \cap B) = \frac{1}{52}, P(A) = \frac{4}{52} = \frac{1}{13}, P(B) = \frac{13}{52} = \frac{1}{4}.$$

Nous en déduisons que les événements A et B sont indépendants, conformément à la définition précédente.

Plus généralement, dans un tel tirage, la figure et la couleur de la carte sont indépendantes.

Définition 12.4.6 Soient $(A_1, \dots, A_n) \in \mathcal{A}^n$ des événements.

Nous dirons que ces événements sont 2 à 2 indépendants si pour tout $(i, j) \in \{1, \dots, n\}$ avec $i \neq j$, on a $P(A_i \cap A_j) = P(A_i)P(A_j)$.

Nous dirons que ces événements sont mutuellement indépendants (ou encore indépendants dans leur ensemble) si pour toute sous-famille d'indices $\{i_1, \dots, i_k\} \subset \{1, \dots, n\}$, on a :

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \times P(A_{i_2}) \times \dots \times P(A_{i_k}).$$

Remarque. Des événements mutuellement indépendants sont clairement indépendants 2 à 2 mais la réciproque est fausse, comme nous allons le voir dans le contre-exemple ci-dessous. Autrement dit, la propriété de mutuelle indépendance est plus forte que la propriété d'indépendance 2 à 2.

Lorsqu'un énoncé dit que des événements A_1, \dots, A_n sont indépendants sans autre précision, c'est toujours l'indépendance mutuelle, c'est-à-dire la plus forte, qu'il faut comprendre.

Contre-exemple: On jette une pièce équilibrée 2 fois de suite et l'on considère les événements : $A = \{\text{on obtient pile au premier jet}\}$, $B = \{\text{on obtient pile au second jet}\}$, $C = \{\text{les résultats des deux jets sont différents}\}$. Il est facile de calculer :

$$P(A) = P(B) = P(C) = 1/2$$

et

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = 1/4.$$

Nous en déduisons que les événements A , B et C sont indépendants 2 à 2. En revanche, ces événements ne sont pas mutuellement indépendants puisque :

$$P(A \cap B \cap C) = P(\emptyset) = 0 \neq P(A)P(B)P(C).$$

□

Proposition 12.4.7 *Des événements $(A_1, \dots, A_n) \in \mathcal{A}^n$ sont mutuellement indépendants si et seulement si on a*

$$P\left(\bigcap_{i=1}^n C_i\right) = \prod_{i=1}^n P(C_i),$$

pour tous $C_i \in \{\Omega, A_i, A_i^c, \emptyset\}$, $1 \leq i \leq n$.

Le sens direct est établi dans [OUV¹ 56-57]. Le sens réciproque est laissé à titre d'exercice.

Nous allons maintenant mettre en évidence une situation générale dans laquelle la *probabilité binomiale* apparaît naturellement.

Proposition 12.4.8 *Soient A_i , $1 \leq i \leq n$, des événements indépendants ayant tous la même probabilité $p \in [0, 1]$. Pour tout $0 \leq k \leq n$, la probabilité que k exactement d'entre eux soient réalisés est égale à :*

$$B_{n,p}(\{k\}) = \binom{n}{k} p^k (1-p)^{n-k}$$

Démonstration: Notons Θ l'ensemble des parties de $\{1, \dots, n\}$ qui sont de cardinal k , de sorte que $\#\Theta = \binom{n}{k}$.

L'événement dont nous voulons calculer la probabilité est le suivant :

$$\bigcup_{\theta \in \Theta} \left[\left(\bigcap_{i \in \theta} A_i \right) \cap \left(\bigcap_{i \notin \theta} A_i^c \right) \right]$$

Un instant de réflexion montre qu'il s'agit d'une union d'événements deux à deux disjoints, si bien que la probabilité cherchée s'écrit :

$$\sum_{\theta \in \Theta} P \left[\left(\bigcap_{i \in \theta} A_i \right) \cap \left(\bigcap_{i \notin \theta} A_i^c \right) \right]$$

À cause de l'hypothèse d'indépendance des $A_i, 1 \leq i \leq n$, cette probabilité s'écrit encore :

$$\sum_{\theta \in \Theta} p^k (1-p)^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

□

Après avoir défini l'indépendance d'une famille finie d'événements, nous terminons ce paragraphe en passant au cas d'une famille infinie dénombrable.

Définition 12.4.9 Une suite d'événements $(A_n)_{n \in \mathbb{N}}$ est dite indépendante si pour toute sous-famille finie d'indices $\{i_1, \dots, i_k\} \subset \mathbb{N}$ (avec $k \in \mathbb{N}^*$), on a l'égalité :

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \times P(A_{i_2}) \times \dots \times P(A_{i_k}).$$

12.5 Exercices.

12.5.1 Clans.

Indicatrices

Soit A un clan sur Ω ; on définit la fonction $\mathbf{1}_A : \Omega \rightarrow \{0, 1\}$ par :

$\mathbf{1}_A(\omega) = 0$ si $\omega \notin A$, $\mathbf{1}_A(\omega) = 1$ si $\omega \in A$.

Montrer les égalités suivantes : $\mathbf{1}_{A \cap B} = \inf(\mathbf{1}_A, \mathbf{1}_B) = \mathbf{1}_A \mathbf{1}_B$,

$\mathbf{1}_{A \cup B} = \sup(\mathbf{1}_A, \mathbf{1}_B) = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A \mathbf{1}_B$,

$\mathbf{1}_{A^c} = 1 - \mathbf{1}_A$,

ainsi que l'équivalence : $\forall \omega \in \Omega, \mathbf{1}_A(\omega) \leq \mathbf{1}_B(\omega) \Leftrightarrow A \subset B$.

On rappelle que la différence symétrique est définie par :

$$A \Delta B = (A \cup B) \setminus (A \cap B) = (A^c \cap B) \cup (A \cap B^c).$$

Montrer que :

$$\mathbf{1}_{A \Delta B} = |\mathbf{1}_A - \mathbf{1}_B| = \mathbf{1}_A + \mathbf{1}_B \pmod{2} = \mathbf{1}_A + \mathbf{1}_B - 2 \mathbf{1}_A \mathbf{1}_B.$$

En déduire que la différence symétrique est commutative, associative et que l'intersection est distributive par rapport à la différence symétrique.

Clan engendré par une partition.

Soit (P_1, \dots, P_n) une partition de Ω . Décrire le clan \mathcal{C} qu'elle engendre.

Réunions d'algèbres.

Montrer que la réunion d'une suite croissante d'algèbres est une algèbre. Est-ce que la réunion d'une suite quelconque d'algèbres est une algèbre ?

***Description de l'algèbre engendrée.**

Soit \mathcal{E} une famille de parties d'un ensemble Ω . Montrer que l'algèbre engendrée par \mathcal{E} est égale à l'ensemble des réunions finies d'intersections finies d'éléments de \mathcal{E} ou de complémentaires d'éléments de \mathcal{E} (parties de la forme $\bigcup_{i=1}^n \bigcap_{j=1}^{m_i} E_{i,j}$, où $\forall i, j$ $E_{i,j} \in \mathcal{E}$ ou $(E_{i,j})^c \in \mathcal{E}$).

Remarque : Il est impossible de décrire de façon analogue la σ -algèbre engendrée par \mathcal{E} .

Qu'est-ce qui ne se généralise pas dans la démonstration précédente ?

12.5.2 Formule de Poincaré.

[FOA 17]

Démonstration par récurrence.

Soit \mathcal{A} une tribu sur Ω et P une probabilité sur \mathcal{A} . On se donne une famille finie (A_1, \dots, A_n) d'éléments de \mathcal{A} .

Etablir par récurrence la formule suivante, dite de Poincaré :

$$P(\cup_{i=1}^n A_i) = \sum_{1 \leq i \leq n} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \dots + (-1)^{n-1} P(A_1 \cap \dots \cap A_n).$$

Application : Indicatrice d'Euler

[FOA 37-38]

On choisit au hasard un nombre parmi $\{1, \dots, n\}$, tous les choix étant équiprobables. Soit p un entier inférieur à n . On note A_p l'événement "le nombre choisi est divisible par p ".

— Calculer $P(A_p)$ lorsque p divise n .

- Soit $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ la décomposition en facteurs premiers de l'entier n . Que représente l'événement $(A_{p_1} \cup \cdots \cup A_{p_r})^c$?
- On note $\phi(n)$ le nombre d'entiers strictement inférieurs à n et premiers avec n . Montrer que l'on a

$$\phi(n) = n \prod_{p \text{ premier, } p|n} \left(1 - \frac{1}{p}\right).$$

Application : Facteur distrait.

[FOA 38-39]

Un facteur dispose de n lettre adressées à n destinataires distincts. Il fait la distribution au hasard entre ces n personnes. Quelle est la probabilité de l'événement "une lettre au moins parvient à son destinataire" ? En donner une approximation pour n assez grand.

12.5.3 Équiprobabilité.

Trois dés.

[COT 5-6]

On jette trois dés à 6 faces bien équilibrés.

- Calculer la probabilité d'obtenir au moins un as.
- Que vaut la probabilité d'obtenir au moins deux faces portant le même chiffre ?
- Calculer la probabilité que la somme des points marqués sur les trois faces soit paire.

Happy birthday to you !

[FOA 31]

Quelle est la probabilité que dans votre promotion deux agrégatifs au moins aient leur anniversaire le même jour ?

Remarque : Pour simplifier, on ne s'occupera pas du 29 Février.

Boules et cases.

[DAC¹ exo page 7]

Soit Ω l'ensemble des configurations que l'on peut obtenir en répartissant r boules indiscernables dans n cases numérotées.

1. Montrer que le cardinal de Ω est égal au nombre de solutions $(r_1, \dots, r_n) \in \mathbb{N}^n$ de l'équation : $r_1 + r_2 + \dots + r_n = r$.
2. Montrer que ce nombre vaut C_{n+r-1}^r de 3 façons différentes :
 - (a) par récurrence sur n à l'aide de l'identité : $C_n^p = C_{n-1}^{p-1} + C_{n-2}^{p-1} + \dots + C_{p-1}^{p-1}$.
 - (b) en développant en série entière les deux membres de l'égalité : $\frac{1}{(1-t)^n} = \left(\sum_{k=0}^{\infty} t^k \right)^n$.
 - (c) en identifiant Ω à l'ensemble des façons d'ordonner en ligne r boules indiscernables et $n-1$ séparations (indiscernables aussi).
3. On répartit au hasard r boules indiscernables dans n cases. Calculer la probabilité qu'aucune case ne soit vide.

12.5.4 Probabilités conditionnelles.

Maladie rare.

[COT 11]

On considère une certaine maladie qui touche 1/1000e de la population. Un laboratoire d'analyse de sang assure avec une fiabilité de 99% la détection de cette maladie lorsqu'elle est effectivement présente. Cependant, le test indique aussi un résultat faussement positif pour 0,2% des personnes réellement saines à qui on l'applique.

Quelle est la probabilité qu'une personne soit vraiment malade sachant que son test est positif ? Commenter le résultat.

Le rouge et le noir.

[ROS 66-67]

On considère 3 cartes à jouer de même forme mais de couleurs différentes : la première est noire des deux côtés, la seconde rouge des deux côtés, tandis que la troisième a une face noire et une face rouge.

On mélange les trois cartes au fond d'un chapeau puis on en tire une carte au hasard, dont on ne montre qu'une face. Sachant que cette face est rouge, quelle est la probabilité que l'autre face soit noire ?

Loi de succession de Laplace.

[COT 15-16]

On dispose de $N+1$ urnes, numérotées de 0 à N . L'urne numéro k contient k boules rouges et $N-k$ boules blanches. On choisit une urne au hasard.

Sans connaître son numéro, on en tire n fois de suite une boule, avec remise après chaque tirage. Quelle est la probabilité que le $(n+1)^e$ tirage donne encore une boule rouge sachant que, au cours des n premiers tirages, seules des boules rouges ont été tirées? Calculer la limite de cette probabilité lorsque $N \rightarrow \infty$.

Taux de panne

[COT 16-18]

Soit T une variable aléatoire prenant ses valeurs dans \mathbb{N} telle que pour tout $n \in \mathbb{N}$, $P(T \geq n) > 0$. On appelle taux de panne la suite $\theta(n) = P(T = n \mid T \geq n)$, $n \in \mathbb{N}$.

1. Calculer $P(T = n)$ en fonction des $\theta(k)$, $k \leq n$.
2. Etablir qu'une suite de réels $\theta(k)$ convient comme taux de panne **ssi** :

$$\forall k \in \mathbb{N} \quad 0 \leq \theta(k) < 1 \quad \text{et} \quad \sum_{k=0}^{\infty} \theta(k) = \infty$$

3. Montrer que T suit une loi géométrique **ssi** $\theta(k) = \text{constante}$ pour tout $k \in \mathbb{N}$. (On dit que T suit une loi géométrique de paramètre $a \in]0, 1[$ si pour tout $k \in \mathbb{N}$, $P(T = k) = a(1-a)^k$.)

12.5.5 Événements indépendants.

Événement auto-indépendant.

Montrer qu'un événement A est indépendant de lui-même **ssi** $P(A) = 0$ ou $P(A) = 1$.

Bruit qui court.

[COT 13]

La personne I_1 reçoit l'information 0 ou 1 et la transmet telle quelle à I_2 avec la probabilité p , I_2 de même à I_3 , etc... I_n la transmet au monde entier. On suppose que les n personnes I_1, \dots, I_n sont indépendantes. Quelle est la probabilité p_n que le monde reçoive la bonne information? Calculer $\lim_{n \rightarrow \infty} p_n$.

Indicatrice d'Euler (bis).

On choisit au hasard un nombre parmi $\{1, \dots, n\}$, tous les choix étant équiprobables. Soit p un entier inférieur à n . On note A_p l'événement "le nombre choisi est divisible par p ".

1. Calculer $P(A_p)$ lorsque p divise n .
2. On suppose que (p_1, \dots, p_k) sont des diviseurs premiers distincts de n , montrer que les événements A_{p_1}, \dots, A_{p_k} sont indépendants.
3. On note $\phi(n)$ le nombre d'entiers strictement inférieurs à n et premiers avec n . Montrer que l'on a

$$\phi(n) = n \prod_{p \text{ premier, } p|n} \left(1 - \frac{1}{p}\right).$$

Corrigé :

1. Puisque p divise n (notation : $p|n$), il existe un entier m tel que $n = mp$; on a alors : $A_p = \{p, 2p, \dots, mp\}$. On en déduit : $P(A_p) = \frac{m}{n} = \frac{1}{p}$.
2. Considérant une sous-famille $A_{p_{i_1}}, \dots, A_{p_{i_l}}$, nous remarquons que

$$\cap_{j=1}^l A_{p_{i_j}} = A_{\prod_{j=1}^l p_{i_j}},$$

puisque les p_{i_j} sont premiers. Pour la même raison, $(\prod_{j=1}^l p_{i_j})|n$ donc on peut appliquer a. pour obtenir :

$$P(\cap_{j=1}^l A_{p_{i_j}}) = \frac{1}{\prod_{j=1}^l p_{i_j}} = \prod_{j=1}^l P(A_{p_{i_j}}),$$

ce qui prouve l'indépendance demandée.

3. Prenons ici p_1, \dots, p_k **tous** les diviseurs premiers de n . D'après b), on a :

$$P(\cap_{i=1}^k A_{p_i}^c) = \prod_{p \text{ premier, } p|n} \left(1 - \frac{1}{p}\right).$$

Or un entier appartient à $\cap_{i=1}^k A_{p_i}^c$ **ssi** il n'est divisible par aucun des diviseurs premiers de n , autrement dit **ssi** il est premier avec n . En passant aux probabilités, on en déduit :

$$\prod_{p \text{ premier, } p|n} \left(1 - \frac{1}{p}\right) = \frac{\phi(n)}{n}.$$

Principe de Hardy-Weinberg.

[MEY 43-44]

Les gènes se présentent le plus souvent en paires et sous deux formes alléliques A,a, ce qui donne trois génotypes AA,Aa,aa. Ces génotypes s'expriment par différents phénotypes ; par exemple, dans le cas de l'albinisme, les individus aa sont albinos, contrairement aux individus AA ou Aa.

Chaque individu reçoit au hasard un gène de chacun de ses parents et donc chaque allèle composant le gène d'un des parents a la probabilité $1/2$ de passer à l'enfant. Les génotypes des parents sont indépendants. On note p, q, r les probabilités qu'un adulte, dans une population donnée, ait les génotypes AA,Aa,aa.

1. Calculer les probabilités P, Q, R qu'un enfant ait les génotypes AA,Aa,aa (on pourra poser $\theta = p + q/2$).
2. Montrer que $Q^2 = 4PR$.
3. A quelle condition sur p, q, r a-t-on $P = p, Q = q, R = r$? Montrer qu'elle est toujours réalisée dès la 2ème génération.

Corrigé :

1. Nous prenons pour espace Ω l'ensemble des triplets (génotype de l'enfant, génotype du père, génotype de la mère).

L'évènement « L'enfant a un génotype AA » s'écrit donc :

$$\{(AA, AA, AA); (AA, Aa, AA); (AA, AA, Aa); (AA, Aa, Aa)\}$$

et , d'après les lois de la génétique rappelées dans l'énoncé, admet pour probabilité :

$$P = p^2 + pq/2 + pq/2 + q^2/4 = \theta^2.$$

Un raisonnement symétrique (en échangeant les lettres A et a) nous donne :

$$R = r^2 + rq/2 + rq/2 + q^2/4 = (1 - \theta)^2.$$

Enfin, puisque $P + Q + R = 1$, on obtient : $Q = 2\theta(1 - \theta)$.

2. Cette relation résulte immédiatement de ce qui précède.
3. D'après la question précédente, on a nécessairement $q^2 = 4pr$.

Réciproquement, si cette condition est vérifiée, les égalités obtenues dans la première question nous donnent :

$$P = p^2 + pq + q^2/4 = p^2 + pq + pr = p(p + q + r) = p$$

$$\text{et } R = r^2 + rq + q^2/4 = r^2 + rq + rp = r(p + q + r) = r.$$

Il en résulte $Q = q$ puisqu'on a : $P + Q + R = p + q + r = 1$.

Finalement, l'égalité $q^2 = 4pr$ est une condition nécessaire et suffisante de "stationnarité" des probabilités des différents génotypes au cours des générations.

Ainsi, le résultat du 2) nous dit que dans tous les cas, la stationnarité s'établit dès la deuxième génération.

Chapitre 13

Variables aléatoires réelles

13.1 La tribu borélienne réelle

Définition 13.1.1 Si (E, d) est un espace métrique, on appelle tribu borélienne sur E et l'on note $\mathcal{B}(E)$ la tribu engendrée par l'ensemble des ouverts de E . Les éléments de $\mathcal{B}(E)$ sont appelés ensembles boréliens de E .

Remarque. Nous constatons immédiatement que $\mathcal{B}(E)$ est également la tribu engendrée par l'ensemble des fermés de E . \square

Si l'espace métrique sous-jacent est égal à \mathbb{R} muni de la distance usuelle, nous parlerons de tribu borélienne réelle et nous la noterons donc $\mathcal{B}(\mathbb{R})$.

On ne sait pas décrire explicitement tous les boréliens réels mais il sera suffisant pour la suite de ce cours de savoir que $\mathcal{B}(\mathbb{R})$ contient tous les intervalles réels (fermés, ouverts, semi-ouverts) et leurs unions finies ou dénombrables. On peut construire un sous-ensemble réel qui n'est pas borélien mais ce n'est pas si facile... On peut également démontrer que $\#\mathcal{B}(\mathbb{R}) = \#\mathbb{R}$.

Proposition 13.1.2 La tribu $\mathcal{B}(\mathbb{R})$ est engendrée par chacun des ensembles suivants de parties réelles :

- $\mathcal{E}_1 = \{]a, b[, \text{ avec } (a, b) \in \mathbb{Q}^2 \text{ et } a < b\}$
- $\mathcal{E}_2 = \{[a, b] , \text{ avec } (a, b) \in \mathbb{Q}^2 \text{ et } a < b\}$
- $\mathcal{E}_3 = \{[a, b] , \text{ avec } (a, b) \in \mathbb{Q}^2 \text{ et } a \leq b\}$
- $\mathcal{E}_4 = \{]-\infty, a], \quad a \in \mathbb{Q}\}$
- $\mathcal{E}_5 = \{]-\infty, a[, \quad a \in \mathbb{Q}\}$

Démonstration:

1. Puisque tous les éléments de \mathcal{E}_1 sont des ouverts, on a $\sigma(\mathcal{E}_1) \subset \mathcal{B}(\mathbb{R})$. Prouvons l'inclusion inverse.

Tout ouvert de \mathbb{R} est une union de boules ouvertes, c'est-à-dire d'intervalles de la forme $]a, b[$, avec $(a, b) \in \mathbb{R}^2$ et $a < b$.

