Cours de mathématiques
MP

Alain TROESCH

Version du 18 mars 2025

Table des matières

1 Suites numériques **7**
I Topologie réelle et complexe . . . 7
I.1 Intervalles . . . 7
I.2 Voisinages . . . 8
II Convergence . . . 9
II.1 Convergence, réexpression topologique . . . 9
II.2 Rappel des propriétés opératoires relatives à la convergence . . . 10
II.3 Comportement des limites vis-à-vis des inégalités . . . 11
II.4 Propriétés des suites monotones . . . 12
III Comparaisons asymptotiques . . . 12
IV Suites récurrentes . . . 15
IV.1 Suites récurrentes linéaires . . . 15
IV.2 Suites récurrentes linéaires perturbées . . . 16
IV.3 Suites récurrentes $u_{n+1} = f(u_n)$ . . . 17

2 Cardinaux et dénombrabilité **21**
I Rappels sur les cardinaux finis . . . 21
I.1 Ensembles finis, cardinaux . . . 21
I.2 Rappel des propriétés calculatoires sur les cardinaux . . . 22
I.3 Grands principes combinatoires . . . 23
II Dénombrabilité . . . 25
II.1 Cardinaux infinis . . . 25
II.2 Dénombrabilité . . . 26
II.3 Règles opératoires sur les ensembles dénombrables . . . 27
II.4 Des ensembles infinis non dénombrables . . . 28

3 Espaces vectoriels normés **31**
I Normes sur un espace vectoriel . . . 31
I.1 Définitions et propriétés élémentaires . . . 31
I.2 Normes sur des espaces de dimension finie . . . 33
I.3 Normes sur des espaces de fonctions . . . 34
I.4 Normes sur $\mathbb{K}[X]$ . . . 35
I.5 Normes sur $\mathcal{M}_n(\mathbb{K})$ . . . 35
I.6 Produits d'e.v.n. . . . 36
II Topologie d'un espace vectoriel normé . . . 37

II.1 Distance et boules . . . 37
II.2 Ouverts, fermés, voisinages . . . 39
II.3 Intérieur, adhérence, frontière . . . 41
II.4 Topologie relative . . . 44
III Convergences . . . 45
III.1 Suites convergentes . . . 45
III.2 Caractérisations séquentielles . . . 47
IV Valeurs d'adhérence et compacité . . . 48
IV.1 Valeur d'adhérence d'une suite . . . 48
IV.2 Théorème de Bolzano-Weierstass sur $\mathbb{R}$ . . . 50
IV.3 Sous-ensembles compacts . . . 51
V Comparaison de normes . . . 52
V.1 Domination . . . 52
V.2 Normes équivalentes . . . 54
VI Espaces métriques (HP) . . . 54

4 Intégrales généralisées 57
I Intégrales sur un intervalle quelconque . . . 57
I.1 Intégrale généralisée sur un intervalle $[a, b[$. . . . 58
I.2 Intégrales de référence . . . 60
I.3 Théorèmes de comparaison pour les intégrales de fonctions positives . . . 61
I.4 Extension aux intégrales sur un intervalle quelconque I . . . 63
I.5 Intégrabilité . . . 65
I.6 Bilan méthodologique . . . 68
II Méthodes calculatoires . . . 68
II.1 Changement de variables . . . 69
II.2 Intégration par parties . . . 70
III Études asymptotiques . . . 71
III.1 Intégration des relations de comparaison . . . 71
III.2 Théorème de convergence dominée . . . 72

5 Fonctions continues sur un e.v.n. 73
I Limites dans un e.v.n. . . . 73
I.1 Limites finies . . . 73
I.2 Limites infinies lorsque $F = \mathbb{R}$ . . . 74
I.3 Limite en l'infini . . . 75
I.4 Limites : point de vue topologique . . . 76
I.5 Opérations sur les limites . . . 78
I.6 Limites partielles . . . 79
II Continuité . . . 80
II.1 Continuité en un point . . . 80
II.2 Opérations sur les fonctions continues . . . 82
II.3 Continuité uniforme, fonctions lipschitziennes . . . 82
II.4 Exemples importants de fonctions continues . . . 83
II.5 Prolongements . . . 84
III Continuité et topologie . . . 85
III.1 Images réciproques d'ouverts et fermés . . . 85
III.2 Fonction continue sur compact . . . 86
IV Connexité . . . 87
IV.1 Parties connexes par arcs . . . 87
IV.2 Image continue d'une partie connexe par arcs . . . 88
V Continuité des applications linéaires . . . 88