Par densité de \mathbb{Q} dans \mathbb{R} , nous pouvons trouver une suite strictement décroissante $(a_n) \in \mathbb{Q}^{\mathbb{N}}$ qui converge vers a et une suite strictement croissante $(b_n) \in \mathbb{Q}^{\mathbb{N}}$ qui converge vers b , de sorte que :

$$]a, b[= \bigcup_{n \in \mathbb{N}}]a_n, b_n[.$$

Par conséquent, tout ouvert de \mathbb{R} est une union d'intervalles ouverts de la forme $]a, b[$, avec $(a, b) \in \mathbb{Q}^2$ et $a < b$. L'ensemble des intervalles ouverts à extrémités rationnelles étant dénombrable, nous avons prouvé que tout ouvert de \mathbb{R} s'écrivait comme une union dénombrable d'éléments de \mathcal{E}_1 , d'où $\mathcal{B}(\mathbb{R}) \subset \sigma(\mathcal{E}_1)$.

2. On montre que $\sigma(\mathcal{E}_2) = \sigma(\mathcal{E}_1)$ en utilisant des égalités de la forme :

$$]a, b] = \bigcap_{n \in \mathbb{N}^*}]a, b + \frac{1}{n}[\quad ; \quad]a, b[= \bigcup_{n \in \mathbb{N}^*}]a, b - \frac{1}{n}]$$

3. On montre que $\sigma(\mathcal{E}_3) = \sigma(\mathcal{E}_2)$ en utilisant des égalités de la forme :

$$[a, b] = \bigcap_{n \in \mathbb{N}^*}]a - \frac{1}{n}, b] \quad ; \quad]a, b] = \bigcup_{n \in \mathbb{N}^*} [a - \frac{1}{n}, b]$$

4. On montre que $\sigma(\mathcal{E}_4) = \sigma(\mathcal{E}_2)$ en utilisant des égalités de la forme :

$$]a, b] =] - \infty, b] -] - \infty, a] \quad ; \quad] - \infty, a] = \bigcup_{n \in \mathbb{N}^*}] - n, a]$$

5. On montre que $\sigma(\mathcal{E}_5) = \sigma(\mathcal{E}_4)$ en utilisant des égalités de la forme :

$$] - \infty, a] = \bigcap_{n \in \mathbb{N}^*}] - \infty, a + \frac{1}{n}[\quad ; \quad] - \infty, a[= \bigcup_{n \in \mathbb{N}^*}] - \infty, a - \frac{1}{n}].$$

□

Remarque. Par passage aux complémentaires, nous constatons immédiatement que la tribu $\mathcal{B}(\mathbb{R})$ est aussi engendrée par les demi-droites de la forme $[a, +\infty[$ (respectivement $]a, +\infty[$), avec $a \in \mathbb{Q}$.

13.2 Les variables aléatoires réelles et leurs lois

Dans toute la suite de ce chapitre, nous considérons un espace probabilisé (Ω, \mathcal{A}, P) .

Dans de très nombreuses situations aléatoires, nous allons nous intéresser à une fonction du résultat ω de l'expérience plutôt qu'à ce résultat lui-même.

Exemple : Un joueur lance une fléchette en visant le centre d'une cible. Si nous prenons pour unité le rayon de la cible, nous pouvons modéliser ceci comme une expérience aléatoire dont le résultat est un élément ω du disque unité fermé :

$$\Omega = \bar{D}(0, 1) = \{\omega \in \mathbb{R}^2, \omega_1^2 + \omega_2^2 \leq 1\}.$$

En réalité, la seule chose qui nous intéresse vraiment n'est pas la position exacte $\omega = (\omega_1, \omega_2)$ de l'impact de la fléchette sur la cible mais seulement sa distance au centre de la cible, qui nous est donnée par la variable aléatoire réelle :

$$\begin{aligned} X : \quad \Omega &\longrightarrow \mathbb{R} \\ \omega = (\omega_1, \omega_2) &\longmapsto \sqrt{\omega_1^2 + \omega_2^2} \end{aligned}$$

Pour mesurer la qualité de la performance du joueur, on va se poser des questions telles que "est-ce que $X(\omega) \in B$?", où B est un sous-ensemble de $[0, 1]$. Par exemple, on peut décider que le joueur a gagné si $X(\omega) \in [0, \frac{1}{10}]$.

Nous allons donc poser la définition suivante.

Définition 13.2.1 *On appelle variable aléatoire réelle (en abrégé v.a.r.) définie sur (Ω, \mathcal{A}, P) toute application $X : \Omega \rightarrow \mathbb{R}$ vérifiant :*

$$\forall B \in \mathcal{B}(\mathbb{R}) \quad X^{-1}(B) \in \mathcal{A}. \quad (13.1)$$

Rappelons que $X^{-1}(B)$ est la partie de Ω définie par $\{\omega \in \Omega, X(\omega) \in B\}$ et que nous la notons encore sous la forme abrégée $\{X \in B\}$.

La définition précédente nous assure que, pour tout borélien réel B , cela a un sens de calculer la probabilité $P(\{X \in B\})$ – que l'on note plus simplement $P(X \in B)$ – puisque $\{X \in B\} \in \mathcal{A}$.

Pour vérifier qu'une application $X : \Omega \rightarrow \mathbb{R}$ est une variable aléatoire réelle, nous serons souvent amenés à utiliser les deux propositions suivantes, que nous admettrons.

Proposition 13.2.2 *Soit \mathcal{E} un ensemble de parties de \mathbb{R} qui engendre la tribu $\mathcal{B}(\mathbb{R})$. Alors une application $X : \Omega \rightarrow \mathbb{R}$ est une variable aléatoire réelle sur (Ω, \mathcal{A}, P) si et seulement si la condition suivante est satisfaite :*

$$\forall B \in \mathcal{E} \quad X^{-1}(B) \in \mathcal{A}.$$

Nous pourrions appliquer cette proposition avec chacune des classes $\mathcal{E}_1, \dots, \mathcal{E}_5$ définies dans la proposition 13.1.2

Par exemple, le lecteur pourra montrer en utilisant la classe \mathcal{E}_4 que, si X et Y sont deux variables aléatoires réelles sur (Ω, \mathcal{A}, P) , alors il en est de même de $\inf(X, Y)$ et $\sup(X, Y)$.

Proposition 13.2.3 *Muni des lois $+$, \times , \cdot , l'ensemble des variables aléatoires réelles définies sur (Ω, \mathcal{A}, P) est une algèbre sur \mathbb{R} .*

Connaître les quantités de la forme $P(X \in B)$ est très intéressant car cela mesure les chances que la valeur prise par la variable X “tombe” dans tel ou tel sous-ensemble borélien réel (par exemple, typiquement, un intervalle).

Ainsi, si nous revenons à l'exemple précédent, une information telle que $P(X \in [0, \frac{1}{10}]) = 0,95$ nous dit que le joueur est très habile.

Théorème et définition 13.2.4 *Considérons un espace probabilisé (Ω, \mathcal{A}, P) et une variable aléatoire réelle $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Alors l'application $P_X : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ définie par :*

$$\forall B \in \mathcal{B}(\mathbb{R}) \quad P_X(B) = P(X \in B)$$

est une mesure de probabilité sur l'espace mesurable $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, appelée loi de probabilité (ou simplement loi) de la v.a.r. X .

Démonstration: Il est immédiat d'après la définition que P_X est une application de $\mathcal{B}(\mathbb{R})$ dans $[0, 1]$ telle que $P_X(\emptyset) = P(X^{-1}(\emptyset)) = P(\emptyset) = 0$ et $P_X(\mathbb{R}) = P(X^{-1}(\mathbb{R})) = P(\Omega) = 1$.

Il nous reste donc à prouver la σ -additivité de P_X . Pour cela, nous considérons une suite $(B_n)_{n \in \mathbb{N}} \in \mathcal{B}(\mathbb{R})^{\mathbb{N}}$ telle que $i \neq j \Rightarrow B_i \cap B_j = \emptyset$.

Nous allons montrer dans un premier temps que :

$$X^{-1}\left(\biguplus_{n \in \mathbb{N}} B_n\right) = \biguplus_{n \in \mathbb{N}} X^{-1}(B_n).$$

En effet, on a :

$$X^{-1}\left(\bigcup_{n \in \mathbb{N}} B_n\right) = \{\omega \in \Omega, \exists n \in \mathbb{N}, X(\omega) \in B_n\} = \bigcup_{n \in \mathbb{N}} X^{-1}(B_n)$$

et, pour $i \neq j$,

$$X^{-1}(B_i) \cap X^{-1}(B_j) = X^{-1}(B_i \cap B_j) = X^{-1}(\emptyset) = \emptyset,$$

d'où le résultat voulu.

Nous déduisons de cette première étape les égalités suivantes :

$$\begin{aligned} P_X \left(\bigcup_{n \in \mathbb{N}} B_n \right) &= P \left(X^{-1} \left(\bigcup_{n \in \mathbb{N}} B_n \right) \right) = P \left(\bigcup_{n \in \mathbb{N}} X^{-1}(B_n) \right) \\ &= \sum_{n \in \mathbb{N}} P(X^{-1}(B_n)) = \sum_{n \in \mathbb{N}} P_X(B_n), \end{aligned}$$

d'où la conclusion. □

Remarque. C'est la condition (13.1) définissant une variable aléatoire réelle X sur (Ω, \mathcal{A}, P) qui nous a permis de "transporter" la probabilité P définie sur son ensemble de départ en une nouvelle probabilité P_X sur son ensemble d'arrivée $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On dit parfois que P_X est la probabilité image de la probabilité P par la variable aléatoire X .

La loi de X nous dit donc avec quelles chances cette v.a.r. prend ses valeurs dans tel ou tel sous-ensemble borélien réel, en particulier dans les intervalles.

Exemple. Soit $a \in \mathbb{R}$ fixé et la variable constante $X \equiv a$. On a donc $P_X(\{a\}) = 1$. C'est la loi la plus simple qui soit : elle est constituée d'une unique masse ponctuelle de poids 1 située au point a . On l'appelle mesure de Dirac au point a et on la note δ_a . C'est la mesure définie sur tout $A \in \mathcal{B}(\mathbb{R})$ par $\delta_a(A) = 1$ si $a \in A$ et $\delta_a(A) = 0$ si $a \notin A$.

Remarque. Il est fréquent d'utiliser de façon interchangeable les termes *mesure de probabilité* et *loi de probabilité* sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

En effet, par définition même, la loi de probabilité d'une v.a.r. est bien une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Mais réciproquement, nous pouvons faire apparaître n'importe quelle mesure de probabilité μ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ comme la loi d'une certaine variable aléatoire que nous allons construire maintenant.

Prenons pour espace probabilisé $(\Omega, \mathcal{A}, P) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ et définissons X comme étant l'application identité sur $\Omega = \mathbb{R}$. Le lecteur vérifiera aisément que X est une variable aléatoire réelle et que la loi de X est égale à μ .

13.3 Fonction de répartition d'une variable aléatoire réelle

A priori, pour déterminer la loi d'une v.a.r. X , il faut connaître la valeur de $P(X \in B)$ pour tout borélien réel B . En fait, nous allons voir dans ce paragraphe qu'il suffit de connaître $P(X \in B)$ pour des boréliens réels bien particuliers : les demi-droites $B =] - \infty, x]$, $x \in \mathbb{R}$.

Nous ne démontrerons pas complètement ce résultat car nous admettrons le résultat intermédiaire suivant : la loi d'une v.a.r. X est entièrement déterminée par ses valeurs sur tous les intervalles réels.

Définition 13.3.1 *On appelle fonction de répartition (en abrégé f.r.) de la v.a.r. X l'application $F_X : \mathbb{R} \rightarrow [0, 1]$ définie par :*

$$\forall x \in \mathbb{R}, \quad F_X(x) = P(X \leq x)$$

Remarque. On peut très bien parler de fonction de répartition de la loi de X car en réalité F_X ne dépend que de P_X :

$$\forall x \in \mathbb{R} \quad F_X(x) = P(X \in] - \infty, x]) = P_X(] - \infty, x]).$$

Ainsi, une autre façon de présenter cette notion consiste à définir d'abord la fonction de répartition d'une mesure de probabilité μ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ par :

$$\forall x \in \mathbb{R} \quad F_\mu(x) = \mu(] - \infty, x])$$

puis à dire que la fonction de répartition d'une variable aléatoire réelle est égale par définition à celle de sa loi de probabilité. \square

Le théorème suivant nous dit que la connaissance de la fonction de répartition d'une loi de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ détermine cette loi de façon unique.

Théorème 13.3.2 *Si deux variables aléatoires réelles X et Y ont même fonction de répartition, alors elles ont même loi :*

$$F_X = F_Y \implies P_X = P_Y$$

Autrement dit, l'application qui à une loi associe sa fonction de répartition est injective.

Démonstration: Connaître F_X , c'est connaître la valeur de la loi P_X sur les demi-droites $] -\infty, x]$, $x \in \mathbb{R}$. Nous allons montrer que ceci suffit à calculer la valeur de P_X sur n'importe quel intervalle réel I et nous admettrons que la loi P_X est alors parfaitement déterminée sur tous les boréliens réels B . Nous énumérons les différentes formes possibles pour l'intervalle I et donnons à chaque fois une formule de calcul de $P_X(I)$ à partir de la connaissance de F_X .

Si I est une demi-droite de la forme $] -\infty, x]$, c'est un cas évident puisque par définition $P_X(]-\infty, x]) = F_X(x)$.

Si I est une demi-droite ouverte de la forme $] -\infty, x[$, par passage à la limite croissante dans une probabilité, nous pouvons écrire l'égalité :

$$P_X(]-\infty, x[) = \lim_n \uparrow P_X(]-\infty, x - \frac{1}{n}]) = \lim_n \uparrow F_X(x - \frac{1}{n})$$

Par passage au complémentaire dans ces deux premiers cas, nous traitons les demi-droites de la forme $]x, +\infty[$ ou $[x, +\infty[$.

$$P_X([x, +\infty[) = 1 - P_X(]-\infty, x]) = 1 - F_X(x)$$

$$P_X(]x, +\infty[) = 1 - P_X(]-\infty, x]) = 1 - \lim_n \uparrow F_X(x - \frac{1}{n})$$

Nous pouvons maintenant traiter les cas où I est un intervalle borné, en le faisant apparaître comme différence propre de demi-droites. Soient a et b deux réels tels que $a \leq b$.

$$P_X([a, b]) = P_X(]-\infty, b]) - P_X(]-\infty, a]) = F_X(b) - F_X(a)$$

De la même façon, on trouve :

$$P_X([a, b]) = P_X(]-\infty, b]) - P_X(]-\infty, a]) = F_X(b) - \lim_n \uparrow F_X(a - \frac{1}{n})$$

Nous laissons au lecteur le soin de traiter de façon similaire les cas restants. \square

Comment utiliser ce théorème ? Dans la suite, nous citerons un certain nombre de lois de probabilité "classiques", c'est-à-dire d'un usage fréquent en calcul des probabilités. Pour la plupart d'entre elles, il est possible de calculer explicitement les fonctions de répartition correspondantes. Une façon de déterminer la loi d'une v.a.r. X sera de calculer sa fonction de répartition. Si nous reconnaissons alors la fonction de répartition d'une loi classique, en vertu du théorème précédent, nous pourrions affirmer que la loi de X est égale à cette

loi classique.

Nous terminons cette section par un théorème permettant de caractériser les fonctions de répartition d'une loi de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ou encore les fonctions de répartition des variables aléatoires réelles, ce qui revient au même d'après la remarque qui termine la section précédente.

Théorème 13.3.3 *Considérons une application $F : \mathbb{R} \rightarrow \mathbb{R}$. Alors F est la fonction de répartition d'une loi de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ si et seulement si elle possède simultanément les quatre propriétés suivantes : F est croissante, continue à droite et vérifie*

$$\lim_{t \rightarrow -\infty} F(t) = 0 \text{ et } \lim_{t \rightarrow +\infty} F(t) = 1.$$

Démonstration: Le sens direct sera démontré en cours. Le sens réciproque est admis. \square

La section et le chapitre suivants concernent deux types particulièrement intéressants de variables aléatoires réelles : les variables discrètes et les variables admettant une densité. Notons que ces deux types sont loin de couvrir toutes les possibilités concernant les v.a.r. Ce sont simplement deux cas particuliers importants et eux seuls figurent au programme officiel.

13.4 Variables discrètes

Une variable aléatoire réelle définie sur (Ω, \mathcal{A}, P) est dite discrète si le sous-ensemble réel $X(\Omega)$ est fini ou dénombrable. Autrement dit, une v.a.r. est dite discrète si elle ne prend qu'un nombre fini ou dénombrable de valeurs. C'est un cas particulièrement simple et si nous notons $X(\Omega) = \{x_0, \dots, x_n, \dots\}$ (dans le cas dénombrable), alors la loi de la variable aléatoire X est entièrement déterminée par la donnée de la suite $(p_n)_{n \in \mathbb{N}}$ définie par :

$$\forall n \in \mathbb{N} \quad p_n = P(X = x_n).$$

Cette suite est bien sûr à valeurs positives ou nulles et telle que

$$\sum_{n=0}^{+\infty} p_n = 1$$

et le lecteur pourra faire le rapprochement avec la proposition 12.3.3.

Le lecteur est invité à se référer à [DAN 418-436 (ou 431-452 pour l'édition 2016)] qui traite de façon exhaustive les notions du programme officiel qui sont relatives aux variables aléatoires discrètes, en particulier les lois

discrètes usuelles suivantes : Bernoulli, binomiale, hypergéométrique, géométrique, Poisson.

Nous n'écrivons ici que la définition de l'espérance d'une variable discrète et le *théorème de transfert* dans le cas discret.

Définition 13.4.1 Soit X une variable aléatoire réelle discrète telle que $X(\Omega) = \{x_0, \dots, x_n, \dots\}$. On appelle espérance de la variable X le nombre réel :

$$E[X] = \sum_{n=0}^{+\infty} x_n P(X = x_n),$$

dès lors que cette série est absolument convergente. On dit alors que la variable X admet une espérance (ou encore qu'elle est intégrable).

Théorème 13.4.2 (de transfert, cas discret) Soit $\psi : \mathbb{R} \rightarrow \mathbb{R}$ une application ; alors la variable aléatoire $\psi(X)$ admet pour espérance le nombre réel

$$E[\psi(X)] = \sum_n \psi(x_n) P(X = x_n)$$

dès que cette série est absolument convergente.

Un outil très utile pour étudier les variables aléatoires à valeurs dans \mathbb{N} est présenté maintenant sous la forme d'un problème corrigé.

Problème : Fonctions génératrices On considère une variable aléatoire X à valeurs dans \mathbb{N} .

1. Montrer que pour tout $s \in [-1, 1]$, la variable aléatoire s^X admet une espérance (on pose $s^0 = 1$ pour tout s) et que l'on a : $E[s^X] = \sum_{n=0}^{\infty} P(X = n) s^n$.
On notera $G_X(s) = E[s^X]$ et l'on appellera G_X la fonction génératrice de X .
2. Montrer que X et Y ont même loi si et seulement si $G_X = G_Y$.
3. Calculer les fonctions génératrices des lois suivantes :
 - (a) Bernoulli de paramètre $p \in [0, 1]$.
 - (b) Géométrique de paramètre $a \in]0, 1[$.
 - (c) Poisson de paramètre λ .
4. Si X et Y sont deux variables indépendantes, calculer G_{X+Y} en fonction de G_X et G_Y . En déduire :
 - (a) la fonction génératrice d'une loi binomiale de paramètres $(n, p) \in \mathbb{N}^* \times [0, 1]$.

- (b) la loi de la somme de deux variables indépendantes suivant des lois de Poisson de paramètres respectifs λ et μ .
5. Montrer que si X est intégrable, alors G_X est dérivable sur $[-1, 1]$, de dérivée : $G'_X(s) = E[Xs^{X-1}]$.
En déduire les espérances des lois introduites dans la question 3.

Corrigé :

1. Pour tout $s \in [-1, 1]$, nous avons l'inégalité $|P(X = n)s^n| \leq P(X = n)$ qui entraîne l'absolue convergence de la série

$$\sum_{n=0}^{\infty} P(X = n)s^n$$

donc s^X est intégrable et admet pour espérance

$$E[s^X] = \sum_{n=0}^{\infty} P(X = n)s^n.$$

2. L'égalité obtenue dans la question précédente nous donne le développement en série entière de G_X sur $[-1, 1]$. En particulier, G_X admet des dérivées à tous les ordres au point $s = 0$ et l'on a, pour tout $n \in \mathbb{N}$:

$$P(X = n) = \frac{G_X^{(n)}(0)}{n!}$$

Il en résulte que $G_X = G_Y \implies P_X = P_Y$; la réciproque est évidente. La fonction génératrice de X caractérise donc la loi de X .

3. (a) $G_X(s) = 1 - p + ps$.
(b) $G_X(s) = \sum_{n=0}^{+\infty} (1-a)a^n s^n = \frac{1-a}{1-as}$.
(c) $G_X(s) = \sum_{n=0}^{+\infty} e^{-\lambda} \frac{\lambda^n}{n!} s^n = e^{-\lambda(1-s)}$.
4. On a par définition : $G_{X+Y}(s) = E[s^{X+Y}] = E[s^X s^Y]$.
Si X et Y sont indépendantes, on en déduit, en utilisant une propriété des variables discrètes indépendantes énoncée dans [DAN 427] :

$$\forall s \in [-1, 1] \quad G_{X+Y}(s) = E[s^X]E[s^Y] = G_X(s)G_Y(s)$$

On peut également faire une preuve directe sans utiliser cette propriété, en considérant le produit de Cauchy de deux séries entières absolument convergentes [DAN 430].

- (a) Une récurrence immédiate prouve que si X_1, \dots, X_n sont des variables aléatoires à valeurs dans \mathbb{N} indépendantes, alors :

$$\forall s \in [-1, 1] \quad G_{X_1 + \dots + X_n}(s) = G_{X_1}(s) \cdots G_{X_n}(s).$$

En particulier, si X_1, \dots, X_n suivent la loi de Bernoulli de paramètre p , alors $X_1 + \dots + X_n$ suit la loi binomiale de paramètres (n, p) , dont la fonction génératrice vaut donc (en utilisant 3(a)) : $G(s) = (1 - p + ps)^n$.

- (b) Si $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$ sont indépendantes, alors l'égalité obtenue en 3(c) nous permet d'écrire :

$$G_{X+Y}(s) = G_X(s)G_Y(s) = e^{-\lambda(1-s)}e^{-\mu(1-s)} = e^{-(\lambda+\mu)(1-s)}.$$

On reconnaît la fonction génératrice de la loi $\mathcal{P}(\lambda + \mu)$, ce qui d'après la question 2 prouve que $X + Y$ suit la loi de Poisson de paramètre $\lambda + \mu$.

5. D'après la première question, la série entière qui définit la fonction génératrice de la variable aléatoire X :

$$G_X(s) = \sum_{n=0}^{\infty} P(X = n)s^n$$

a un rayon de convergence $R \geq 1$. Nous en déduisons que G_X est dérivable sur $] -1, 1[$ et que :

$$\forall s \in] -1, 1[\quad G'_X(s) = \sum_{n=1}^{\infty} P(X = n)ns^{n-1}.$$

Notre hypothèse d'intégrabilité sur X équivaut à la convergence de la série à termes positifs suivante :

$$\sum_{n=1}^{\infty} nP(X = n).$$

Un résultat classique sur les séries entières nous permet d'en déduire que la série :

$$\sum_{n=1}^{\infty} P(X = n)ns^{n-1}$$

converge uniformément sur $[0,1]$ vers une application continue. En particulier, $G'_X(s)$ admet une limite quand $s \nearrow 1$ et nous savons que cela entraîne la dérivabilité à gauche de G_X au point 1, ainsi que l'égalité :

$$G'_X(1) = \lim_{s \nearrow 1} G'_X(s) = \lim_{s \nearrow 1} \sum_{n=1}^{\infty} P(X = n) n s^{n-1} = \sum_{n=1}^{\infty} n P(X = n) = E[X].$$

Dans les trois exemples suivants, nous allons donc utiliser l'égalité :

$$E[X] = G'_X(1).$$

Si $X \sim \text{Bernoulli}(p)$, $G'_X(s) = 1 - p$ donc $E[X] = 1 - p$.

Si $X \sim \text{Géom}(a)$, $G'_X(s) = \frac{a(1-a)}{(1-as)^2}$ donc $E[X] = \frac{a}{1-a}$.

Enfin, pour $X \sim \mathcal{P}(\lambda)$, $G'_X(s) = \lambda e^{-\lambda(1-s)}$ d'où $E[X] = \lambda$.

Chapitre 14

Variables à densité

À l'extrême opposé d'une variable aléatoire discrète, dont la loi n'est faite que de masses ponctuelles, la loi d'une variable aléatoire absolument continue correspond à une répartition continue de masse sur la droite réelle et en particulier n'admet aucune masse ponctuelle. Plus précisément, si nous reprenons l'analogie entre une loi sur \mathbb{R} et une répartition de masse sur un fil infiniment long et infiniment mince de masse totale 1, nous allons maintenant nous intéresser à la situation correspondant à un fil admettant une densité de masse linéique, à savoir une application f à valeurs dans \mathbb{R}_+ telle que la masse du segment élémentaire $[x, x + dx]$ soit $f(x) dx$.

Nous rencontrerons des variables aléatoires absolument continues dans de nombreuses situations de modélisation où interviennent des grandeurs continues comme le temps (1er instant de panne d'une machine, temps d'attente à un guichet etc.), l'espace (mesure d'une longueur perturbée par de petites erreurs aléatoires) ou autre (pression atmosphérique mesurée dans une station météo).

14.1 Définitions

Définition 14.1.1 On appelle densité de probabilité sur \mathbb{R} toute application $f \in C_M(\mathbb{R}, \mathbb{R}_+)$ telle que :

$$\int_{-\infty}^{+\infty} f(x) dx = 1.$$

Définition 14.1.2 Soit f une densité de probabilité sur \mathbb{R} . On dit qu'une variable aléatoire réelle X admet f pour densité si, pour tout intervalle réel I , on a l'égalité :

$$P(X \in I) = \int_I f(x) dx.$$

Remarques.

1. Nous avons admis que la connaissance des quantités $P(X \in I)$ pour tout intervalle réel I déterminait entièrement la loi de X . Ainsi, si X admet la densité f , sa loi P_X est entièrement déterminée. On dit que P_X est la loi de densité f ou encore que X suit la loi de densité f .
2. Il n'est pas nécessaire de vérifier à l'avance que f est une densité de probabilité sur \mathbb{R} mais seulement que $f \in C_M(\mathbb{R}, \mathbb{R}_+)$. En effet, en appliquant l'égalité précédente avec $I = \mathbb{R}$, on en déduit que f est bien une densité de probabilité.

14.2 Exemples classiques de lois à densité

La loi uniforme sur $[a, b]$

C'est la loi de densité $f(x) = \frac{1}{b-a} \mathbf{1}_{[a,b]}(x)$ (on vérifie immédiatement que $f \in C_M(\mathbb{R}, \mathbb{R}_+)$ et $\int_{\mathbb{R}} f(x) dx = 1$).

Cette loi est notée $U([a, b])$. Elle correspond à l'expression "tirer un nombre au hasard entre a et b ". La fonction 'random' d'une calculatrice ou d'un ordinateur est censée simuler une telle loi pour $a = 0$ et $b = 1$.

Si la v.a.r. X admet f pour densité, on dit qu'elle suit la loi uniforme sur $[a, b]$ et l'on note $X \sim U([a, b])$.

La loi exponentielle

Soit $\lambda > 0$ fixé. L'application f_λ définie pour tout $x \in \mathbb{R}$ par :

$$f_\lambda(x) = \lambda e^{-\lambda x} \mathbf{1}_{\mathbb{R}_+^*}(x)$$

est une densité de probabilité sur \mathbb{R} .

En effet, on constate facilement que $f_\lambda \in C_M(\mathbb{R}, \mathbb{R}_+)$ et nous calculons :

$$\int_{-\infty}^{+\infty} f_\lambda(x) dx = \int_0^{+\infty} \lambda e^{-\lambda x} dx = [-e^{-\lambda x}]_0^{+\infty} = 1$$

On appelle loi exponentielle de paramètre λ la loi de densité f_λ . Comme nous le verrons dans l'exercice 14.6.2, cette loi est caractérisée par son "absence de mémoire". Par exemple, la durée de vie d'un matériel idéal qui ne subirait aucun vieillissement est une variable aléatoire exponentielle.

La loi gaussienne (ou loi normale)

Nous montrerons plus tard que l'application f définie par :

$$\forall x \in \mathbb{R}, \quad f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

vérifie bien $\int_{\mathbb{R}} f(x) dx = 1$. Comme il est clair que $f \in C_M(\mathbb{R}, \mathbb{R}_+)$ (elle est même continue), l'application f représentée ci-dessous est une densité de probabilité appelée *densité gaussienne (ou normale) centrée réduite*. La loi associée est notée $\mathcal{N}(0, 1)$.

Si nous fixons deux paramètres $\mu \in \mathbb{R}$ et $\sigma > 0$, nous déduisons de ce qui précède que l'application $f_{\mu, \sigma}$ définie par :

$$\forall x \in \mathbb{R}, \quad f_{\mu, \sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

est aussi une densité de probabilité (il suffit de faire un simple changement de variable dans une intégrale pour le voir). On l'appelle densité gaussienne (ou normale) de paramètres μ et σ^2 . La loi associée est notée $\mathcal{N}(\mu, \sigma^2)$.

La loi de Cauchy

Pour tout $a > 0$ fixé, nous définissons l'application f_a par :

$$\forall x \in \mathbb{R}, \quad f_a(x) = \frac{a}{\pi} \frac{1}{a^2 + x^2}.$$

On a bien sûr $f \in C_M(\mathbb{R}, \mathbb{R}_+)$ (elle est même continue) et l'on calcule :

$$\int_{-\infty}^{+\infty} f_a(x) dx = \frac{1}{\pi} \left[\arctan \frac{x}{a} \right]_{-\infty}^{+\infty} = 1$$

L'application f_a est donc une densité de probabilité appelée densité de Cauchy de paramètre a . La loi associée est notée $C(a)$.

La loi gamma

Pour $a > 0$ et $\lambda > 0$, nous définissons l'application $g_{a, \lambda}$ par :

$$g_{a, \lambda}(x) = \frac{\lambda^a}{\Gamma(a)} e^{-\lambda x} x^{a-1} \mathbf{1}_{\mathbb{R}_+^*}(x).$$

Dans le cas $a \geq 1$ et $\lambda > 0$, on vérifie immédiatement que $g_{a, \lambda} \in C_M(\mathbb{R}, \mathbb{R}_+)$ et l'on calcule, avec le changement de variable $y = \lambda x$,

$$\int_{-\infty}^{+\infty} g_{a, \lambda}(x) dx = \frac{1}{\Gamma(a)} \int_0^{+\infty} e^{-\lambda x} (\lambda x)^{a-1} \lambda dx = \frac{1}{\Gamma(a)} \int_0^{+\infty} e^{-y} y^{a-1} dy = 1.$$

L'application $g_{a,\lambda}$ est donc une densité de probabilité appelée densité gamma de paramètres a et λ . La loi associée est notée $\gamma(a, \lambda)$.

Remarque. Dans le cas $(a, \lambda) \in]0, 1[\times \mathbb{R}_+^*$, on a $g_{a,\lambda} \notin C_M(\mathbb{R}, \mathbb{R}_+)$ puisque :

$$\lim_{x \rightarrow 0+} g_{a,\lambda}(x) = \lim_{x \rightarrow 0+} \frac{\lambda^a}{\Gamma(a)} e^{-\lambda x} x^{a-1} = +\infty.$$

L'application $g_{a,\lambda}$ n'est donc pas une densité de probabilité sur \mathbb{R} , au sens de la Définition 14.1.1, page 255. Néanmoins, on a $g_{a,\lambda} \in C_M(\mathbb{R}_+^*, \mathbb{R}_+)$ et $\int_{\mathbb{R}_+^*} g_{a,\lambda} = 1$. Nous dirons que $g_{a,\lambda}$ est une densité de probabilité sur \mathbb{R}_+^* , permettant de définir la loi $\gamma(a, \lambda)$ dans le cas général $(a, \lambda) \in (\mathbb{R}_+^*)^2$.

Cette difficulté technique n'aura aucune incidence quand nous étudierons la loi $\gamma(a, \lambda)$ dans la suite car nous prendrons toujours soin de nous ramener à des calculs d'intégrales sur \mathbb{R}_+^* .

14.3 Fonction de répartition

Proposition 14.3.1 *Soit $f \in C_M(\mathbb{R}, \mathbb{R}_+)$ et X une variable aléatoire réelle. L'application f est une densité de probabilité et la variable X admet f pour densité si et seulement si la fonction de répartition de X est donnée par :*

$$\forall x \in \mathbb{R}, \quad F_X(x) = \int_{-\infty}^x f(t) dt. \quad (14.1)$$

Démonstration: Dans le sens direct, cela résulte de la définition 14.1.2 avec $I =]-\infty, x]$.

Dans le sens réciproque, en faisant tendre x vers l'infini dans l'égalité (14.1), on obtient grâce au théorème 13.3.3, page 250,

$$\int_{-\infty}^{+\infty} f(t) dt = 1,$$

et donc f est bien une densité de probabilité. En outre, la variable X suit bien la loi de densité f puisque la fonction de répartition caractérise la loi d'une variable aléatoire réelle.

□

À titre d'application, nous allons démontrer la proposition suivante.

Proposition 14.3.2 *Soit X une v.a.r. admettant la densité f . Nous définissons la v.a.r. $Y = \sigma X + \mu$, où $\mu \in \mathbb{R}$ et $\sigma > 0$ sont deux paramètres fixés.*

Alors Y admet la densité $f_{\mu,\sigma}$ définie par la formule :

$$\forall x \in \mathbb{R}, \quad f_{\mu,\sigma}(x) = \frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}\right) \quad (14.2)$$

Démonstration: Calculons la f.r. de Y pour tout $t \in \mathbb{R}$:

$$F_Y(t) = P(Y \leq t) = P(\sigma X + \mu \leq t) = P\left(X \leq \frac{t - \mu}{\sigma}\right).$$

Puisque la variable X admet la densité f , nous en déduisons :

$$\forall t \in \mathbb{R}, \quad F_Y(t) = \int_{-\infty}^{\frac{t - \mu}{\sigma}} f(x) dx.$$

Dans cette intégrale, nous effectuons le changement de variable $y = \sigma x + \mu \Leftrightarrow x = \frac{y - \mu}{\sigma}$, ce qui nous donne :

$$\forall t \in \mathbb{R}, \quad F_Y(t) = \int_{-\infty}^t f\left(\frac{y - \mu}{\sigma}\right) \frac{1}{\sigma} dy.$$

La proposition 14.3.1 nous permet d'en déduire que l'application $f_{\mu,\sigma}$ définie par la formule (14.2) est une densité de probabilité et que Y suit la loi de densité $f_{\mu,\sigma}$. □

Proposition 14.3.3 *Soit X une variable aléatoire réelle de fonction de répartition F_X et $f \in C(\mathbb{R}, \mathbb{R}_+)$. Alors la variable X admet f pour densité si et seulement si F_X est dérivable sur \mathbb{R} avec :*

$$\forall x \in \mathbb{R}, \quad F'_X(x) = f(x).$$

Démonstration: Supposons que X admet la densité f . Alors la fonction de répartition de X est donnée par :

$$\forall x \in \mathbb{R}, \quad F_X(x) = \int_{-\infty}^x f(t) dt.$$

Comme f est continue, nous en déduisons que F_X est dérivable sur \mathbb{R} et que $F'_X = f$.

Dans le sens réciproque, puisque F_X est une primitive de f , nous avons pour tous réels $w \leq x$:

$$\int_w^x f(t) dt = F_X(x) - F_X(w).$$

Si nous faisons tendre w vers $-\infty$ dans l'égalité précédente, nous obtenons grâce au théorème 13.3.3 :

$$\forall x \in \mathbb{R}, \quad \int_{-\infty}^x f(t) dt = F_X(x).$$

Ceci nous permet de conclure, en vertu de la proposition 14.3.1. \square

Remarque. Cet énoncé admet une généralisation au cas f continue par morceaux. Le lecteur pourra consulter [DAN 408] à ce sujet. \square

Voici une application de la proposition 14.3.3 : Soit X v.a.r. de loi uniforme sur l'intervalle $] -\frac{\pi}{2}, \frac{\pi}{2}[$. Que vaut la loi de $Y = \tan X$?

Calculons la f.r. de Y pour tout $t \in \mathbb{R}$:

$$F_Y(t) = P(Y \leq t) = P(X \leq \arctan t) = \int_{-\frac{\pi}{2}}^{\arctan t} \frac{1}{\pi} dx = \frac{\arctan t + \frac{\pi}{2}}{\pi}.$$

Cette fonction étant de classe C^1 sur \mathbb{R} , en calculant $f = F'_Y$, nous déduisons de la proposition précédente que Y admet pour densité l'application f définie par $f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$ pour tout $x \in \mathbb{R}$.

Autrement dit, Y suit la loi de Cauchy de paramètre 1.

Remarque. La fonction de répartition d'une v.a.r. de densité f s'écrivant $F_X(t) = \int_{-\infty}^t f(x) dx$, c'est en particulier une application continue, d'où le qualificatif « continue » donné à la v.a.r.

Pourquoi précise-t-on v.a.r. *absolument* continue ? Parce que les v.a.r. admettant une densité ne sont qu'un cas particulier de variables dont la f.r. est continue. C'est néanmoins le seul type de variables continues que nous examinerons dans le cadre de ce cours.

14.4 Espérance. Théorème de transfert

Définition 14.4.1 Soit X une variable aléatoire réelle de densité f . On appelle espérance de X le nombre réel

$$E[X] = \int_{-\infty}^{+\infty} x f(x) dx,$$

dès lors que cette intégrale impropre est absolument convergente.

Nous énonçons simplement ici un important résultat sur lequel nous reviendrons plus précisément dans le chapitre suivant dans le cas des vecteurs aléatoires. L'implication $1 \Rightarrow 2$ s'appelle le *théorème de transfert* pour une variable aléatoire réelle à densité.

Proposition 14.4.2 *Soit X une variable aléatoire réelle et $f : \mathbb{R} \rightarrow \mathbb{R}_+$ une application intégrable sur \mathbb{R} . Alors il y a équivalence entre les deux propositions suivantes :*

1. *L'application f est une densité de probabilité sur \mathbb{R} et la variable aléatoire réelle X admet f pour densité.*
2. *Pour toute application $\psi \in C_M(\mathbb{R}, \mathbb{R})$ telle que $|\psi|f$ soit intégrable sur \mathbb{R} , la variable aléatoire $\psi(X)$ admet pour espérance :*

$$E[\psi(X)] = \int_{\mathbb{R}} \psi(x)f(x) dx.$$

Remarque. La variable aléatoire $\psi(X)$ n'admet pas forcément de densité, comme nous le constatons en prenant $\psi \equiv 0$. \square

Le lecteur est invité à consulter [DAN 438-450] pour des exemples et propriétés de variables à densité (que Dantzer appelle des variables *continues*).

14.5 Moments d'une variable aléatoire

Le lecteur pourra étudier dans le livre d'ESCOFFIER ou dans le tome 1 d'OUVRARD (par exemple) les notions de *moments*, *variance* et *écart-type* d'une variable aléatoire réelle en distinguant le cas discret et le cas où la variable admet une densité.

Une application importante de la notion de variance est l'inégalité de Bienaymé-Tchebychev, qui est une clé dans la démonstration de la loi faible des grands nombres. Dans l'énoncé suivant, la v.a.r. X est soit discrète, soit à densité.

Proposition 14.5.1 (Inégalité de Bienaymé-Tchebychev) *Considérons une variable aléatoire réelle X admettant un moment d'ordre 2. Pour tout $t > 0$, nous avons :*

$$P(|X - E[X]| \geq t) \leq \frac{\text{Var} X}{t^2}$$

Dans les ouvrages cités ci-dessus, le lecteur trouvera encore les notions de *covariance* et de *coefficient de corrélation* pour un couple de variables aléatoires. Ces notions sont introduites en distinguant à nouveau le cas discret et le cas où les variables admettent des densités, conformément au programme officiel.

14.6 Exercices sur les fonctions de répartition

Dans cette section, nous regroupons quelques exercices sur les variables aléatoires réelles qui peuvent tous être résolus en utilisant la notion de fonction de répartition. En outre, les exercices 14.6.3, 14.6.4 et 14.6.5 peuvent être résolus par une seconde méthode : utiliser la proposition 14.4.2 en faisant un changement de variable dans une intégrale sur \mathbb{R} .

14.6.1 Minimum de variables exponentielles.

Soient X_1, \dots, X_n des v.a.r. suivant des lois exponentielles de paramètres respectifs $\lambda_1, \dots, \lambda_n$. On suppose que ces v.a.r. sont indépendantes, ce qui équivaut à :

$\forall (x_i)_{1 \leq i \leq n} \in \mathbb{R}^n$, les événements $\{X_i \leq x_i\}, 1 \leq i \leq n$, sont mutuellement indépendants.

Calculer la loi de $X = \min_{1 \leq i \leq n} X_i$.