VI E.v.n. de dimension finie, et conséquences sur la continuité . . . 90
VI.1 Équivalence des normes . . . 90
VI.2 Conséquences topologiques . . . 91
VI.3 Conséquences sur la continuité des applications linéaires . . . 92
VI.4 Applications polynomiales . . . 93

6 Séries numériques et vectorielles **95**
I Rappels et compléments sur les séries numériques . . . 95
I.1 Convergences . . . 95
I.2 Comparaisons séries-intégrales . . . 96
I.3 Théorèmes de comparaison pour la convergence absolue . . . 98
I.4 Semi-convergence . . . 99
I.5 Techniques asymptotiques . . . 100
II Séries dans un e.v.n. . . . 101
II.1 Définitions et propriétés générales . . . 101
II.2 Séries à valeurs dans un e.v.n. de dimension finie . . . 103

7 Structures algébriques **105**
I Groupes . . . 105
I.1 Rappels . . . 105
I.2 Sous-groupes . . . 106
I.3 Le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$ . . . 108
I.4 Groupes monogènes . . . 109
I.5 Ordre d'un élement . . . 110
I.6 Un deuxième théorème de Lagrange, HP . . . 111
II Anneaux et idéaux . . . 112
II.1 Rappels . . . 112
II.2 Idéaux . . . 115
II.3 Idéaux de $\mathbb{Z}$ et arithmétique . . . 116
III Anneau $\mathbb{Z}/n\mathbb{Z}$ . . . 117
III.1 Propriétés d'inversibilité . . . 117
III.2 Théorème chinois, et systèmes de congruences . . . 118
III.3 L'indicatrice d'Euler . . . 119
IV Arithmétique dans l'anneau $\mathbb{K}[X]$ . . . 120
IV.1 Idéaux de $\mathbb{K}[X]$ . . . 120
IV.2 Décomposition en facteurs irréductibles . . . 121

8 Compléments d'algèbre linéaire **123**
I Sommes directes et blocs . . . 124
I.1 Somme directe de n sous-espaces vectoriels . . . 124
I.2 Détermination d'une application linéaire . . . 126
I.3 Calcul matriciel par blocs . . . 128
I.4 Déterminants par blocs . . . 130
II Algèbres . . . 132
II.1 Définitions générales . . . 132
II.2 Spécialisation d'un polynôme . . . 134
III Polynômes d'endomorphismes . . . 136
III.1 Polynômes annulateurs et polynôme minimal . . . 136
III.2 Lemme de décomposition des noyaux . . . 138

9 Réduction des endomorphismes **139**
I Élements propres d'un endomorphisme 139
I.1 Valeurs propres, vecteurs propres, sous-espaces propres 139
I.2 Éléments propres d'une matrice carrée 142
I.3 Polynôme caractéristique d'un endomorphisme en dimension finie 143
I.4 Recherche de valeurs propres 145
I.5 Multiplicité géométrique et algébrique d'une valeur propre 145
I.6 Théorème de Cayley-Hamilton 146
II Diagonalisation 147
II.1 Diagonalisation d'un endomorphisme, point de vue géométrique 147
II.2 Diagonalisation d'un endomorphisme, point de vue algébrique 148
II.3 Diagonalisation d'une matrice 149
III Trigonalisation 150
III.1 Trigonalisabilité et caractérisations algébriques 150
III.2 Endomorphismes nilpotents 152
III.3 Sous-espaces caractéristiques 152

10 Suites et séries de fonctions **155**
I Modes de convergences 155
I.1 Convergence simple 155
I.2 Convergence uniforme 156
I.3 Convergence uniforme locale 157
II Régularité des limites de suites de fonctions 158
II.1 Continuité d'une limite uniforme 158
II.2 Double-limite 159
II.3 Interversion limite-intégrale 160
II.4 Dérivabilité et dérivée d'une limite uniforme 162
III Séries de fonctions 163
III.1 Modes de convergence 163
III.2 Régularité et interversions sur des sommes uniformes 165
III.3 Techniques asymptotiques 167
IV Approximations uniformes 168
IV.1 Principes généraux 168
IV.2 Approximations uniformes classiques 169

11 Intégrales à paramètres **171**
I Suites et séries d'intégrales 171
I.1 Interversion limite/intégrale 171
I.2 Intégration terme à terme 173
II Limite et continuité d'intégrales à paramètre dans un e.v.n. 176
III Dérivation d'intégrales à paramètre réel 177
IV La fonction Γ (HP mais classique) 179

12 Endomorphismes d'un espace euclidien **181**
I Rappels sur les espaces euclidiens 181
I.1 Formes bilinéaires et produit scalaire 181
I.2 Orthogonalité 183
I.3 Projection orthogonale 184
I.4 Coordonnées en base orthonormale 184
I.5 Changements de base orthonormales, matrices orthogonales 185
I.6 Orientation d'un espace 187
II Endomorphismes autoadjoints 188

II.1 Adjoint d'un endomorphisme . . . 188
II.2 Endomorphismes autoadjoints . . . 190
II.3 Théorème spectral . . . 191
II.4 Endomorphismes autoadjoints positifs . . . 192
III Isométries vectorielles . . . 193
III.1 Définitions . . . 193
III.2 Matrices orthogonales de $\mathcal{M}_2(\mathbb{R})$ . . . 195
III.3 Isométries en dimension 2 . . . 196
III.4 Réduction des isométries . . . 198

13 Séries entières **199**
I Définitions, convergence . . . 199
I.1 Définition . . . 199
I.2 Domaine de convergence, rayon de convergence . . . 200
I.3 Opérations sur les séries entières . . . 203
II Régularité de la somme d'une série entière . . . 204
II.1 Convergence uniforme . . . 204
II.2 Continuité et limite radiale . . . 204
II.3 Dérivabilité et classe $\mathcal{C}^\infty$ . . . 205
III Développements en séries entières . . . 207
III.1 Généralités sur les fonctions développables en séries entières . . . 207
III.2 DSE de l'exponentielle . . . 208
III.3 DSE usuels . . . 209

14 Espaces probabilisés **211**
I Espaces probabilisés . . . 211
I.1 Tribus et espaces probabilisables . . . 211
I.2 Mesures de probabilité, espaces probabilisés . . . 213
I.3 Espaces probabilisés discrets . . . 214
I.4 Continuité monotone . . . 215
I.5 Évenements négligeables, presque sûrs . . . 217
II Conditionnement et indépendance . . . 218
II.1 Probabilités conditionnelles . . . 218
II.2 Formules liées aux probabilités conditionnelles . . . 219
II.3 Indépendance . . . 221
III Bilan des méthodes calculatoires en probabilités . . . 223

15 Variables aléatoires discrètes **227**
I Loi d'une variable aléatoire discrète . . . 227
I.1 Généralités sur les variables aléatoires discrètes . . . 227
I.2 Loi d'une variable aléatoire discrète . . . 229
I.3 Loi de $f(X)$ . . . 230
I.4 Loi d'un vecteur aléatoire réel discret . . . 231
I.5 Variables indépendantes . . . 232
II Espérance et variance . . . 235
II.1 Espérance d'une variable positive . . . 235
II.2 Espérance d'une variable complexe . . . 236
II.3 Formule de transfert et conséquences . . . 236
II.4 Moment d'ordre 2 et variance d'une variable réelle . . . 238
II.5 Covariance . . . 240
II.6 Variance d'une somme . . . 242
III Variables aléatoires à valeurs dans $\mathbb{N}$, lois usuelles . . . 242

III.1 Fonction génératrice d'une variable à valeurs dans $\mathbb{N}$ . . . 243
III.2 Rappels sur les lois usuelles finies . . . 244
III.3 Loi géométrique . . . 246
III.4 Loi de Poisson . . . 247
III.5 Stabilités (HP classique) . . . 248
IV Inégalités probabilistes . . . 248
IV.1 Inégalités de Markov et de Bienaymé-Tchebychev . . . 248
IV.2 Loi faible des grands nombres . . . 249