14.6.2 Variables amnésiques.

[COT 91-92] Soit T une v.a. à valeurs réelles telle que, pour tous $s, t \geq 0$, on ait :

$$P(T > t + s) = P(T > t)P(T > s).$$

Le but de cet exercice est de montrer que, soit $P(T > 0) = 0$, soit T suit une loi exponentielle.

1. Montrer que si $P(T > 0) > 0$, alors pour tout $t > 0$, $P(T > t) > 0$.
2. On définit alors l'application $f(t) = \ln P(T > t)$, $t > 0$.
Montrer que $f(x) = xf(1)$ pour tout x rationnel positif, puis pour tout x réel positif.
3. Conclure.

14.6.3 Loi du χ^2 à un degré de liberté.

[COT 83 et 86]

Soit $X \sim \mathcal{N}(0, 1)$. Montrer que la variable aléatoire réelle $Y = X^2$ admet pour densité l'application $f : \mathbb{R} \rightarrow \mathbb{R}_+$ définie par :

$$\forall y \in \mathbb{R} \quad f(y) = \frac{1}{\sqrt{2\pi y}} \exp\left(-\frac{y}{2}\right) \mathbf{1}_{\mathbb{R}_+^*}(y).$$

Cette densité définit la loi du χ^2 à un degré de liberté.

14.6.4 Loi gaussienne dans \mathbb{R} .

Soit $X \sim \mathcal{N}(0, 1)$. Montrer que la variable aléatoire réelle $Y = m + \sigma X$, où $m \in \mathbb{R}$ et $\sigma > 0$, admet une densité que l'on calculera. Comment s'appelle la loi de Y ?

14.6.5 Avec une loi de Cauchy.

[COT 52-53] Soit X une variable aléatoire réelle suivant une loi de Cauchy de paramètre 1.

1. Quelle est la loi de aX , où $a \in \mathbb{R}^*$?
2. Montrer que $Y = \ln |X|$ admet la densité : $p(y) = \frac{1}{\pi \cosh y}$.

Chapitre 15

Vecteurs aléatoires et indépendance

Remarque préliminaire : Les expressions *vecteur aléatoire de dimension n* et *variable aléatoire à valeurs dans \mathbb{R}^n* sont synonymes. Il est donc approprié de s'intéresser aux vecteurs aléatoires dans le cadre d'une leçon telle que « Variables aléatoires possédant une densité. Exemples. »

Pour introduire la notion de vecteur aléatoire, nous nous intéressons à quelques situations particulières :

- Nous choisissons une personne au hasard dans la population française ; soit X son âge et Y son taux de cholestérol. Est-ce qu'il y a un lien entre les v.a.r. X et Y ou bien prennent-elles leurs valeurs de façon indépendante ?
- Nous observons la trajectoire de la fusée Ariane en repérant sa position à différents instants $t_i, 1 \leq i \leq p$, grâce à ses coordonnées $(X_{t_i}, Y_{t_i}, Z_{t_i}), 1 \leq i \leq p$ dans un certain repère orthonormé. Pour tenir compte des erreurs entre la trajectoire idéale et la trajectoire réelle, chacune de ces coordonnées est considérée comme une v.a.r. Une idée de la trajectoire réelle nous est alors donnée par le “grand vecteur” $(X_{t_1}, Y_{t_1}, Z_{t_1}, \dots, X_{t_p}, Y_{t_p}, Z_{t_p})$ qui est aléatoire.
- Dans une chaîne de production, le responsable du contrôle qualité prélève p pièces au hasard sur la production de la journée et en mesure les masses respectives X_1, \dots, X_p . Comment peut-il estimer le poids moyen d'une pièce produite dans la journée ou encore l'écart-type de la masse d'une pièce ?

Dans toutes ces situations, nous constatons notre besoin d'étudier plusieurs variables aléatoires réelles simultanément et non plus une par une. De fa-

çon générale, nous sommes amenés à étudier des n -uplets dont chacune des composantes est une variable aléatoire réelle.

15.1 Les vecteurs aléatoires et leurs lois

Dans toute la suite, nous considérerons un espace probabilisé (Ω, \mathcal{A}, P) .

Avant de définir les vecteurs aléatoires à valeurs dans \mathbb{R}^n , nous rappelons que la tribu borélienne sur \mathbb{R}^n est la tribu engendrée par l'ensemble des ouverts de \mathbb{R}^n (pour la topologie usuelle).

Proposition 15.1.1 *Soit $n \in \mathbb{N}^*$. La tribu $\mathcal{B}(\mathbb{R}^n)$ est engendrée par l'ensemble des pavés ouverts à extrémités rationnelles, c'est-à-dire par :*

$$\mathcal{E} = \left\{ \prod_{i=1}^n]a_i, b_i[\text{ avec } (a_1, \dots, a_n, b_1, \dots, b_n) \in \mathbb{Q}^{2n} \right\}$$

Démonstration: Puisque ces pavés sont ouverts, on a $\sigma(\mathcal{E}) \subset \mathcal{B}(\mathbb{R}^n)$. Prouvons l'inclusion inverse.

Si nous munissons \mathbb{R}^n de la norme infinie ($\|(x_1, \dots, x_n)\|_\infty = \sup_{1 \leq i \leq n} |x_i|$), nous constatons que les boules ouvertes sont des pavés ouverts. On conclut alors en utilisant le même schéma de démonstration que pour le premier point de la proposition 13.1.2

□

Remarque. On montre facilement que la tribu $\mathcal{B}(\mathbb{R}^n)$ est aussi engendrée par l'ensemble des pavés fermés à extrémités rationnelles, c'est-à-dire par :

$$\mathcal{E}' = \left\{ \prod_{i=1}^n [a_i, b_i] \text{ avec } (a_1, \dots, a_n, b_1, \dots, b_n) \in \mathbb{Q}^{2n} \right\}$$

De façon similaire à ce qui a été fait dans le chapitre précédent à propos des variables aléatoires réelles, nous définissons les vecteurs aléatoires à valeurs dans \mathbb{R}^n comme suit.

Définition 15.1.2 *On appelle variable aléatoire à valeurs dans \mathbb{R}^n (ou vecteur aléatoire à valeurs dans \mathbb{R}^n , ou vecteur aléatoire de dimension n) défini sur (Ω, \mathcal{A}, P) toute application $X : \Omega \rightarrow \mathbb{R}^n$ vérifiant :*

$$\forall B \in \mathcal{B}(\mathbb{R}^n) \quad X^{-1}(B) \in \mathcal{A}. \quad (15.1)$$

Un résultat général de la théorie de la mesure permet de démontrer la proposition suivante, que nous admettrons.

Proposition 15.1.3 *Soit \mathcal{E} un ensemble de parties de \mathbb{R}^n qui engendre la tribu $\mathcal{B}(\mathbb{R}^n)$. Alors une application $X : \Omega \rightarrow \mathbb{R}^n$ est un vecteur aléatoire n -dimensionnel sur (Ω, \mathcal{A}, P) si et seulement si la condition suivante est satisfaite :*

$$\forall B \in \mathcal{E} \quad X^{-1}(B) \in \mathcal{A}.$$

Cette proposition admet un corollaire important, qui sera notre outil principal pour démontrer qu'une application $X : \Omega \rightarrow \mathbb{R}^n$ est un vecteur aléatoire défini sur (Ω, \mathcal{A}, P) . Avant de l'énoncer, nous introduisons quelques notations. Pour tout $1 \leq i \leq n$, nous appellerons $\phi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ la i -ème projection canonique, définie par :

$$\forall x = (x_1, \dots, x_n) \in \mathbb{R}^n \quad \phi_i(x) = x_i.$$

Pour tout $1 \leq i \leq n$, nous considérons alors l'application $X_i = \phi_i \circ X : \Omega \rightarrow \mathbb{R}$ de sorte que nous pouvons écrire l'application X à valeurs dans \mathbb{R}^n sous la forme $X = (X_1, \dots, X_n)$.

Corollaire 15.1.4 *L'application $X : \Omega \rightarrow \mathbb{R}^n$ est un vecteur aléatoire de dimension n sur (Ω, \mathcal{A}, P) si et seulement si pour tout $1 \leq i \leq n$, l'application $X_i = \phi_i \circ X : \Omega \rightarrow \mathbb{R}$ est une variable aléatoire réelle sur (Ω, \mathcal{A}, P) .*

Démonstration: Supposons que X est un vecteur aléatoire n -dimensionnel et montrons que pour tout $1 \leq i \leq n$ et tout intervalle réel $]a, b[$, nous avons $X_i^{-1}(]a, b[) \in \mathcal{A}$, ce qui suffira à établir le sens direct de notre corollaire d'après la proposition 13.2.2. Or nous avons l'égalité :

$$X_i^{-1}(]a, b[) = X^{-1}(\mathbb{R} \times \dots \underbrace{]a, b[}_{i\text{-ème position}} \times \dots \times \mathbb{R}) \in \mathcal{A}$$

puisque $\mathbb{R} \times \dots]a, b[\times \dots \times \mathbb{R}$ est un ouvert donc un borélien de \mathbb{R}^n .

Passons au sens réciproque. D'après la proposition précédente, il suffit de vérifier que pour tout pavé ouvert $\prod_{i=1}^n]a_i, b_i[$, nous avons :

$$X^{-1}\left(\prod_{i=1}^n]a_i, b_i[\right) \in \mathcal{A}.$$

Or ceci résulte de l'égalité :

$$X^{-1}\left(\prod_{i=1}^n]a_i, b_i[\right) = \bigcap_{i=1}^n \underbrace{X_i^{-1}(]a_i, b_i[)}_{\in \mathcal{A}}.$$

□

Remarque. Autrement dit, un vecteur aléatoire de dimension n n'est autre qu'un n -uplet dont toutes les composantes sont des variables aléatoires réelles. Une conséquence immédiate de ce corollaire est que l'ensemble des vecteurs aléatoires n -dimensionnels définis sur (Ω, \mathcal{A}, P) est un \mathbb{R} -espace vectoriel.

Nous continuons à généraliser ce qui a été fait dans le chapitre précédent à propos des variables aléatoires réelles, en définissant la loi d'un vecteur aléatoire à valeurs dans \mathbb{R}^n comme suit.

Théorème et définition 15.1.5 *Considérons un espace probabilisé (Ω, \mathcal{A}, P) et un vecteur aléatoire n -dimensionnel $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Alors l'application $P_X : \mathcal{B}(\mathbb{R}^n) \rightarrow [0, 1]$ définie par :*

$$\forall B \in \mathcal{B}(\mathbb{R}^n) \quad P_X(B) = P(X \in B)$$

est une mesure de probabilité sur l'espace mesurable $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, appelée loi de probabilité (ou simplement loi) du vecteur aléatoire X .

Démonstration: Tout-à-fait similaire à celle qui a été faite dans le cas des variables aléatoires réelles. \square

15.2 Vecteurs aléatoires discrets

Définition 15.2.1 *Un vecteur aléatoire n -dimensionnel $X = (X_1, \dots, X_n)$ est dit discret si toutes ses composantes X_1, \dots, X_n sont des variables aléatoires réelles discrètes.*

Pour tout $1 \leq i \leq n$, nous notons $E_i = X_i(\Omega)$ qui est donc un sous-ensemble réel fini ou dénombrable. Notons qu'alors $X(\Omega) \subset E_1 \times \dots \times E_n$ est un sous-ensemble de \mathbb{R}^n également fini ou dénombrable. Le lecteur pourra montrer que connaître la loi d'un vecteur aléatoire discret équivaut à connaître toutes les quantités suivantes :

$$P(X_1 = x_1, \dots, X_n = x_n), \quad x_1 \in E_1, \dots, x_n \in E_n.$$

Nous renvoyons le lecteur à [MEY] IV.3, IV.5, V et VII pour les notions liées aux vecteurs aléatoires discrets.

15.3 Vecteurs aléatoires à densité

Dans cette section, nous écrirons les énoncés dans le cas de la dimension 2 pour alléger les écritures mais tout se généralise en dimension n . Un vecteur aléatoire de dimension 2 est encore appelé *couple aléatoire*.

Définition 15.3.1 On appelle densité de probabilité sur \mathbb{R}^2 toute application $f : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ intégrable et telle que :

$$\int_{\mathbb{R}^2} f(x, y) \, dx \, dy = 1.$$

Exemple. Si $D \subset \mathbb{R}^2$ admet une aire non nulle et finie, alors l'application $f : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ définie par :

$$\forall (x, y) \in \mathbb{R}^2 \quad f(x, y) = \frac{1}{\text{aire}(D)} \mathbf{1}_D(x, y)$$

est une densité de probabilité appelée densité uniforme sur D .

Définition 15.3.2 On dit qu'un couple aléatoire (X, Y) admet f pour densité si pour tous intervalles réels I et J , on a l'égalité :

$$P(X \in I, Y \in J) = \int_{I \times J} f(x, y) \, dx \, dy.$$

Remarques.

1. L'égalité précédente s'écrit évidemment de façon équivalente :

$$P((X, Y) \in I \times J) = \int_{I \times J} f(x, y) \, dx \, dy.$$

Un résultat de théorie de la mesure nous dit alors que

$$P((X, Y) \in B) = \int_B f(x, y) \, dx \, dy,$$

pour tout sous-ensemble $B \subset \mathbb{R}^2$ pour lequel l'intégrale du membre de droite a un sens (par exemple un *domaine simple* du plan).

2. Il n'est pas nécessaire de vérifier à l'avance que f est une densité de probabilité sur \mathbb{R}^2 mais seulement que f est une application positive et intégrable sur \mathbb{R}^2 . En effet, en appliquant l'égalité précédente avec $I = J = \mathbb{R}$, on en déduit que f est bien une densité de probabilité.

Théorème 15.3.3 (de transfert) Soit (X, Y) un couple aléatoire de densité f et $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ une application telle que $|\psi|f$ soit intégrable sur \mathbb{R}^2 . Alors $\psi(X, Y)$ est une variable aléatoire réelle qui admet pour espérance :

$$E[\psi(X, Y)] = \int_{\mathbb{R}^2} \psi(x, y) f(x, y) \, dx \, dy.$$

Conformément au programme, nous admettons ce théorème. Remarquons simplement que $\psi(X, Y)$ n'est pas nécessairement une variable à densité, comme nous le voyons en prenant $\psi \equiv 0$.

En fait, on peut raffiner le théorème précédent en énonçant une équivalence.

Proposition 15.3.4 *Soit (X, Y) un couple aléatoire et $f : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ une application intégrable sur \mathbb{R}^2 . Alors, il y a équivalence entre les propositions :*

1. *L'application f est une densité de probabilité sur \mathbb{R}^2 et le couple aléatoire (X, Y) admet f pour densité.*
2. *Pour toute application $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ telle que $|\psi|f$ soit intégrable sur \mathbb{R}^2 , on a l'égalité :*

$$E[\psi(X, Y)] = \int_{\mathbb{R}^2} \psi(x, y) f(x, y) dx dy.$$

Démonstration: L'implication $1 \Rightarrow 2$ n'est autre que le théorème de transfert. Montrons donc $2 \Rightarrow 1$.

En appliquant une première fois notre égalité avec $\psi \equiv 1$, nous obtenons que f est bien une densité de probabilité sur \mathbb{R}^2 .

Pour tous intervalles réels I et J , nous appliquons maintenant l'égalité avec $\psi = \mathbf{1}_{I \times J}$. Remarquons d'abord que

$$E[\mathbf{1}_{I \times J}(X, Y)] = E[\mathbf{1}_{\{X \in I, Y \in J\}}] = P(X \in I, Y \in J).$$

Nous obtenons donc :

$$P(X \in I, Y \in J) = \int_{I \times J} f(x, y) dx dy,$$

d'où la conclusion. □

En utilisant le théorème du changement de variable dans des intégrales doubles, on prouve [FOA 197-198] le résultat suivant qui est très utile pour effectuer des calculs pratiques de densités de couples aléatoires.

Proposition 15.3.5 *Soit Φ un C^1 -difféomorphisme de U sur V , où U et V sont deux ouverts de \mathbb{R}^2 . Considérons un couple aléatoire (X, Y) admettant une densité f nulle en dehors de U (donc $P[(X, Y) \in U] = 1$) et définissons un nouveau couple aléatoire en posant $(S, T) = \Phi(X, Y)$.*

Alors (S, T) admet pour densité l'application g définie par :

$$\forall (s, t) \in \mathbb{R}^2, \quad g(s, t) = f \circ \Phi^{-1}(s, t) |J_{\Phi^{-1}}(s, t)| \mathbf{1}_V(s, t),$$

où $J_{\Phi^{-1}}(s, t)$ désigne le déterminant jacobien du difféomorphisme Φ^{-1} .

Le lecteur, après avoir lu la section suivante sur l'indépendance de variables aléatoires, pourra utiliser ce corollaire pour résoudre les exercices 15.6.4 et 15.6.5.

Nous terminons cette section par un résultat sur l'existence de densités pour les v.a.r. coordonnées d'un vecteur aléatoire admettant une densité.

Proposition 15.3.6 *Si le couple aléatoire (X, Y) admet une densité f sur \mathbb{R}^2 , alors les variables aléatoires réelles X et Y admettent respectivement des densités de probabilité sur \mathbb{R} données par :*

$$f_X(x) = \int_{\mathbb{R}} f(x, y) dy \quad , \quad f_Y(y) = \int_{\mathbb{R}} f(x, y) dx,$$

dès lors que ces applications f_X et f_Y sont continues par morceaux. On les appelle les densités marginales.

Démonstration: Nous faisons la démonstration pour la variable X ; la démonstration pour Y s'en déduit par symétrie. Soit I un intervalle réel quelconque; nous avons alors les égalités :

$$P(X \in I) = P((X, Y) \in I \times \mathbb{R}) = \int_{I \times \mathbb{R}} f(x, y) dx dy.$$

Si nous définissons l'application $f_X : \mathbb{R} \rightarrow \mathbb{R}_+$ par :

$$\forall x \in \mathbb{R} \quad f_X(x) = \int_{\mathbb{R}} f(x, y) dy,$$

alors le théorème de Fubini-Tonelli implique que f_X est intégrable sur \mathbb{R} et

$$\int_{I \times \mathbb{R}} f(x, y) dx dy = \int_I f_X(x) dx.$$

Finalement, nous avons bien montré que pour tout intervalle réel I ,

$$P(X \in I) = \int_I f_X(x) dx.$$

□

Remarque. Ainsi, si un vecteur aléatoire admet une densité, alors chacune de ses composantes est une variable aléatoire à densité. Il est à noter que la réciproque est fautive. En effet, soit X une variable aléatoire réelle à densité (nous pouvons prendre $X \sim \mathcal{N}(0, 1)$ par exemple). Nous allons prouver que le couple aléatoire (X, X) n'admet aucune densité, ce qui nous fournira un contre-exemple.

Notons $\Delta = \{(x, y) \in \mathbb{R}^2, x = y\}$ la première diagonale du plan. Nous avons bien sûr $P[(X, X) \in \Delta] = 1$.

Raisonnons par l'absurde en supposant que le couple aléatoire (X, X) admet une densité f . Nous avons alors :

$$P[(X, X) \in \Delta] = \int_{\Delta} f(x, y) dx dy = \int_{\mathbb{R}^2} \mathbf{1}_{x=y} f(x, y) dx dy.$$

D'après le théorème de Fubini-Tonelli, cette dernière intégrale vaut :

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbf{1}_{x=y} f(x, y) dx \right) dy = \int_{\mathbb{R}} \left(\int_y^y f(x, y) dx \right) dy = 0,$$

d'où une contradiction. □

15.4 Indépendance de p v.a.r.

Dans toute cette section, on fixe un entier $p \geq 2$.

Définition 15.4.1 On dit que p variables aléatoires réelles X_1, \dots, X_p sont indépendantes (sous-entendu dans leur ensemble) si pour tous intervalles réels I_1, \dots, I_p , on a l'égalité :

$$P(X_1 \in I_1, \dots, X_p \in I_p) = P(X_1 \in I_1) \cdots P(X_p \in I_p)$$

Un instant de réflexion permet de constater que cette condition équivaut à : Pour tous interv. réels I_1, \dots, I_p , les événements $\{X_1 \in I_1\}, \dots, \{X_p \in I_p\}$ sont indépendants dans leur ensemble.

Proposition 15.4.2 On considère p variables aléatoires réelles X_1, \dots, X_p indépendantes et admettant des densités respectives f_{X_1}, \dots, f_{X_p} . Alors le vecteur aléatoire (X_1, \dots, X_p) admet pour densité l'application $f : \mathbb{R}^p \rightarrow \mathbb{R}_+$ définie par :

$$\forall (x_1, \dots, x_p) \in \mathbb{R}^p \quad f(x_1, \dots, x_p) = f_{X_1}(x_1) \cdots f_{X_p}(x_p)$$

Remarque. L'existence de cette densité est déjà un résultat intéressant en soi. En outre, sous nos hypothèses, nous obtenons une formule explicite de cette densité.