16 Calcul différentiel et optimisation **251**
I Compléments sur les fonctions vectorielles . . . 252
I.1 Extention des notations o et O . . . 252
I.2 Dérivation de fonctions vectorielles d'une variable réelle . . . 253
I.3 Fonctions vectorielles de classe $\mathcal{C}^k$ . . . 256
II Dérivées partielles et différentielles . . . 257
II.1 Dérivées selon un vecteur . . . 257
II.2 Différentielles . . . 261
II.3 Gradient . . . 264
II.4 Opérations sur les applications différentiables . . . 266
III Applications de classe $\mathcal{C}^n$ . . . 269
III.1 Applications de classe $\mathcal{C}^1$ . . . 269
III.2 Plan tangent et noyau de la différentielle . . . 272
III.3 Applications de classe $\mathcal{C}^k$ . . . 275
III.4 Hessienne et formule de Taylor-Young à l'ordre 2 . . . 278
III.5 Équations aux dérivées partielles, EDP . . . 279
IV Optimisation . . . 280
IV.1 Principes généraux . . . 280
IV.2 Condition nécessaire du premier ordre . . . 281
IV.3 Optimisation sous contrainte . . . 283
IV.4 Étude à l'ordre 2 . . . 284

17 Intégration vectorielle et équations différentielles linéaires **287**
I Compléments d'intégration . . . 287
I.1 Définition et propriétés générales . . . 287
I.2 Théorème fondamental . . . 289
I.3 Formules de Taylor . . . 290
II Propriétés générales des EDL . . . 291
II.1 Notations et définitions . . . 291
II.2 Structure de l'ensemble des solutions . . . 292
II.3 Forme normalisée . . . 294
III EDL d'ordre 1 . . . 294
III.1 Situation . . . 294
III.2 Problème de Cauchy et espace des solutions . . . 295
III.3 Méthodes de résolution dans le cas scalaire . . . 297
III.4 Méthodes de résolution dans le cas vectoriel à coefficients constants . . . 299
IV EDL scalaires d'ordre supérieur . . . 300
IV.1 Comment se ramener à l'ordre 1 . . . 300
IV.2 Équations à coefficients constants . . . 301
IV.3 EDL d'ordre 2 . . . 302

1

Suites numériques – rappels et compléments

Nous rappelons dans ce chapitre les définitions et propriétés importantes sur les suites, ainsi que quelques aspects techniques et calculatoires en fin de chapitre. Il s'agit essentiellement de révisions du cours de MPSI, mais nous profitons de ce chapitre de révision pour introduire des notions fondamentales en analyse, comme les notions topologiques.
L'objectif n'est pas de refaire le cours de MPSI, mais uniquement des rappels. Ainsi, en particulier, les démonstrations, et les exemples de base sont à reprendre dans votre cours de MPSI. Certaines démonstrations importantes seront néanmoins reprises rapidement.
Nous rappelons que $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

I Topologie réelle et complexe

I.1 Intervalles

Définition 1.1.1 – Intervalle de $\mathbb{R}$

Un intervalle I de $\mathbb{R}$ est un sous-ensemble convexe I de $\mathbb{R}$, c'est-à-dire tel que :

$$\forall (a,b) \in I^2,\ \forall x \in \mathbb{R},\ a \leqslant x \leqslant b \Longrightarrow x \in I.$$

On rappelle que la propriété de la borne supérieure permet de décrire tous les intervalles :

Théorème 1.1.2 – Inventaire des intervalles réels

Tout intervalle I de $\mathbb{R}$ est d'une des formes suivantes, pour certaines valeurs réelles a et b :

- $[a,b] = \{x \in \mathbb{R},\ a \leqslant x \leqslant b\}$, $a \leqslant b$;
- $]a,b[= \{x \in \mathbb{R},\ a < x < b\}$, $a < b$;
- $[a,b[= \{x \in \mathbb{R},\ a \leqslant x < b\}$, $a < b$;
- $]a,b] = \{x \in \mathbb{R},\ a < x \leqslant b\}$, $a < b$;
- $[a,+\infty[= \{x \in \mathbb{R},\ x \geqslant a\}$;
- $]a,+\infty[= \{x \in \mathbb{R},\ x > a\}$;
- $]-\infty,b] = \{x \in \mathbb{R},\ x \leqslant b\}$;
- $]-\infty,b[= \{x \in \mathbb{R},\ x < b\}$;
- $]-\infty,+\infty[= \mathbb{R}$;
- $\varnothing$.

◁ **Éléments de preuve.**
Si I est non vide, sa borne inférieure a et sa borne supérieure b existent (éventuellement infinies). Montrer, par la propriété de convexité, que $]a,b[\subset I \subset [a,b]$. ▷

Remarques 1.1.3

1. Cette propriété est spécifique à $\mathbb{R}$. Par exemple, elle est fausse dans $\mathbb{Q}$ (avec $a, b \in \mathbb{Q}$).
2. La définition et le théorème s'étendent à $\overline{\mathbb{R}}$. Puisque $+\infty$ et $-\infty$ sont des éléments particuliers de $\mathbb{R}$, on a besoin de moins de cas pour décrire tous les intervalles : $[a, b]$, $]a, b]$, $[a, b[$ et $]a, b[$, avec $a, b \in \overline{\mathbb{R}}$. L'ensemble vide est par exemple obtenu en considérant $]0, 0[$, alors que $\mathbb{R} =]-\infty, +\infty[$ et $\overline{\mathbb{R}} = [-\infty, +\infty]$.
   La seule différence avce les intervalles de $\mathbb{R}$ est la possibilité d'inclure les bornes infinies.

I.2 Voisinages

On note $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. La notation $|z|$ désignera donc suivant le cas soit la valeur absolue dans $\mathbb{R}$, soit le module dans $\mathbb{C}$.

Définition 1.1.4 – Voisinage dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$

Soit $a \in \mathbb{K}$, et $V \subset \mathbb{K}$. On dit que V est un voisinage de a s'il existe $\varepsilon > 0$ tel que :

$$\forall y \in \mathbb{K}, \quad |y - x| < \varepsilon \Longrightarrow y \in V,$$

c'est-à-dire $B(a, \varepsilon) \subset V$, où $B(a, \varepsilon) = \{y \in \mathbb{R} \mid |y - a| < \varepsilon\}$ (boule ouverte de centre a et de rayon ε.
On note $\mathcal{V}(a) \subset \mathcal{P}(\mathbb{K})$ l'ensemble des voisinages de a.

Définition 1.1.5 – Voisinage de $\pm\infty$ dans $\mathbb{R}$

Soit $V \in \mathcal{P}(\mathbb{R})$.

- On dit que V est un voisinage de $+\infty$ si et seulement s'il existe b tel que $]b, +\infty[\subset V$.
- On dit que V est un voisinage de $a = -\infty$ si et seulement s'il existe b tel que $]-\infty, b[\subset V$.

On note $\mathcal{V}(+\infty)$ et $\mathcal{V}(-\infty)$ l'ensemble des voisinages de $+\infty$ et $-\infty$ respectivement.

Proposition 1.1.6 – Caractérisation des voisinages de $\mathbb{R}$

Soit $a \in \mathbb{R}$ et $V \subset \mathbb{R}$. Alors $V \in \mathcal{V}(a)$ si et seulement s'il existe un intervalle ouvert I tel que

$$a \in I \subset V.$$

Proposition 1.1.7 – Union et intersection de voisinages

Soit $a \in \overline{\mathbb{R}}$ ou $a \in \mathbb{C}$. Soit $I \neq \varnothing$, et $(V_i)_{i \in I}$ une famille de voisinages de a (dans $\mathbb{R}$ ou $\mathbb{C}$ selon le cas). Alors :

1. $\displaystyle\bigcup_{i \in I} V_i$ est un voisinage de a
2. Si I est fini, $\displaystyle\bigcap_{i \in I} V_i$ est un voisinage de a.

Exemple 1.1.8

Donner un exemple de famille de voisinages de 0 dans $\mathbb{R}$ dont l'intersection n'est pas un voisinage de 0.

Terminologie 1.1.9 – Propriété au voisinage de a

On dit qu'une propriété $\mathcal{P}(x)$ dépendant d'un élément $x \in \mathbb{K}$ est vérifiée au voisinage de a s'il existe

$V \in \mathcal{V}(a)$ telle que pour tout $x \in V$, $\mathcal{P}(x)$ est vraie.

Par exemple, dire qu'une fonction f est continue au voisinage de a signifie qu'il existe un voisinage V de a tel que pour tout $x \in V$, f soit continue en x, donc tel que f soit continue sur V.