Démonstration: Pour tous intervalles réels I_1, \dots, I_p , on a l'égalité :

$$P(X_1 \in I_1, \dots, X_p \in I_p) = \prod_{i=1}^p P(X_i \in I_i) = \prod_{i=1}^p \int_{I_i} f_{X_i}(x_i) dx_i$$

Comme les applications $f_{X_i}, 1 \leq i \leq p$ sont toutes positives, le théorème de Fubini nous permet d'en déduire :

$$P((X_1, \dots, X_p) \in I_1 \times \dots \times I_p) = \int_{I_1 \times \dots \times I_p} f(x_1, \dots, x_p) dx_1 \dots dx_p,$$

où l'application f est définie comme dans l'énoncé, d'où la conclusion. \square

Réciproquement, nous avons le résultat suivant :

Proposition 15.4.3 *Supposons que le vecteur aléatoire (X_1, \dots, X_p) admet pour densité une application $f : \mathbb{R}^p \rightarrow \mathbb{R}_+$ à variables séparables, c'est-à-dire vérifiant*

$$\forall (x_1, \dots, x_p) \in \mathbb{R}^p, \quad f(x_1, \dots, x_p) = f_1(x_1) \cdots f_p(x_p),$$

où les applications $f_i : \mathbb{R} \rightarrow \mathbb{R}_+$ sont intégrables. Alors :

1. *Les variables aléatoires réelles X_1, \dots, X_p sont indépendantes.*
2. *Pour tout $1 \leq i \leq p$, la variable aléatoire réelle X_i admet une densité de la forme*

$$f_{X_i}(x_i) = c_i f_i(x_i),$$

où les c_i sont des constantes positives telles que $\prod_{i=1}^p c_i = 1$.

Démonstration: Dans un souci d'alléger les notations, nous l'écrirons dans le cas $p = 2$ mais elle se généralise immédiatement.

D'après la proposition 15.3.6, la variable aléatoire réelle X_1 admet pour densité

$$f_{X_1}(x_1) = \int_{\mathbb{R}} f(x_1, x_2) dx_2 = \int_{\mathbb{R}} f_1(x_1) f_2(x_2) dx_2 = c_1 f_1(x_1),$$

où l'on a posé $c_1 := \int_{\mathbb{R}} f_2(x_2) dx_2 \in \mathbb{R}_+$ puisque f_2 a été supposée intégrable. Par un raisonnement symétrique, on montre que X_2 admet une densité de la forme $f_{X_2}(x_2) = c_2 f_2(x_2)$.

D'une part, nous avons :

$$\int_{\mathbb{R}^2} f_{X_1}(x_1) f_{X_2}(x_2) dx_1 dx_2 = \left(\int_{\mathbb{R}} f_{X_1}(x_1) dx_1 \right) \left(\int_{\mathbb{R}} f_{X_2}(x_2) dx_2 \right) = 1$$

puisque f_{X_1} et f_{X_2} sont des densités de probabilité sur \mathbb{R} .

D'autre part, nous avons, en utilisant les égalités précédentes :

$$\int_{\mathbb{R}^2} f_{X_1}(x_1) f_{X_2}(x_2) dx_1 dx_2 = c_1 c_2 \int_{\mathbb{R}^2} f_1(x_1) f_2(x_2) dx_1 dx_2 = c_1 c_2$$

puisque $f(x_1, x_2) = f_1(x_1) f_2(x_2)$ est une densité de probabilité sur \mathbb{R}^2 .

Nous en déduisons l'égalité $c_1 c_2 = 1$. Notons qu'elle a pour conséquence :

$$\forall (x_1, x_2) \in \mathbb{R}^2 \quad f(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) \quad (15.2)$$

Il nous reste à démontrer l'indépendance des v.a.r. X_1 et X_2 .

Pour tous intervalles réels I_1, I_2 , nous avons l'égalité :

$$P(X_1 \in I_1, X_2 \in I_2) = P((X_1, X_2) \in I_1 \times I_2) = \int_{I_1 \times I_2} f(x_1, x_2) dx_1 dx_2$$

D'après l'égalité (15.2), nous avons donc :

$$P(X_1 \in I_1, X_2 \in I_2) = \int_{I_1 \times I_2} f_{X_1}(x_1) f_{X_2}(x_2) dx_1 dx_2$$

Comme les applications f_{X_1} et f_{X_2} sont positives, le théorème de Fubini nous permet d'en déduire :

$$P(X_1 \in I_1, X_2 \in I_2) = \left(\int_{I_1} f_{X_1}(x_1) dx_1 \right) \left(\int_{I_2} f_{X_2}(x_2) dx_2 \right) = P(X_1 \in I_1) P(X_2 \in I_2)$$

d'où la conclusion. \square

15.5 Produit de convolution de densités

Définition 15.5.1 Soient f et g des densités de probabilité sur \mathbb{R} . On appelle produit de convolution des densités f et g , noté $f * g$, l'application de \mathbb{R} dans \mathbb{R}_+ définie comme suit :

$$\forall x \in \mathbb{R}, \quad f * g(x) = \int_{-\infty}^{+\infty} f(x-y) g(y) dy.$$

Proposition 15.5.2 Soient X et Y deux variables aléatoires réelles indépendantes admettant des densités respectives f et g . Si $f * g \in C_M(\mathbb{R}, \mathbb{R})$, alors la variable aléatoire réelle $X + Y$ admet pour densité $f * g$.

Démonstration: D'après la Proposition 15.4.2, le couple aléatoire (X, Y) admet pour densité $f(x, y) = f_X(x)f_Y(y)$ sur \mathbb{R}^2 .

Appliquons la Proposition 15.3.5 avec $U = V = \mathbb{R}^2$ et $\Phi(x, y) = (x + y, y)$ si bien que $\Phi^{-1}(s, t) = (s - t, t)$ et $J_{\Phi^{-1}}(s, t) = 1$.

Nous en déduisons que le couple aléatoire $(X + Y, Y)$ admet la densité $g(s, t) = f(s - t, t) = f_X(s - t)f_Y(t)$ sur \mathbb{R}^2 .

La Proposition 15.3.6 nous permet alors de conclure en calculant la première densité marginale de ce couple aléatoire. \square

Remarque. Le lecteur trouvera dans [OUV¹ 203-204] une démonstration alternative, bien adaptée au programme officiel.

Proposition 15.5.3 Soient X et Y deux v.a.r. indépendantes de lois respectives $\gamma(a, \lambda)$ et $\gamma(b, \lambda)$ avec $(a, b, \lambda) \in (\mathbb{R}_+^*)^3$. Alors, $X + Y \sim \gamma(a + b, \lambda)$.

Démonstration: Notons comme suit la densité de la loi $\gamma(a, \lambda)$:

$$g_{a,\lambda}(x) = \frac{\lambda^a}{\Gamma(a)} e^{-\lambda x} x^{a-1} \mathbf{1}_{\mathbb{R}_+^*}(x).$$

Nous allons montrer que :

$$\boxed{g_{a,\lambda} * g_{b,\lambda} = g_{a+b,\lambda}} \quad (15.3)$$

Ceci nous permettra de conclure puisque la variable aléatoire $X + Y$ admet alors $g_{a+b,\lambda}$ pour densité.

Par définition du produit de convolution, nous avons :

$$\forall x \in \mathbb{R}, \quad g_{a,\lambda} * g_{b,\lambda}(x) = \int_{-\infty}^{+\infty} g_{a,\lambda}(x - y) g_{b,\lambda}(y) dy$$

En tenant compte des indicatrices de \mathbb{R}_+^* intervenant dans les densités, nous constatons que cette intégrale est nulle pour $x \leq 0$ et que

$$\forall x \in \mathbb{R}_+^*, \quad g_{a,\lambda} * g_{b,\lambda}(x) = \int_0^x \frac{\lambda^a}{\Gamma(a)} e^{-\lambda(x-y)} (x-y)^{a-1} \frac{\lambda^b}{\Gamma(b)} e^{-\lambda y} y^{b-1} dy$$

Ainsi, nous avons :

$$g_{a,\lambda} * g_{b,\lambda}(x) = \begin{cases} \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} e^{-\lambda x} \int_0^x (x-y)^{a-1} y^{b-1} dy & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases} \quad (15.4)$$

Arrêtons nous un instant sur cette dernière intégrale. En posant le changement de variable $y = xt \iff t = \frac{y}{x}$, nous obtenons :

$$\int_0^x (x-y)^{a-1} y^{b-1} dy = \int_0^1 x^{a-1} (1-t)^{a-1} x^{b-1} t^{b-1} x dt = x^{a+b-1} B(b, a)$$

avec, par définition de la fonction Bêta, $B(b, a) = \int_0^1 t^{b-1} (1-t)^{a-1} dt$. On sait que $B(b, a) = B(a, b)$ (il suffit de poser le changement de variable $s = 1-t$) donc l'égalité (15.4) se réécrit finalement :

$$\forall x \in \mathbb{R}, \quad g_{a,\lambda} * g_{b,\lambda}(x) = \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} e^{-\lambda x} x^{a+b-1} B(a, b) \mathbf{1}_{\mathbb{R}_+^*}(x)$$

Puisque $g_{a,\lambda} * g_{b,\lambda} \in C_M(\mathbb{R}, \mathbb{R})$, la Proposition 15.5.2 nous dit que cette application est une densité de probabilité. D'autre part, nous constatons que $g_{a,\lambda} * g_{b,\lambda}$ est proportionnelle à la densité $g_{a+b,\lambda}$. Nous en déduisons que ces densités sont égales : nous obtenons donc bien l'égalité (15.3). En outre, la constante de proportionnalité valant à la fois 1 et $\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} B(a, b)$, nous avons « en bonus » l'égalité : $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$. \square

Application 15.5.4 (Loi d'Erlang) Si X_1, \dots, X_n sont indépendantes de même loi $\mathcal{E}(\lambda)$, alors $X_1 + \dots + X_n$ admet la densité :

$$\frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} \mathbf{1}_{\mathbb{R}_+^*}(x)$$

(Cette densité définit la loi d'Erlang de paramètres n et λ : elle a été introduite historiquement pour étudier le nombre de clients contactant simultanément un centre d'appels téléphoniques.)

Démonstration: On vérifie que $\gamma(1, \lambda) = \mathcal{E}(\lambda)$ et la proposition précédente nous permet d'en déduire que $[\mathcal{E}(\lambda)]^{*n} = \gamma(n, \lambda)$ qui a bien la densité annoncée. \square

Application 15.5.5 (Loi $\chi^2(n)$) Si X_1, \dots, X_n sont indépendantes de même loi $\mathcal{N}(0, 1)$, alors $X_1^2 + \dots + X_n^2$ admet la densité :

$$\frac{1}{2^{n/2} \Gamma\left(\frac{n}{2}\right)} x^{n/2-1} e^{-x/2} \mathbf{1}_{\mathbb{R}_+^*}(x)$$

(Cette densité définit la loi du Khi-deux à n degrés de liberté, notée $\chi^2(n)$, qui permet de construire un test statistique très utilisé.)

Démonstration: Si $Y \sim \mathcal{N}(0, 1)$, alors $Y^2 \sim \gamma\left(\frac{1}{2}, \frac{1}{2}\right)$: cela se vérifie en calculant la fonction de répartition de Y^2 .

Comme les v.a.r. X_1^2, \dots, X_n^2 sont indépendantes, leur somme suit la loi $[\gamma\left(\frac{1}{2}, \frac{1}{2}\right)]^{*n} = \gamma\left(\frac{n}{2}, \frac{1}{2}\right)$ dont on reconnaît la densité. \square

15.6 Exercices

15.6.1 Calcul de densité marginale.

Soit (X, Y) un couple aléatoire à valeurs dans \mathbb{R}^2 , admettant pour densité :

$$p(x, y) = \frac{1}{\pi} \frac{ye^{-y}}{x^2 + y^2} \mathbf{1}_{\mathbb{R}_+^*}(y).$$

Vérifier que p est bien une densité sur \mathbb{R}^2 et calculer la densité marginale de la variable aléatoire réelle Y .

15.6.2 Densité gaussienne en dimension 2.

Soit (X, Y) un vecteur aléatoire à valeurs dans \mathbb{R}^2 de densité :

$$p(x, y) = \frac{1}{2\pi\sigma\tau\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma^2} - 2\rho\frac{xy}{\sigma\tau} + \frac{y^2}{\tau^2}\right)\right\},$$

où σ et τ sont deux réels positifs strictement et $|\rho| < 1$.

Montrer, par un calcul de densités marginales, que :

$$X \sim \mathcal{N}(0, \sigma^2) \text{ et } Y \sim \mathcal{N}(0, \tau^2).$$

15.6.3 Loi uniforme sur le disque

[OUV¹ 223]

Soit (X, Y) un couple aléatoire qui suit la loi uniforme sur le disque $D(0, 1) = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 \leq 1\}$. Calculer les densités marginales de X et Y et en déduire que ces deux variables aléatoires réelles ne sont pas indépendantes.

Remarque. Avec la notion de covariance qui apparaîtra dans la suite du cours, il est facile de montrer que $\text{Cov}(X, Y) = 0$: on dit que les variables X et Y sont décorrélées. Cet exercice nous fournit donc un exemple de variables décorrélées mais non indépendantes.

15.6.4 Gauss et Cauchy.

[FOA 201]

On considère deux v.a.r. indépendantes S et T de même loi $\mathcal{N}(0, 1)$.

Montrer que le couple $(\frac{S}{T}, T)$ admet une densité que l'on calculera.

En déduire, par calcul de densité marginale, que $\frac{S}{T}$ suit une loi de Cauchy de paramètre 1.

15.6.5 Avec des lois exponentielles.

[OUV² 75-77]

Soient $X \sim \text{Exp}(\lambda)$ et $Y \sim \text{Exp}(\mu)$ deux variables aléatoires réelles indépendantes ($\lambda > 0, \mu > 0$).

On définit les variables aléatoires réelles $U = \inf(X, Y)$ et $V = |X - Y|$.

Montrer que U et V sont indépendantes.

15.6.6 Méthode de Box-Muller

Soient X et Y deux variables aléatoires réelles indépendantes de même loi uniforme sur $]0, 1[$. Nous définissons :

$$(S, T) = (\sqrt{-2 \ln X} \cos(2\pi Y), \sqrt{-2 \ln X} \sin(2\pi Y))$$

Montrer que S et T sont deux variables indépendantes de même loi $\mathcal{N}(0, 1)$.

Corrigé. Montrons que nous définissons bien un C^1 -difféomorphisme de l'ouvert $U =]0, 1[^2$ sur l'ouvert $V = \mathbb{R}^2 \setminus (\mathbb{R}_+ \times \{0\})$ en posant :

$$\Phi(x, y) = (\sqrt{-2 \ln x} \cos(2\pi y), \sqrt{-2 \ln x} \sin(2\pi y))$$

En remarquant que $\rho(x) = \sqrt{-2 \ln x}$ et $\theta(y) = 2\pi y$ ne sont autre que les coordonnées polaires du point $\Phi(x, y)$, le lecteur constatera que Φ est bien une bijection de U sur V .

En outre, il est immédiat de vérifier que Φ est de classe C^1 . Il nous reste donc à prouver que Φ^{-1} est bien de classe C^1 mais nous rencontrons une difficulté pour expliciter cette bijection réciproque. En effet, si nous posons $\Phi(x, y) = (s, t) \in V$, nous calculons facilement

$$x = \exp\left(-\frac{s^2 + t^2}{2}\right)$$

mais il n'est pas possible de donner une formule pour y , sauf à « couper en morceaux » l'ensemble V . Nous préférons procéder indirectement en utilisant

un théorème de calcul différentiel.

Pour ce faire, nous calculons d'abord le déterminant jacobien de Φ , et non de Φ^{-1} comme d'habitude :

$$J_{\Phi}(x, y) = \begin{vmatrix} -\frac{\cos(2\pi y)}{x\sqrt{-2\ln x}} & -2\pi\sqrt{-2\ln x}\sin(2\pi y) \\ -\frac{\sin(2\pi y)}{x\sqrt{-2\ln x}} & 2\pi\sqrt{-2\ln x}\cos(2\pi y) \end{vmatrix} = -\frac{2\pi}{x}$$

Ce déterminant jacobien ne s'annulant jamais sur U , le *théorème d'inversion globale* affirme que Φ^{-1} est bien de classe C^1 sur V .

En outre, nous avons l'égalité :

$$J_{\Phi^{-1}}(s, t) = \frac{1}{J_{\Phi}(x, y)} = -\frac{x}{2\pi} = -\frac{1}{2\pi} \exp\left(-\frac{s^2 + t^2}{2}\right)$$

Nous pouvons maintenant appliquer la Proposition 15.3.5 pour obtenir la densité du couple (S, T) :

$$\frac{1}{2\pi} \exp\left(-\frac{s^2 + t^2}{2}\right) \mathbf{1}_V(s, t)$$

Si nous ôtons $\mathbf{1}_V(s, t)$ de cette formule, nous ne modifions la densité que sur la demi-droite $\mathbb{R}_+ \times \{0\}$, qui est de « surface nulle » : cela ne change donc rien aux valeurs des intégrales doubles dans lesquelles cette fonction intervient et il est donc correct de dire que (S, T) admet pour densité :

$$g(s, t) = \frac{1}{2\pi} \exp\left(-\frac{s^2 + t^2}{2}\right)$$

Cette fonction étant à variables séparables, la Proposition 15.4.3 nous permet de conclure.

Chapitre 16

Loi des grands nombres.

En ce qui concerne les *théorèmes limites* (c'est-à-dire les théorèmes qui énoncent des résultats asymptotiques) en calcul des probabilités, le programme comprend la loi faible des grands nombres avec sa démonstration par l'inégalité de Bienaymé-Tchebychev, ainsi que deux théorèmes admis : la loi forte des grands nombres et le théorème central limite. Nous traiterons ce dernier théorème, encore appelé *théorème-limite central*, et ses conséquences dans le chapitre suivant, en nous concentrant dans le présent chapitre sur les deux énoncés (faible et fort) de loi des grands nombres.

Introduction Un célèbre fabricant de piles électriques affirme : “elles durent deux fois plus longtemps”. Pour déterminer si c'est de la publicité mensongère ou non, il faudrait connaître l'espérance de vie d'une pile de ce type et la comparer à celle d'une pile “ordinaire”. Mais comment procéder dans la pratique pour évaluer, au moins approximativement, la quantité abstraite $E[X]$, où X est la durée de vie d'une pile en fonctionnement continu, que nous modélisons par une variable aléatoire réelle ?

Une idée qui vient naturellement à l'esprit est de tester indépendamment les unes des autres N piles du même type et, en notant $X_i(\omega)$ la durée de vie de la i -ème pile observée à la fin de cette expérience, de faire l'approximation suivante pour N suffisamment grand :

$$\frac{X_1(\omega) + \cdots + X_N(\omega)}{N} \approx E[X].$$

Pour justifier cette démarche, nous aurons besoin de théorèmes limites en probabilité. Le premier, appelé *loi des grands nombres*, montre que l'on a bien la convergence suivante dans un sens à préciser :

$$\lim_{n \rightarrow \infty} \frac{X_1 + \cdots + X_n}{n} = E[X].$$

Le second, appelé théorème central limite, précisera à quelle vitesse la convergence a lieu dans la loi des grands nombres, ce qui a une importance énorme dans les applications pratiques (statistique, méthodes de Monte-Carlo...)

Notons que pour donner un sens mathématique rigoureux à la limite précédente, il ne suffit pas d'avoir un “grand nombre” de variables aléatoires réelles X_1, \dots, X_N modélisant les durées de vie des piles testées mais une suite $(X_i)_{i \in \mathbb{N}^*}$ de variables aléatoires réelles. Notre hypothèse de base est que les durées de vie des différentes piles sont indépendantes les unes des autres mais qu'elles ont toute le même comportement aléatoire, plus précisément la même loi de probabilité. Avant d'aborder les théorèmes limites, nous aurons donc besoin d'étudier quelques propriétés d'une suite $(X_i)_{i \in \mathbb{N}^*}$ de variables aléatoires réelles indépendantes ayant toutes même loi.

Dans toute la suite de ce chapitre, les variables aléatoires considérées sont définies sur un espace de probabilité (Ω, \mathcal{A}, P) .

16.1 Suite indépendante équirépartie.

Définition 16.1.1 Une suite de variables aléatoires réelles $(X_i)_{i \in \mathbb{N}^*}$ est dite indépendante si pour tout $k \in \mathbb{N}^*$ et pour tous indices $1 \leq i_1 < i_2 < \dots < i_k$, les variables aléatoires réelles X_{i_1}, \dots, X_{i_k} sont indépendantes.

Une suite de variables aléatoires réelles $(X_i)_{i \in \mathbb{N}^*}$ est dite identiquement distribuée (ou équirépartie) si toutes les variables aléatoires réelles $X_i, i \in \mathbb{N}^*$ ont même loi de probabilité.

Une suite $(X_i)_{i \in \mathbb{N}^*}$ possédant ces deux propriétés est donc dite indépendante identiquement distribuée (i.i.d. en abrégé).

Nous retrouverons cette hypothèse tout au long de ce chapitre. Elle correspond à une expérience répétée à l'infini dans des conditions identiques (bien sûr, dans la pratique, elle n'est répétée qu'un “grand nombre” de fois).

Par exemple, un jeu de pile ou face répété à l'infini est modélisé par une suite i.i.d. de variables de Bernoulli de paramètre p .

Proposition 16.1.2 Soit $(X_i)_{i \in \mathbb{N}^*}$ une suite i.i.d. et $f \in C(\mathbb{R}, \mathbb{R})$. Alors la suite $(f(X_i))_{i \in \mathbb{N}^*}$ est encore i.i.d.

Démonstration: Montrons d'abord que la suite $(f(X_i))_{i \in \mathbb{N}^*}$ est indépendante, c'est-à-dire que pour tout $k \in \mathbb{N}^*$ et tous indices $1 \leq i_1 < \dots < i_k$, les variables aléatoires réelles $f(X_{i_1}), \dots, f(X_{i_k})$ sont indépendantes.

Nous considérons donc des intervalles ouverts réels arbitraires I_1, \dots, I_k et nous calculons la probabilité suivante :

$$P[f(X_{i_1}) \in I_1, \dots, f(X_{i_k}) \in I_k] = P[X_{i_1} \in f^{-1}(I_1), \dots, X_{i_k} \in f^{-1}(I_k)].$$

Remarquons que les sous-ensembles réels $f^{-1}(I_1), \dots, f^{-1}(I_k)$ sont ouverts en tant qu'images réciproques d'ouverts par une application continue; ce sont donc des éléments de la tribu borélienne $\mathcal{B}(\mathbb{R})$. Or, par hypothèse, les variables X_{i_1}, \dots, X_{i_k} sont indépendantes, si bien que nous pouvons encore écrire la probabilité précédente sous forme du produit :

$$P[X_{i_1} \in f^{-1}(I_1)] \cdots P[X_{i_k} \in f^{-1}(I_k)] = P[f(X_{i_1}) \in I_1] \cdots P[f(X_{i_k}) \in I_k],$$

ce qui nous permet de conclure.

Passons à la démonstration de l'équidistribution de la suite $(f(X_i))_{i \in \mathbb{N}^*}$.

En notant P_{X_i} la loi de probabilité de la variable aléatoire réelle X_i , nous avons par hypothèse :

$$\forall i \in \mathbb{N}^* \quad P_{X_i} = P_{X_1}.$$

Pour tout $i \in \mathbb{N}^*$ et tout intervalle ouvert réel I , nous avons donc :

$$P[f(X_i) \in I] = P[X_i \in f^{-1}(I)] = P_{X_i}[f^{-1}(I)] = P_{X_1}[f^{-1}(I)] = P[f(X_1) \in I],$$

ce qui s'écrit encore :

$$P_{f(X_i)}(I) = P_{f(X_1)}(I).$$

Une loi de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ étant entièrement déterminée par ses valeurs sur les intervalles ouverts, nous avons ainsi démontré :

$$\forall i \in \mathbb{N}^* \quad P_{f(X_i)} = P_{f(X_1)}.$$

□

Remarque. Si nous avons considéré une suite $(f_i)_{i \in \mathbb{N}^*}$ d'applications continues de \mathbb{R} dans \mathbb{R} , alors la suite $(f_i(X_i))_{i \in \mathbb{N}^*}$ serait encore indépendante mais n'aurait aucune raison d'être identiquement distribuée. Le lecteur pourra adapter la première partie de la démonstration précédente à titre d'exercice. Il est très facile de trouver un contre-exemple en ce qui concerne l'équidistribution.

Notations : Dans la suite, nous noterons $L^1(\Omega, \mathcal{A}, P)$ l'ensemble des variables aléatoires réelles définies sur l'espace de probabilité (Ω, \mathcal{A}, P) et admettant une espérance.

De même, nous noterons $L^2(\Omega, \mathcal{A}, P)$ l'ensemble des variables aléatoires réelles X définies sur l'espace de probabilité (Ω, \mathcal{A}, P) et admettant un moment d'ordre deux, c'est-à-dire telles que X^2 admet une espérance.

Nous ferons l'abus de langage courant consistant à appeler *variable de variance finie* une variable admettant un moment d'ordre deux. C'est un abus

de langage parce que si X n'est pas une variable de variance finie, cela ne signifie pas qu'elle a une variance infinie mais que sa variance n'est pas définie du tout ! C'est le cas par exemple si X suit une loi de Cauchy.

On démontre l'inclusion $L^2(\Omega, \mathcal{A}, P) \subset L^1(\Omega, \mathcal{A}, P)$ en utilisant l'inégalité $|x| \leq x^2 + 1$ valable pour tout réel x .

Proposition 16.1.3 *Soit $(X_i)_{i \in \mathbb{N}^*}$ une suite identiquement distribuée (mais pas nécessairement indépendante).*

Alors nous avons l'équivalence suivante :

$$[\forall i \in \mathbb{N}^* \quad X_i \in L^1(\Omega, \mathcal{A}, P)] \iff X_1 \in L^1(\Omega, \mathcal{A}, P).$$

Nous dirons alors que la suite (X_i) est dans $L^1(\Omega, \mathcal{A}, P)$ et, dans ce cas, toutes les variables aléatoires réelles ont la même espérance $m = E[X_i]$, $i \in \mathbb{N}^$.*

De même, nous avons l'équivalence :

$$[\forall i \in \mathbb{N}^* \quad X_i \in L^2(\Omega, \mathcal{A}, P)] \iff X_1 \in L^2(\Omega, \mathcal{A}, P).$$

Nous dirons alors que la suite (X_i) est dans $L^2(\Omega, \mathcal{A}, P)$ et, dans ce cas, toutes les variables aléatoires réelles ont la même variance $\sigma^2 = \text{Var}X_i$, $i \in \mathbb{N}^$.*

Démonstration: Nous ne ferons la démonstration que dans le cas où la loi commune aux variables X_i admet une densité de probabilité g .

Le lecteur pourra à titre d'exercice faire la preuve dans le cas où cette loi commune est discrète. Nous admettrons cette proposition dans les autres cas. D'après le théorème de transfert, nous avons pour tout $i \in \mathbb{N}^*$:

$$E[|X_i|] = \int_{-\infty}^{+\infty} |x|g(x) dx = E[|X_1|],$$

ce qui nous donne la première équivalence.

En outre, si l'intégrale impropre précédente est convergente, ce qui équivaut à :

$$\forall i \in \mathbb{N}^* \quad X_i \in L^1(\Omega, \mathcal{A}, P),$$

le théorème de transfert nous donne encore :

$$E[X_i] = \int_{-\infty}^{+\infty} x g(x) dx = E[X_1]$$

et nous posons $m := E[X_1]$.

De façon similaire, nous obtenons la deuxième équivalence en écrivant, pour tout $i \in \mathbb{N}^*$:

$$E[X_i^2] = \int_{-\infty}^{+\infty} x^2 g(x) dx = E[X_1^2].$$

Si cette espérance est finie, en gardant la notation $m = E[X_i]$ (qui a un sens puisque $L^2(\Omega, \mathcal{A}, P) \subset L^1(\Omega, \mathcal{A}, P)$), nous avons pour tout $i \in \mathbb{N}^*$:

$$\text{Var}X_i = E[X_i^2] - m^2 = E[X_1^2] - m^2 = \text{Var}X_1 = \sigma^2.$$

□

16.2 Convergences de suites de variables aléatoires réelles

16.2.1 Événements négligeables, égalité presque sûre.

Définition 16.2.1 *Un événement $A \in \mathcal{A}$ est dit négligeable si $P(A) = 0$.*

En termes courants, un événement négligeable n'a donc "aucune chance de se produire", il est "extrêmement improbable" mais il ne faudrait pas en conclure pour autant qu'il est impossible : seul l'événement \emptyset est qualifié d'impossible. Pour illustrer cette distinction, prenons l'exemple d'un nombre réel tiré au hasard uniformément entre 0 et 1. Nous pouvons modéliser cette expérience en prenant $\Omega = [0, 1]$, $\mathcal{A} = \mathcal{B}([0, 1])$, qui est la plus petite tribu sur Ω contenant tous les sous-intervalles de $[0, 1]$, et en prenant pour P la *mesure de Lebesgue*, qui est (nous admettons son existence et son unicité) l'unique probabilité sur (Ω, \mathcal{A}) telle que :

$$\forall [a, b] \subset [0, 1] \quad P([a, b]) = b - a.$$

Soit $x \in [0, 1]$ arbitraire fixé. Nous constatons que l'événement $A = \{x\} = [x, x]$ est négligeable. Il n'est néanmoins pas impossible qu'un nombre tiré au hasard entre 0 et 1 soit égal à x .

Proposition 16.2.2 *Toute union finie ou dénombrable d'événements négligeables est encore un événement négligeable.*

Démonstration: Soit $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ une suite d'événements tels que :

$$\forall n \in \mathbb{N} \quad P(A_n) = 0.$$

Par propriété de sous- σ -additivité de la mesure de probabilité P (cf. formule (12.2) page 218), nous avons alors :

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} P(A_n) = 0,$$

d'où la conclusion.

□

Remarque. Cela n'a plus aucune raison d'être vrai pour une union quelconque. Ainsi, si nous reprenons l'exemple précédent, nous avons :

$$\Omega = \bigcup_{x \in [0,1]} \{x\},$$

qui n'est évidemment pas négligeable.

Définition 16.2.3 *Un événement A est dit presque sûr si l'événement complémentaire A^c est négligeable.*

Deux variables aléatoires réelles X et Y sont dites égales presque sûrement si l'événement $\{\omega \in \Omega, X(\omega) = Y(\omega)\}$ est presque sûr.

Un événement presque sûr est donc un événement de probabilité 1.

Deux variables aléatoires réelles sont égales presque sûrement si et seulement si l'événement $\{X \neq Y\}$ est négligeable. À titre d'exercice, le lecteur pourra en déduire que l'égalité presque sûre est une relation d'équivalence sur l'ensemble des variables aléatoires réelles définies sur (Ω, \mathcal{A}, P) .

16.2.2 Convergences en probabilité et presque sûre.

Nous définissons maintenant deux types de convergence concernant les suites de variables aléatoires réelles.

Définition 16.2.4 *Soient $(Y_n)_{n \in \mathbb{N}^*}$ et Y des variables aléatoires réelles définies sur un espace de probabilité (Ω, \mathcal{A}, P) .*

On dit que la suite (Y_n) converge presque sûrement vers Y et l'on note $Y_n \xrightarrow{p.s.} Y$ s'il existe un événement négligeable $N \in \mathcal{A}$ tel que :

$$\forall \omega \notin N \quad Y_n(\omega) \longrightarrow Y(\omega).$$

On dit que la suite (Y_n) converge en probabilité vers Y et l'on note $Y_n \xrightarrow[n \rightarrow +\infty]{(P)} Y$ si :

$$\forall \epsilon > 0 \quad P(|Y_n - Y| > \epsilon) \longrightarrow 0.$$

Proposition 16.2.5 *Une limite en probabilité est unique modulo l'égalité presque sûre.*

Démonstration: Supposons que la suite (Y_n) converge en probabilité à la fois vers Y et Z . Pour tout $\epsilon > 0$, nous écrivons alors l'inclusion (facile à vérifier en passant aux événements complémentaires et en appliquant l'inégalité triangulaire) :

$$\{|Y - Z| > \epsilon\} \subset \{|Y_n - Y| > \frac{\epsilon}{2}\} \cup \{|Y_n - Z| > \frac{\epsilon}{2}\}.$$

Par passage aux probabilités puis en faisant tendre n vers l'infini, nous en déduisons : $P(|Y - Z| > \epsilon) = 0$.

Comme $\epsilon > 0$ était arbitraire, nous en déduisons que pour tout $p \in \mathbb{N}^*$, $P(|Y - Z| > \frac{1}{p}) = 0$.

Nous concluons alors en écrivant :

$$P(Y \neq Z) = P\left(\lim \uparrow \left\{|Y - Z| > \frac{1}{p}\right\}\right) = \lim \uparrow P\left(|Y - Z| > \frac{1}{p}\right) = 0.$$

□

Il serait facile de démontrer directement qu'une limite au sens de la convergence presque sûre est unique modulo l'égalité presque sûre (vous pouvez le faire à titre d'exercice) mais cela va résulter de la proposition suivante.

Proposition 16.2.6 *Si (Y_n) converge presque sûrement vers Y , alors (Y_n) converge en probabilité vers Y .*

Démonstration: Le lecteur est invité à lire [DAN 452].

□

Remarque. La réciproque est fausse : le lecteur pourra s'en convaincre en étudiant le contre-exemple suivant.

Nous définissons une suite $(I_n)_{n \in \mathbb{N}^*}$ de sous-intervalles de $[0, 1]$ comme suit :

$$\forall k \in \mathbb{N} \quad \forall j \in \{0, 1, \dots, 2^k - 1\} \quad I_{2^k+j} = \left] \frac{j}{2^k}, \frac{j+1}{2^k} \right].$$

Nous prenons alors pour espace de probabilité $(\Omega, \mathcal{A}, P) = ([0, 1], \mathcal{B}([0, 1]), \lambda_{[0,1]})$ et nous définissons pour tout $n \in \mathbb{N}^*$ la variable aléatoire $Y_n = \mathbf{1}_{I_n}$.

Nous constatons alors que $Y_n \xrightarrow[n \rightarrow +\infty]{(P)} 0$ mais que la suite (Y_n) ne converge pas presque sûrement puisque :

$$\forall \omega \in \Omega \quad \limsup Y_n(\omega) = 1 \text{ et } \liminf Y_n(\omega) = 0.$$

16.3 Loi des grands nombres

Théorème 16.3.1 (Loi faible des grands nombres) Soit $(X_i)_{i \in \mathbb{N}^*}$ une suite indépendante dans $L^2(\Omega, \mathcal{A}, P)$. Nous supposons que les variables X_i , $i \in \mathbb{N}^*$ ont toutes même espérance, notée m , et même variance. Pour tout $n \in \mathbb{N}^*$, nous définissons la moyenne empirique d'ordre n :

$$\bar{X}_n = \frac{X_1 + \cdots + X_n}{n}.$$

Alors nous avons la convergence en probabilité suivante :

$$\bar{X}_n \xrightarrow[n \rightarrow +\infty]{(P)} m.$$

Démonstration: Pour tout $n \in \mathbb{N}^*$, définissons la variable $S_n = \sum_{i=1}^n X_i$ de sorte que $\bar{X}_n = S_n/n$. Comme $L^2(\Omega, \mathcal{A}, P)$ est un espace vectoriel, nous avons $S_n \in L^2(\Omega, \mathcal{A}, P)$ et nous calculons immédiatement $E[S_n] = nm$ puis, en utilisant l'indépendance des variables X_i ,

$$\text{Var} S_n = \sum_{i=1}^n \text{Var} X_i = n\sigma^2,$$

où l'on a noté σ^2 la variance commune aux variables X_i .

Nous en déduisons que, pour tout $n \in \mathbb{N}^*$, $E[\bar{X}_n] = m$ et $\text{Var} \bar{X}_n = \frac{\sigma^2}{n}$.

Pour tout $\epsilon > 0$, l'inégalité de Bienaymé-Tchebychev nous permet alors d'écrire :

$$P(|\bar{X}_n - m| > \epsilon) \leq \frac{1}{\epsilon^2} \text{Var} \bar{X}_n = \frac{\sigma^2}{n\epsilon^2} \rightarrow 0,$$

d'où la conclusion. □

Remarques.

1. La loi des grands nombres nous dit que, conformément à notre intuition, la moyenne empirique (observable après l'expérience) converge vers la moyenne probabiliste théorique, c'est-à-dire l'espérance.
2. Ce théorème s'applique en particulier à toute suite (X_i) i.i.d. dans $L^2(\Omega, \mathcal{A}, P)$. Au prix d'une démonstration plus difficile, on peut démontrer alors qu'il se généralise au cas d'une suite (X_i) i.i.d. dans $L^1(\Omega, \mathcal{A}, P)$. Cependant, le théorème suivant va nous conduire à une conclusion encore plus forte sous la même hypothèse d'indépendance et équidistribution dans $L^1(\Omega, \mathcal{A}, P)$.

Théorème 16.3.2 (Loi forte des grands nombres, Kolmogorov) *Soit $(X_i)_{i \in \mathbb{N}^*}$ une suite indépendante identiquement distribuée. Alors il y a équivalence entre les deux conditions suivantes :*

1. *La suite (\bar{X}_n) converge presque sûrement vers une certaine variable aléatoire X à valeurs dans \mathbb{R} .*
2. *La suite (X_i) est dans $L^1(\Omega, \mathcal{A}, P)$.*

En supposant ces conditions satisfaites et en notant m l'espérance commune aux variables X_i , nous avons alors $X = m$ presque sûrement, c'est-à-dire :

$$\bar{X}_n \xrightarrow{p.s.} m = E[X_i].$$

Ce théorème, que nous admettons, porte le nom de loi *forte* des grands nombres parce que la convergence presque sûre est plus forte que la convergence en probabilité qui apparaissait dans le théorème précédent.

En outre, il nous donne une condition *nécessaire et suffisante* de convergence presque sûre de la suite des moyennes empiriques.

Chapitre 17

Loi normale

17.1 La loi normale centrée réduite

Proposition 17.1.1 (Intégrale de Gauss)

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$

[GOURDON 163,335] démontre cette égalité par 3 méthodes différentes :

- dérivation d’une intégrale dépendant d’un paramètre
- inégalité de convexité, changement de variables et formule de Wallis
- passage en coordonnées polaires

□

Nous suivons maintenant la présentation de Foata et Fuchs dans leur ouvrage “Calcul des probabilités” chez Dunod, 2ème édition, pages 178 à 181.

Proposition et définition 17.1.2 *L’application $g : \mathbb{R} \rightarrow \mathbb{R}_+$ définie par :*

$$\forall x \in \mathbb{R} \quad g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

est une densité de probabilité. La loi de densité g est appelée loi de Gauss ou loi normale (centrée réduite) et notée $\mathcal{N}(0, 1)$.

Gauss a introduit cette loi en 1809 à propos d’un problème statistique d’estimation de paramètre.

Le graphe de la densité de Gauss g est une «courbe en cloche» aplatie (voir figure page ??). On le trace facilement en notant que f est paire, qu’elle

admet un maximum global en $x = 0$ égal à $\frac{1}{\sqrt{2\pi}} \simeq 0,399$ et que la courbe admet deux points d'inflexion en $x = -1$ et $x = 1$.

Si $X \sim \mathcal{N}(0, 1)$, on calcule immédiatement $E[X] = 0$ par le théorème de transfert en utilisant la parité de g et

$$\text{Var}X = E[X^2] = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$$

par intégration par parties. C'est pourquoi on parle de loi normale *centrée réduite*.

La fonction de répartition de la loi de Gauss est donnée par

$$\forall x \in \mathbb{R} \quad \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

On vérifie facilement que Φ est de classe C^∞ , strictement croissante et donc bijective de \mathbb{R} sur son image $]0, 1[$, telle que $\Phi(-x) = 1 - \Phi(x)$ pour tout $x \in \mathbb{R}$.

Son graphe est une «courbe en S» assez étalée, symétrique par rapport à $(0, 1/2)$, où elle admet un point d'inflexion et où la pente de sa tangente vaut $\frac{1}{\sqrt{2\pi}} \simeq 0,399$.

On n'a pas de formule plus explicite pour Φ mais cette fonction et sa bijection réciproque Φ^{-1} sont tabulées, notamment pour des applications statistiques. Une valeur que nous utiliserons dans la suite est $\Phi(1,96) \approx 0,975$.

Si $X \sim \mathcal{N}(0, 1)$, alors $P(|X| \leq x) = 2\Phi(x) - 1$ pour tout $x \geq 0$, de sorte que $P(|X| \leq 1,96) \approx 0,95$.

On peut démontrer l'équivalence suivante lorsque $x \rightarrow +\infty$:

$$1 - \Phi(x) \sim \frac{1}{x} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

Comme cela tend «vite» vers 0, on dit que la loi de Gauss a une queue de distribution peu épaisse.

Proposition 17.1.3 *La loi de Gauss admet des moments de tous les ordres. Si $X \sim \mathcal{N}(0, 1)$, nous avons, pour tout $n \in \mathbb{N}$:*

$$E[X^{2n+1}] = 0 \quad ; \quad E[X^{2n}] = \frac{(2n)!}{2^n n!}$$

Démonstration: Pour tout $k \in \mathbb{N}^*$, nous avons $|x|^k g(x) = o(1/x^2)$ quand $x \rightarrow \pm\infty$ donc X admet un moment d'ordre k donné par :

$$E[X^k] = \int_{-\infty}^{+\infty} x^k \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dt$$

Lorsque $k = 2n + 1$, cette intégrale est nulle pour des raisons de parité. On démontre par intégration par parties que $E[X^{2n}] = (2n - 1)E[X^{2n-2}]$ pour tout $n \in \mathbb{N}^*$, ce qui nous donne par récurrence la formule annoncée puisque $E[X^0] = 1$ par convention. □

Pour simuler numériquement la loi de Gauss, on utilise la proposition suivante, qui se démontre grâce à un changement de variable dans une intégrale double. [OUV² 67-68]

Proposition 17.1.4 (Méthode de Box-Muller) Soient U_1 et U_2 deux v.a.r. indépendantes, de même loi uniforme sur $[0, 1]$. Nous posons :

$$X := \sqrt{-2 \ln U_1} \cos(2\pi U_2) \quad , \quad Y := \sqrt{-2 \ln U_1} \sin(2\pi U_2).$$

Alors, les v.a.r. X et Y sont indépendantes et de même loi $\mathcal{N}(0, 1)$.

17.2 La loi normale générale

Proposition et définition 17.2.1 Soient $\mu \in \mathbb{R}$ et $\sigma > 0$ deux paramètres fixés. L'application $g_{\mu, \sigma^2} : \mathbb{R} \rightarrow \mathbb{R}_+$ définie par :

$$\forall x \in \mathbb{R} \quad g_{\mu, \sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

est une densité de probabilité. La loi de densité g_{μ, σ^2} est appelée loi normale de paramètres (μ, σ^2) et notée $\mathcal{N}(\mu, \sigma^2)$.

On vérifie que g_{μ, σ^2} est bien une densité en effectuant un changement de variable affine dans une intégrale simple, puis en utilisant la proposition 17.1.2. Notons que pour $\mu = 0$ et $\sigma^2 = 1$, on retrouve bien la densité de la loi $\mathcal{N}(0, 1)$ telle qu'elle a été définie dans la sous-section précédente. Nous allons maintenant préciser les rapports entre la loi normale générale et la loi normale centrée réduite.

Proposition 17.2.2 Soient $\mu \in \mathbb{R}$ et $\sigma > 0$ deux paramètres fixés. On considère deux variables aléatoires réelles X et Y telles que $Y = \mu + \sigma X$. Alors on a l'équivalence suivante :

$$X \sim \mathcal{N}(0, 1) \iff Y \sim \mathcal{N}(\mu, \sigma^2)$$

Démonstration: On vérifie facilement que cette équivalence résulte de la proposition 14.3.2. □

Pour simuler numériquement la loi $\mathcal{N}(\mu, \sigma^2)$, on applique d'abord la méthode de Box-Muller pour simuler $X \sim \mathcal{N}(0, 1)$ puis l'on calcule $Y = \mu + \sigma X$.

Corollaire 17.2.3 Soit $Y \sim \mathcal{N}(\mu, \sigma^2)$. Alors $E[Y] = \mu$ et $\text{Var}Y = \sigma^2$

Ce résultat justifie que le premier paramètre d'une loi normale soit traditionnellement noté μ ou m (comme «moyenne» probabiliste, c'est-à-dire espérance) et le second σ^2 puisqu'il est égal à la variance de la loi.

Corollaire 17.2.4 La fonction de répartition Φ_{μ, σ^2} de la loi $\mathcal{N}(\mu, \sigma^2)$ est donnée par :

$$\forall x \in \mathbb{R} \quad \Phi_{\mu, \sigma^2}(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

Démonstration: En reprenant les notations de la proposition 17.2.2, nous avons, pour tout $x \in \mathbb{R}$:

$$\Phi_{\mu, \sigma^2}(x) = P(Y \leq x) = P(\mu + \sigma X \leq x) = P\left(X \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

□

La proposition suivante énonce la stabilité de la loi normale par produit de convolution.

Proposition 17.2.5 Soient X et Y deux v.a.r. indépendantes de lois respectives $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_2, \sigma_2^2)$, avec $(m_1, m_2) \in \mathbb{R}^2$ et $(\sigma_1, \sigma_2) \in (\mathbb{R}_+^*)^2$. Alors la v.a.r. $X + Y$ suit la loi $\mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

Le résultat essentiel établi par cette proposition est que la somme de deux variables gaussiennes indépendantes est encore gaussienne. Il est alors facile de déterminer ses paramètres en calculant l'espérance et la variance.

Démonstration: Commençons par réduire le problème au cas où X est centrée réduite et Y centrée en supposant démontrée la proposition suivante : si $X' \sim \mathcal{N}(0, 1)$ et $Y' \sim \mathcal{N}(0, s^2)$, avec $s > 0$, sont indépendantes, alors $X' + Y' \sim \mathcal{N}(0, 1 + s^2)$. Nous pouvons alors en déduire la proposition dans le cas général en définissant les variables aléatoires réelles :

$$X' = \frac{X - m_1}{\sigma_1} \quad , \quad Y' = \frac{Y - m_2}{\sigma_1} \quad ,$$

si bien que les hypothèses précédentes sont satisfaites avec $s = \sigma_2/\sigma_1$ et :

$$X + Y = m_1 + \sigma_1 X' + m_2 + \sigma_1 Y' = m_1 + m_2 + \sigma_1 (X' + Y').$$

Par transformation affine de la variable $X' + Y' \sim \mathcal{N}(0, 1 + (\sigma_2^2/\sigma_1^2))$, nous obtenons alors la conclusion dans le cas général.

Nous supposons donc désormais $m_1 = m_2 = 0$, $\sigma_1 = 1$ et $\sigma_2 = s > 0$. D'après la Proposition 15.5.2, page 274, la variable aléatoire réelle $X + Y$ admet une densité proportionnelle à :

$$\int_{-\infty}^{+\infty} e^{-\frac{(x-y)^2}{2}} e^{-\frac{y^2}{2s^2}} dy = \int_{-\infty}^{+\infty} \exp\left(-\frac{s^2 x^2 - 2s^2 xy + (1+s^2)y^2}{2s^2}\right) dy.$$

Cette dernière expression s'écrit encore :

$$\int_{-\infty}^{+\infty} \exp\left(-\frac{(1+s^2)\left(y - \frac{s^2}{1+s^2}x\right)^2}{2s^2}\right) \exp\left(\left(\frac{s^2}{1+s^2} - 1\right)\frac{x^2}{2}\right) dy,$$

ou encore, puisque l'intégration porte sur la variable y ,

$$\exp\left(-\frac{x^2}{2(1+s^2)}\right) \int_{-\infty}^{+\infty} \exp\left(-\frac{(1+s^2)\left(y - \frac{s^2}{1+s^2}x\right)^2}{2s^2}\right) dy.$$

Nous constatons que cette dernière intégrale est une constante en effectuant, à x fixé, le simple changement de variable $u = y - \frac{s^2}{1+s^2}x$.

Nous avons donc prouvé que la variable aléatoire réelle $X + Y$ admet une densité proportionnelle à $\exp(-x^2/(2(1+s^2)))$. Nous en déduisons que la seule constante de proportionnalité possible pour que ce soit effectivement une densité de probabilité vaut $1/\sqrt{2\pi(1+s^2)}$ et que $X + Y \sim \mathcal{N}(0, 1 + s^2)$. \square

Pour une présentation de la loi normale dans \mathbb{R}^d , encore appelée loi de Laplace-Gauss en dimension d , le lecteur pourra lire [ESC 147ss] qui en fait une présentation cadrant bien avec le programme officiel.

17.3 Approximation normale de la loi binomiale

Historiquement, Abraham de Moivre, mathématicien anglais d'origine française, avait mis en évidence dès 1733 une approximation normale de la loi binomiale en étudiant un modèle de pile ou face avec une pièce équilibrée.

Sa preuve reposait sur des calculs laborieux d'estimation des coefficients binomiaux.

Pierre-Simon de Laplace avait généralisé ce résultat en 1820 au cas d'une pièce éventuellement biaisée. Il avait également mis en évidence le lien entre la loi normale et la théorie des erreurs d'observation, dont nous reparlerons dans le chapitre suivant.

Nous suivons ici la présentation d'[OUVRARD tome 1, 228-229].

Théorème 17.3.1 (de Moivre, Laplace) *Considérons $p \in]0, 1[$ et, pour tout $n \in \mathbb{N}^*$, une variable aléatoire réelle S_n qui suit la loi binomiale $B(n, p)$. Nous noterons \tilde{S}_n la variable centrée réduite associée à S_n , i.e.*

$$\tilde{S}_n = \frac{S_n - E[S_n]}{\sigma(S_n)} = \frac{S_n - np}{\sqrt{np(1-p)}}.$$

Alors nous avons la convergence suivante, uniforme sur tous les intervalles réels I :

$$\sup_I \left| P(\tilde{S}_n \in I) - \int_I \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt \right| \xrightarrow{n \rightarrow +\infty} 0.$$

Autrement dit, quand n est «grand», \tilde{S}_n suit une loi approximativement égale à la loi normale centrée réduite, ce que nous notons $\tilde{S}_n \approx \mathcal{N}(0, 1)$.

En utilisant un changement de variable affine dans une intégrale simple, nous pouvons dire de façon équivalente que $S_n \approx \mathcal{N}(np, np(1-p))$, toujours pour n «grand». Plus précisément, nous avons le résultat suivant :

Corollaire 17.3.2 *Considérons $p \in]0, 1[$ et, pour tout $n \in \mathbb{N}^*$, une variable aléatoire réelle S_n qui suit la loi binomiale $B(n, p)$. Alors nous avons la convergence suivante, uniforme sur tous les intervalles réels I :*

$$\sup_I \left| P(S_n \in I) - \int_I \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left(-\frac{(t - np)^2}{2np(1-p)}\right) dt \right| \xrightarrow{n \rightarrow +\infty} 0.$$

Pour pouvoir utiliser cette approximation gaussienne de la loi binomiale, il nous reste à préciser ce que signifie n «grand» ! Le théorème de Berry¹-Esseen², dont nous donnerons l'énoncé général dans le chapitre suivant, met en évidence le rôle joué par la quantité $np(1-p)$ dans la qualité de l'approximation.

1. Andrew C. Berry.

2. Carl-Gustav Esseen (1918-2001), mathématicien suédois.

Proposition 17.3.3 (Berry-Esseen, cas binomial) *Nous reprenons les hypothèses et les notations du théorème 17.3.1 et nous notons en outre $F_{\tilde{S}_n}$ la fonction de répartition de la variable \tilde{S}_n et Φ la fonction de répartition de la loi $\mathcal{N}(0, 1)$. Nous avons alors, pour tout $n \in \mathbb{N}^*$:*

$$\sup_{x \in \mathbb{R}} |F_{\tilde{S}_n}(x) - \Phi(x)| \leq \frac{p^2 + (1-p)^2}{\sqrt{np(1-p)}}.$$

Si $np(1-p)$ est suffisamment grand, le théorème de Berry-Esseen garantit que l'approximation de la loi binomiale par la loi normale sera bonne. Dans la pratique, certaines règles empiriques existent, variables d'ailleurs d'un ouvrage à l'autre. Voici celle qui est adoptée par Dacunha-Castelle et Duflo :

$$B(n, p) \approx \mathcal{N}(np, np(1-p)) \text{ dès que } np(1-p) > 18.$$

Voici une autre règle, adoptée dans le programme officiel de probabilités au lycée.

$$B(n, p) \approx \mathcal{N}(np, np(1-p)) \text{ dès que } n \geq 30, np \geq 5 \text{ et } n(1-p) \geq 5.$$

Si les conditions précédentes ne sont pas satisfaites, on pourra envisager l'approximation de la loi binomiale par la loi de Poisson. Voici une règle empirique énoncée par OUVRARD :

$$B(n, p) \approx \mathcal{P}(np) \text{ dès que } n \geq 30 \text{ et } p \leq 0,1.$$

Notons que si $p \geq 0,9$, l'approximation poissonnienne est aussi utilisable en comptant les échecs à la place des succès (ou les faces au lieu des piles!).

Le lecteur pourra trouver des applications en statistique de l'approximation gaussienne d'une loi binomiale dans [MEY 131-132] : Exemple 3.7 (intervalle de confiance dans un sondage) et Exercice 2 (intervalle de fluctuation pour le nombre de clients se connectant à un serveur informatique).

Chapitre 18

Le théorème-limite central

Ce théorème-limite est considéré comme central en théorie des probabilités pour les deux raisons suivantes :

1. Il précise à quelle vitesse a lieu la convergence énoncée par la loi des grands nombres, ce qui a une importance cruciale pour les applications (intervalle de confiance en statistique, méthode de Monte-Carlo pour le calcul approché d'une intégrale etc.)
2. Il permet de répondre à la question suivante : Pourquoi les lois normales (ou gaussiennes) jouent-elles un rôle central dans la modélisation de phénomènes aussi divers qu'une erreur de mesure en Physique, une donnée psychométrique en Psychologie différentielle, une fluctuation dans l'évolution d'un cours de la bourse en Finance ou de la taille d'une population animale en Biologie ?
Pourquoi le champ d'application des lois de Gauss est-il si large qu'on les a qualifiées de lois *normales* ? Après tout, la densité gaussienne n'a pas une forme si simple que cela ; pourquoi donc la considérer comme plus normale que les autres ?

Le théorème-limite central (ou *théorème central limite* selon la terminologie du programme officiel) nous dit *en gros* que **si l'on additionne beaucoup de variables indépendantes de variances finies, alors le résultat est une variable approximativement gaussienne.**

La précision de ce résultat sous des hypothèses aussi faibles est surprenante car si l'on sait qu'une loi est gaussienne – i.e. de la forme $\mathcal{N}(m, \sigma^2)$ – alors on dispose de procédures statistiques qui permettent d'estimer les deux paramètres m et σ^2 et donc d'identifier complètement la loi (approximative) du phénomène étudié. Or notre hypothèse de départ sur les lois des variables indépendantes que l'on additionne est très faible : on demande simplement qu'elle soient de variances finies.

18.1 Énoncé classique

Conformément au programme officiel, nous admettons ce théorème.

Théorème 18.1.1 (Théorème-limite central) *Soit $(X_i)_{i \in \mathbb{N}^*}$ une suite i.i.d. de variables aléatoires réelles dans $L^2(\Omega, \mathcal{A}, P)$; notons m leur espérance commune et σ^2 leur variance commune et posons, pour tout $n \in \mathbb{N}^*$, $S_n = X_1 + \cdots + X_n$. Enfin, soit Z une variable aléatoire réelle de loi $\mathcal{N}(0, 1)$. Alors, pour tout intervalle réel I , nous avons la convergence suivante :*

$$P\left(\frac{S_n - nm}{\sigma\sqrt{n}} \in I\right) \xrightarrow{n \rightarrow +\infty} P(Z \in I). \quad (18.1)$$

On dit que la suite de variables aléatoires réelles $((S_n - nm)/(\sigma\sqrt{n}))_{n \in \mathbb{N}^*}$ converge en loi vers Z (ou encore vers la loi normale $\mathcal{N}(0, 1)$) et l'on note :

$$\frac{S_n - nm}{\sigma\sqrt{n}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

Remarques.

1. La variable aléatoire réelle qui apparaît dans le membre de gauche est la variable centrée réduite associée à S_n par transformation affine. On vérifie en effet immédiatement que $E[S_n] = nm$ et $\text{Var}S_n = n\sigma^2$. Pour alléger les écritures, nous utiliserons la notation :

$$\tilde{S}_n := \frac{S_n - nm}{\sigma\sqrt{n}}$$

2. Si nous notons \bar{X}_n la moyenne empirique d'ordre n associée à la suite $(X_i)_{i \in \mathbb{N}^*}$, alors le théorème-limite central s'écrit encore, pour tout intervalle réel I ,

$$P\left(\frac{\sqrt{n}}{\sigma}(\bar{X}_n - m) \in I\right) \xrightarrow{n \rightarrow +\infty} P(Z \in I).$$

Ainsi, si $I = [a, b]$, nous obtenons en utilisant la densité de la loi normale centrée réduite :

$$P\left(a \leq \frac{\sqrt{n}}{\sigma}(\bar{X}_n - m) \leq b\right) \xrightarrow{n \rightarrow +\infty} \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

Si nous prenons $-a = b = 1,96$, nous obtenons (d'après la fin du paragraphe précédent) que lorsque n est «grand», il y a approximativement 95 chances sur 100 pour que l'on ait : $|\bar{X}_n - m| \leq 1,96\sigma/\sqrt{n}$.

C'est en ce sens que le théorème-limite central implique que l'erreur commise en faisant une approximation de m par \bar{X}_n , comme nous y invite la loi des grands nombres, est d'ordre de grandeur σ/\sqrt{n} pour n «grand». On dit de façon équivalente que la convergence énoncée par la loi des grands nombres a lieu avec une vitesse d'ordre de grandeur \sqrt{n}/σ .

3. La convergence énoncée par la loi des grands nombres était valable sous l'hypothèse $(X_i)_{i \in \mathbb{N}^*}$ i.i.d. dans $L^1(\Omega, \mathcal{A}, P)$ mais pour préciser sa vitesse grâce au théorème-limite central, nous avons besoin d'une hypothèse supplémentaire : la suite $(X_i)_{i \in \mathbb{N}^*}$ est supposée dans $L^2(\Omega, \mathcal{A}, P)$.
4. En prenant $I = [a, b]$ (pour fixer les idées) et en faisant un simple changement de variable dans une intégrale, on montre que la suite $((S_n - nm)/\sqrt{n})_{n \in \mathbb{N}^*}$ converge en loi vers la loi normale $\mathcal{N}(0, \sigma^2)$.
5. La convergence en loi est une notion plus difficile à manipuler que la convergence en probabilité ou la convergence presque sûre. Elle est la plus faible de ces trois types de convergences : on peut montrer que, pour des variables aléatoires réelles quelconques $(Y_n)_{n \in \mathbb{N}^*}$ et Y , on a les implications

$$Y_n \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} Y \implies Y_n \xrightarrow[n \rightarrow +\infty]{(P)} Y \implies Y_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Y.$$

Malgré la difficulté de cette notion, on est obligé d'en passer par la convergence en loi tout simplement parce que les types plus forts de convergence ne permettent pas d'obtenir un théorème du genre théorème-limite central : sous les hypothèses de ce théorème, on peut en effet prouver qu'il n'existe aucun $\alpha > 0$ tel que la suite $(n^\alpha(\bar{X}_n - m))_{n \in \mathbb{N}^*}$ converge en probabilité vers une variable aléatoire non presque sûrement nulle.

18.2 TLC et fonctions de répartition

Corollaire 18.2.1 *Sous les hypothèses du théorème-limite central et en notant Φ la fonction de répartition de la loi $\mathcal{N}(0, 1)$, nous avons la convergence suivante pour tout $x \in \mathbb{R}$:*

$$P\left(\frac{S_n - nm}{\sigma\sqrt{n}} \leq x\right) \xrightarrow[n \rightarrow +\infty]{} \Phi(x). \quad (18.2)$$

Démonstration: Il suffit d'appliquer le théorème 18.1.1 avec $I =] - \infty, x]$.
 \square

Remarques.

1. On peut reformuler cet énoncé en disant que, si l'on note F_n la fonction de répartition de la variable aléatoire réelle \tilde{S}_n , alors la suite $(F_n)_{n \in \mathbb{N}^*}$ converge simplement sur \mathbb{R} vers la fonction de répartition Φ de la loi normale centrée réduite.
2. Le corollaire reste vrai si l'on remplace l'inégalité large par une inégalité stricte. Il suffit en effet de prendre $I =] - \infty, x[$ dans le théorème précédent. En nous souvenant des propriétés des fonctions de répartition, nous pouvons écrire ceci sous la forme :

$$\forall x \in \mathbb{R} \quad F_n(x-) \xrightarrow{n \rightarrow +\infty} \Phi(x) \quad (18.3)$$

3. On peut montrer que cet énoncé 18.2.1 n'est pas seulement un corollaire du théorème 18.1.1 mais lui est en fait équivalent. C'est pourquoi le lecteur trouvera le théorème-central limite énoncé sous cette forme (convergence simple de fonctions de répartition) dans plusieurs ouvrages. De même, la convergence (18.3) est en fait équivalente à l'énoncé du théorème-limite central.

Les convergences (18.2) et (18.3), qui sont donc équivalentes à l'énoncé du théorème-limite central, peuvent être améliorées. En effet, un critère de convergence uniforme appelé *second théorème de Dini*, permet de démontrer le résultat suivant, que nous admettrons.

Lemme 18.2.2 *Soit $(F_n)_{n \in \mathbb{N}^*}$ une suite de fonctions de répartition qui converge simplement sur \mathbb{R} vers une fonction de répartition continue F . Alors $(F_n)_{n \in \mathbb{N}^*}$ converge uniformément vers F sur \mathbb{R} .*

La fonction de répartition Φ de la loi $\mathcal{N}(0, 1)$ étant continue (c'est d'ailleurs vrai pour toute fonction de répartition d'une loi qui admet une densité de probabilité), nous déduisons immédiatement de ce lemme que les convergences (18.2) et (18.3) sont en réalité uniformes sur \mathbb{R} . Par exemple, nous avons, en notant F_n la fonction de répartition de la variable aléatoire réelle \tilde{S}_n :

$$\sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| \xrightarrow{n \rightarrow +\infty} 0 \quad (18.4)$$

Il est alors facile d'en déduire la proposition suivante, que le lecteur pourra prouver à titre d'exercice, en utilisant des égalités telles que :

$$P(\tilde{S}_n \in [a, b]) = F_n(b) - F_n(a-), \quad P(\tilde{S}_n \in]a, b]) = F_n(b-) - F_n(a)$$

Proposition 18.2.3 *Sous les hypothèses et avec les notations du théorème-limite central, la convergence (18.1) est uniforme sur tous les intervalles réels I , ce qui s'écrit :*

$$\sup_I \left| P \left(\frac{S_n - nm}{\sigma\sqrt{n}} \in I \right) - P(Z \in I) \right| \xrightarrow{n \rightarrow +\infty} 0 \quad (18.5)$$

Lorsque n tend vers l'infini, la loi de la variable \tilde{S}_n et la loi de la variable Z , c'est-à-dire la loi $\mathcal{N}(0, 1)$, deviennent arbitrairement proches uniformément sur tous les intervalles réels I . C'est en ce sens que l'on peut dire que, lorsque n est suffisamment grand, la loi de la variable \tilde{S}_n est approximativement gaussienne centrée réduite, ce qu'on écrira (informellement) $\tilde{S}_n \approx \mathcal{N}(0, 1)$. Revenons maintenant au fait que les convergences (18.2) et (18.3) ont lieu uniformément sur \mathbb{R} . Si nous notons F_{S_n} la fonction de répartition de la variable S_n et $\Phi_{nm, n\sigma^2}$ la fonction de répartition de la loi $\mathcal{N}(nm, n\sigma^2)$, on démontre facilement que, pour tout $n \in \mathbb{N}^*$ et tout $x \in \mathbb{R}$:

$$F_{S_n}(x) = F_n \left(\frac{x - nm}{\sigma\sqrt{n}} \right) \text{ et } \Phi_{nm, n\sigma^2}(x) = \Phi \left(\frac{x - nm}{\sigma\sqrt{n}} \right).$$

Nous en déduisons les convergences uniformes suivantes :

$$\sup_{x \in \mathbb{R}} |F_{S_n}(x) - \Phi_{nm, n\sigma^2}(x)| \xrightarrow{n \rightarrow +\infty} 0, \quad \sup_{x \in \mathbb{R}} |F_{S_n}(x-) - \Phi_{nm, n\sigma^2}(x)| \xrightarrow{n \rightarrow +\infty} 0.$$

Par un raisonnement similaire à ce qui a été fait précédemment, nous pouvons en déduire que, lorsque n tend vers l'infini, la loi de la variable S_n et la loi $\mathcal{N}(nm, n\sigma^2)$ deviennent arbitrairement proches uniformément sur tous les intervalles réels I . C'est en ce sens que l'on peut dire que, lorsque n est suffisamment grand, la loi de la variable S_n vaut approximativement la loi $\mathcal{N}(nm, n\sigma^2)$, ce que nous écrivons : $S_n \approx \mathcal{N}(nm, n\sigma^2)$.

Il nous reste à préciser ce que signifie n «grand» ! Comme nous l'avons fait dans le chapitre précédent à propos du théorème de De Moivre-Laplace, nous utilisons le *théorème de Berry-Esseen*, ici dans son énoncé général, lequel nous donne des précisions sur la vitesse à laquelle a lieu la convergence énoncée par le théorème-limite central.

Le lecteur pourra retrouver cet énoncé dans “L'essentiel en théorie des probabilités” par Jean JACOD et Philip PROTTER chez Cassini, page 191 (corriger le membre de droite en recentrant la variable X_1). On ne connaît pas la valeur optimale de la constante C intervenant ci-dessous mais on sait que l'on peut prendre $C \leq 0,8$.

Théorème 18.2.4 (Berry-Esseen) *Soit $(X_i)_{i \in \mathbb{N}^*}$ une suite indépendante identiquement distribuée de variables aléatoires réelles dans $L^3(\Omega, \mathcal{A}, P)$. En reprenant les notations de (18.4), il existe une constante $C > 0$ telle que :*

$$\forall n \in \mathbb{N}^* \quad \sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| \leq C \frac{E[|X_1 - E[X_1]|^3]}{\sigma^3 \sqrt{n}}$$

Pour savoir à quelle vitesse avait lieu la convergence dans la loi des grands nombres, nous avons dû rajouter une condition L^2 et la réponse nous a alors été fournie par le théorème-limite central. Pour savoir à quelle vitesse avait lieu la convergence énoncée par ce dernier, nous avons dû rajouter une condition L^3 et la réponse nous a alors été fournie par le théorème de Berry-Esseen. Remarquons que les questions de vitesse de convergence s'arrêtent là car le théorème de Berry-Esseen n'est pas un théorème asymptotique : la majoration qu'il énonce est valable pour tout entier $n \in \mathbb{N}^*$, même petit.

18.3 Un énoncé sans équidistribution

Nous avons annoncé dans l'introduction que, selon le théorème-limite central, si l'on additionne beaucoup de variables indépendantes admettant un moment d'ordre deux, alors le résultat est une variable approximativement gaussienne. C'est ce que nous venons de montrer mais à une objection près : jusqu'à présent, nous n'avons pas seulement supposé les variables indépendantes et dans $L^2(\Omega, \mathcal{A}, P)$ mais nous leur avons également demandé d'avoir toutes la même loi. Nous dirons simplement à ce sujet qu'il existe divers raffinements du théorème-limite central qui permettent, sous certaines conditions, de s'affranchir de l'hypothèse d'équidistribution. Nous donnons un tel énoncé à titre d'exemple, en l'admettant.

Proposition 18.3.1 (TLC, suite bornée dans $L^{2+\epsilon}$) *Soit $\epsilon > 0$ et $(X_i)_{i \in \mathbb{N}^*}$ une suite de variables aléatoires réelles indépendantes mais non nécessairement de même loi, formant une famille bornée dans $L^{2+\epsilon}(\Omega, \mathcal{A}, P)$, i.e. telle que :*

$$\sup_{i \in \mathbb{N}^*} E[|X_i|^{2+\epsilon}] < +\infty.$$

Par commodité, nous supposons $E[X_i] = 0$ et nous notons $\sigma_i^2 = \text{Var} X_i$ pour tout $i \in \mathbb{N}^$.*

Alors, si $\sum_{i=1}^{+\infty} \sigma_i^2 = +\infty$, nous avons la convergence en loi suivante :

$$\frac{S_n}{\sqrt{\sum_{i=1}^n \sigma_i^2}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

18.4 Test du χ^2 d'ajustement.

La construction de ce test statistique d'usage très courant fait appel à une version multidimensionnelle du théorème-limite central qui est hors programme. Nous admettrons donc certains résultats dans la suite, en présentant les idées essentielles permettant la mise en oeuvre de ce test.

Nous considérons n répétitions d'une expérience aléatoire qui produit un résultat dans un ensemble fini, par exemple $\llbracket 1; k \rrbracket$.

Nous avons des raisons de supposer que la loi sur $\llbracket 1; k \rrbracket$ qui gouverne cette expérience est donnée par $p = (p^1, \dots, p^k)$, avec $p^i \geq 0$ pour tout $i \in \llbracket 1; k \rrbracket$ et $p^1 + \dots + p^k = 1$, mais nous voudrions confirmer ou infirmer cette hypothèse au regard des valeurs $(x_1, \dots, x_n) = (X_1(\omega), \dots, X_n(\omega))$ observées au cours des n répétitions de l'expérience, c'est-à-dire tester l'*ajustement* de notre modèle à la réalité expérimentale.

Pour ce faire, une première étape consiste à introduire une sorte de distance entre les lois de probabilité sur l'ensemble $\llbracket 1; k \rrbracket$, l'idée étant de regarder ensuite si la loi empirique (définie ci-dessous et calculable à partir du résultat de l'expérience) est proche ou éloignée de la loi théorique p dont nous avons fait l'hypothèse.

Définition 18.4.1 Soient p et q deux lois de probabilité sur $\llbracket 1; k \rrbracket$.

On appelle «distance» du χ^2 entre p et q la quantité :

$$d_{\chi^2}(p, q) = \sum_{i=1}^k \frac{(p^i - q^i)^2}{p^i}$$

Le mot distance est entre guillemets car ce n'est pas du tout une distance au sens des espaces métriques : elle n'est visiblement pas symétrique et il est facile de constater qu'elle ne vérifie pas non plus l'inégalité triangulaire. En réalité, son seul rapport avec une distance est la propriété suivante :

$$d_{\chi^2}(p, q) = 0 \Leftrightarrow p = q$$

Remarquons que cette pseudo-distance a tendance à surévaluer les différences entre p et q sur les entiers i où p_i est petit : nous chercherons à limiter ce phénomène dans la suite en imposant des conditions telles que $np_i \geq 5$ pour tout $i \in \llbracket 1; k \rrbracket$.

La deuxième étape de construction du test qui va nous permettre de confirmer ou infirmer notre hypothèse consiste à comparer la loi théorique p avec la loi empirique \bar{p}_n que nous définissons comme suit :

Définition 18.4.2 Si (X_1, \dots, X_n) est la variable aléatoire modélisant les n répétitions de notre expérience, nous posons :

$$\forall i \in \llbracket 1; k \rrbracket, \quad N_n^i = \sum_{j=1}^n \mathbf{1}_{\{X_j=i\}}$$

Nous appelons alors loi empirique, notée \bar{p}_n , la loi sur $\llbracket 1; k \rrbracket$ définie par :

$$\forall i \in \llbracket 1; k \rrbracket, \quad \bar{p}_n^i = \frac{N_n^i}{n}$$

Notons que la valeur de \bar{p}_n dépend du résultat ω de l'expérience, d'où le qualificatif empirique. En toute rigueur, c'est d'ailleurs $\bar{p}_n(\omega)$ (et non pas \bar{p}_n) qui est une loi de probabilité sur $\llbracket 1; k \rrbracket$.

Définition 18.4.3 On appelle χ^2 d'ajustement la variable aléatoire suivante :

$$nd_{\chi^2}(p, \bar{p}_n) = n \sum_{i=1}^k \frac{(p^i - \bar{p}_n^i)^2}{p^i} = \sum_{i=1}^k \frac{(np^i - N_n^i)^2}{np^i}$$

Nous rappelons maintenant la définition d'une loi de probabilité sur \mathbb{R} qui a été introduite dans l'Application 15.5.5, page 276, et identifiée à une loi Gamma, ce qui nous fournit sa densité.

Définition 18.4.4 Considérons $Z = (Z_1, \dots, Z_d)$ un vecteur aléatoire dont les composantes sont indépendantes et de même loi $\mathcal{N}(0, 1)$.

Alors la loi de la variable aléatoire positive $\|Z\|^2 = Z_1^2 + \dots + Z_d^2$ est appelée loi du χ^2 à d degrés de liberté et notée $\chi^2(\mathbf{d})$.

Cette loi est en fait égale à la loi $\gamma(\frac{1}{2}, \frac{d}{2})$ et admet donc pour densité :

$$g_{\frac{1}{2}, \frac{d}{2}}(x) = \frac{1}{2^{\frac{d}{2}} \Gamma(\frac{d}{2})} e^{-\frac{x}{2}} x^{\frac{d}{2}-1} \mathbf{1}_{\mathbb{R}_+^*}(x)$$

Le résultat essentiel qui va nous permettre de construire le test dit du χ^2 est le suivant :

Proposition 18.4.5 (Pearson) Si pour tout $n \in \mathbb{N}^*$, le vecteur aléatoire (X_1, \dots, X_n) suit la loi $p^{\otimes n}$, alors la convergence suivante a lieu :

$$nd_{\chi^2}(p, \bar{p}_n) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \chi^2(\mathbf{k} - \mathbf{1})$$

Autrement dit, si nous notons F_{k-1} la fonction de répartition de la loi $\chi^2(\mathbf{k} - \mathbf{1})$, nous avons la convergence suivante pour tout $x \in \mathbb{R}$:

$$P(nd_{\chi^2}(p, \bar{p}_n) \leq x) \xrightarrow[n \rightarrow +\infty]{} F_{k-1}(x)$$

Comme annoncé plus haut, la démonstration de ce résultat nécessite la version multidimensionnelle du théorème-limite central, hors programme, donc nous l'admettrons. Remarquons simplement que si (X_1, \dots, X_n) suit la loi $p^{\otimes n}$, c'est-à-dire si les variables X_i sont indépendantes et de même loi p , alors le TLC (en dimension 1) implique la convergence suivante :

$$\frac{N_n^i - np^i}{\sqrt{np^i}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1 - p_i)$$

Il est alors assez intuitif que la loi limite du χ^2 d'ajustement, c'est-à-dire de la variable suivante :

$$nd_{\chi^2}(p, \bar{p}_n) = \sum_{i=1}^k \left(\frac{N_n^i - np^i}{\sqrt{np^i}} \right)^2$$

soit une loi du χ^2 mais on pourrait penser que celle-ci a k degrés de liberté, alors que la vraie loi limite est $\chi^2(\mathbf{k} - \mathbf{1})$. En fait, la «perte» d'un degré de liberté peut se comprendre en constatant que nos variables ne sont pas totalement «libres» puisqu'il existe entre elles la relation linéaire suivante :

$$\sum_{i=1}^k (N_n^i - np^i) = 0$$

La proposition de Pearson nous dit donc que, si l'expérience est bien gouvernée par la loi p supposée, alors le χ^2 d'ajustement suit une loi proche de $\chi^2(\mathbf{k} - \mathbf{1})$ lorsque le nombre n de répétitions de l'expérience devient «grand».

En revanche, si l'expérience est en réalité régie par une loi $q \neq p$, alors il existe $1 \leq i \leq k$ tel que $q^i \neq p^i$ et la loi des grands nombres implique la convergence suivante lorsque n tend vers l'infini :

$$d_{\chi^2}(p, \bar{p}_n) = \sum_{i=1}^k \frac{(p^i - \bar{p}_n^i)^2}{p^i} \xrightarrow{\text{p.s.}} d_{\chi^2}(p, q) > 0$$

d'où nous déduisons le comportement asymptotique du χ^2 d'ajustement :

$$nd_{\chi^2}(p, \bar{p}_n) \xrightarrow{\text{p.s.}} +\infty \quad (18.6)$$

C'est la différence entre ces deux comportements asymptotiques qui va nous permettre de tester l'hypothèse H_0 : «L'expérience est gouvernée par la loi p » contre l'hypothèse alternative H_1 : «L'expérience est gouvernée par une loi $q \neq p$ ».

Passons à la construction effective du test.

Si nous notons F_{k-1} la fonction de répartition de la loi $\chi^2(\mathbf{k} - \mathbf{1})$, nous prouvons facilement que F_{k-1} est une bijection de \mathbb{R}_+ sur $[0, 1[$ en constatant que la densité de la loi $\chi^2(\mathbf{k} - \mathbf{1})$ est strictement positive sur \mathbb{R}_+^* et nulle sur \mathbb{R}_- . Par conséquent, pour tout $\alpha \in]0, 1]$, il existe un unique $c_\alpha \in \mathbb{R}_+$ tel que $\chi^2(\mathbf{k} - \mathbf{1})(]c_\alpha, +\infty[) = \alpha$ et l'on a $c_\alpha = F_{k-1}^{-1}(1 - \alpha)$.

Si l'hypothèse H_0 est réalisée, donc si l'expérience est régie par la loi p , la proposition de Pearson implique :

$$P_p(nd_{\chi^2}(p, \bar{p}_n) > c_\alpha) \xrightarrow{n \rightarrow +\infty} \alpha \quad (18.7)$$

En revanche, si c'est l'hypothèse H_1 qui est réalisée, donc si l'expérience est gouvernée par une loi $q \neq p$, alors le comportement asymptotique (18.6) entraîne la convergence suivante lorsque n tend vers l'infini :

$$P_q(nd_{\chi^2}(p, \bar{p}_n) > c_\alpha) \xrightarrow{n \rightarrow +\infty} 1 \quad (18.8)$$

Nous pratiquons donc notre test comme suit :

Nous choisissons une valeur $\alpha \in]0, 1]$ (typiquement α est petit car, comme nous allons le voir, il représente le niveau d'erreur du test) et nous en déduisons la valeur c_α . Pour le résultat ω de l'expérience que nous observons, nous calculons la valeur du χ^2 d'ajustement :

$$nd_{\chi^2}(p, \bar{p}_n(\omega)) = \sum_{i=1}^k \frac{(np^i - N_n^i(\omega))^2}{np^i}$$

Nous comparons alors cette valeur à c_α pour conclure :

- Si $nd_{\chi^2}(p, \bar{p}_n(\omega)) > c_\alpha$, alors nous rejetons l'hypothèse H_0 .
- Si $nd_{\chi^2}(p, \bar{p}_n(\omega)) \leq c_\alpha$, alors nous acceptons l'hypothèse H_0 .

De façon générale, lorsque nous pratiquons un test statistique, notre conclusion peut être erronée de deux façons différentes :

Erreur de 1ère espèce : Je rejette l'hypothèse H_0 alors qu'elle est satisfaite en réalité.

Erreur de 2nde espèce : J'accepte l'hypothèse H_0 alors qu'elle n'est pas satisfaite en réalité.

Dans de nombreuses situations pratiques, ces deux types d'erreurs ne sont pas symétriques, l'erreur de première espèce étant considérée comme plus grave que l'autre :

Par exemple, si je teste le câble d'un ascenseur supposé pouvoir accueillir 10 personnes (750kg) et si je note M la masse critique à partir de laquelle le câble casse, je choisirai H_0 : « $M \leq 750$ » et H_1 : « $M > 750$ ».

L'erreur de 1ère espèce conduirait des usagers de l'ascenseur à un grand plongeon : c'est ce risque que je veux absolument maîtriser. L'erreur de 2nde espèce conduirait à des réparations inutiles sur l'ascenseur : je veux l'éviter mais elle est moins grave que la première.

Dans le test du χ^2 que nous venons de construire, la convergence (18.7) se traduit comme suit : la probabilité de commettre une erreur de 1ère espèce est asymptotiquement égale à α . On dit qu'on a construit un test de niveau d'erreur asymptotique α ou de *niveau de confiance asymptotique* $1 - \alpha$.

Quant à la convergence (18.8), sa traduction est plus vague : lorsque n devient «grand», la probabilité de commettre une erreur de 2nde espèce devient «petite» mais nous ne maîtrisons pas la vitesse à laquelle cette convergence se produit. On dit que la *puissance* du test, c'est-à-dire la probabilité de rejeter l'hypothèse H_0 quand elle n'est effectivement pas satisfaite dans la réalité, tend vers 1 lorsque n tend vers l'infini.

Nous terminons cette introduction au test du χ^2 en le généralisant au cas où l'ensemble E des résultats possibles pour l'expérience est infini. Nous pouvons alors adapter notre méthode comme suit :

Nous choisissons une partition finie (E_1, \dots, E_k) de l'ensemble E . Si ν est la loi sur E supposée gouverner l'expérience, nous posons $p^i = \nu(E_i)$ et nous comptons maintenant le nombre de fois où l'on tombe dans la classe E_i au cours des n répétitions de l'expérience :

$$N_n^i(\omega) = \sum_{j=1}^n \mathbf{1}_{\{X_j(\omega) \in E_i\}}$$

Le reste du test se déroule comme précédemment.

Notons qu'avec cette méthode, nous ne testons en réalité que l'adéquation aux valeurs de ν sur les différentes classes E_i .

Attention, le choix des classes E_i n'est pas du tout innocent ! Une règle d'or, liée à la remarque faite à la suite de la définition de d_{χ^2} , est que les *effectifs théoriques* np_i (espérance de N^i dans le cas où l'expérience est vraiment gouvernée par la loi ν) des classes E_i ne doivent jamais être trop «petits»

Dans la pratique, on impose souvent que la condition suivante soit satisfaite :

$$\boxed{\forall i = \llbracket 1; k \rrbracket \quad np^i \geq 5}$$

D'autres règles pratiques existent : par exemple, on a constaté qu'il était préférable de choisir les classes E_i de sorte que leurs effectifs théoriques respectifs np_i soient à peu près tous équivalents. Nous n'insisterons pas plus sur ce sujet des règles pratiques qui concernent plus les orfèvres du test du χ^2 que les agrégatifs !

ISBN 978-2-91-635238-1



9 782916 352381