II Convergence

II.1 Convergence, réexpression topologique

On rappelle les définitions suivantes :

Définition 1.2.1 – limite d'une suite, convergence, divergence

- Une suite $(u_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ admet une limite $\ell \in \mathbb{K}$ si et seulement si :

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geqslant N,\ u_n \in B(\ell, \varepsilon).$$

- Si $(u_n)_{n\in\mathbb{N}}$ admet une limite $\ell \in \mathbb{K}$, on dit qu'elle est convergente.
- Si $(u_n)_{n\in\mathbb{N}}$ n'admet aucun ℓ de $\mathbb{K}$ comme limite, alors on dit que $(u_n)_{n\in\mathbb{N}}$ est divergente (dans $\mathbb{K}$).

Si (u_n) admet une limite ℓ, on dit aussi que (u_n) tend vers ℓ lorsque n tend vers $+\infty$, et on écrira

$$\ell = \lim_{n\to+\infty} u_n, \qquad \text{ou} \qquad u_n \underset{n\to+\infty}{\longrightarrow} \ell.$$

On pourra omettre $n \to +\infty$ s'il n'y a pas d'ambiguïté sur la variable.

Définition 1.2.2 – Limite infinie d'une suite réelle

Soit une suite $(u_n)_{n\in\mathbb{N}}$ réelle.

- (u_n) admet la limite $+\infty$ si :

$$\forall A \in \mathbb{R}, \quad \exists N \in \mathbb{N}, \quad \forall n \geqslant N, \quad u_n \geqslant A.$$

- (u_n) admet la limite $-\infty$ si :

$$\forall A \in \mathbb{R}, \quad \exists N \in \mathbb{N}, \quad \forall n \geqslant N, \quad u_n \leqslant A.$$

On note respectivement $\lim u_n = +\infty$ et $\lim u_n = -\infty$.

On rappelle le résultat suivant, qui permet de donner un sens à la notation $\lim u_n$.

Théorème 1.2.3 – Unicité de la limite

Soit (u_n) une suite réelle ou complexe. Si elle existe, la limite de (u_n) est unique.

◁ **Éléments de preuve.**
Si ℓ_1 et ℓ_2 sont deux limites (non infinies) distinctes de (u_n), considérer $\varepsilon = |\ell_2 - \ell_1|$ dans les définitions pour obtenir une contradiction.
Discuter les cas infinis de façon similaire. ▷

Pour ne pas avoir constamment à traiter le cas d'une limite infinie de façon particulière, on donne une traduction topologique de la définition de la limite d'une suite, valide tout aussi bien pour des limites finies ou infinies.

Proposition 1.2.4 – Caractérisation topologique de la limite

Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et $a \in \mathbb{K}$ (ou éventuellement $a = \pm\infty$ si $\mathbb{K} = \mathbb{R}$). Alors les propositions suivantes sont équivalentes:

(i) $u_n \underset{a\to+\infty}{\longrightarrow} a$

(ii) $\forall V \in \mathcal{V}(a),\ \exists N \in \mathbb{N},\ \forall n \geqslant N,\ u_n \in V.$

On termine cette section par une propriété importante des suites d'entiers.

Proposition 1.2.5 – Suites convergentes à valeurs entières

Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathbb{Z}$. La suite (u_n) est convergente si et seulement si elle est stationnaire (*i.e.* constante à partir d'un certain rang).

◁ **Éléments de preuve.**
Réciproque évidente. Sens direct : considérer $\varepsilon < \frac{1}{2}$. ▷

II.2 Rappel des propriétés opératoires relatives à la convergence

Les règles opératoires usuelles sur les limites vous sont en général bien connues :

Théorème 1.2.6 – Opérations sur les limites finies

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles ou complexes convergentes, et λ et μ deux scalaires (réels ou complexes). Alors :

1. $(|u_n|)_{n\in\mathbb{N}}$ est convergente, et $\lim |u_n| = |\lim u_n|$;
2. $(\lambda u_n + \mu v_n)_{n\in\mathbb{N}}$ est convergente, et $\lim(\lambda u_n + \mu v_n) = \lambda \lim u_n + \mu \lim v_n$;
3. $(u_n v_n)_{n\in\mathbb{N}}$ est convergente, et $\lim(u_n v_n) = \lim u_n \cdot \lim v_n$;
4. si $\lim v_n \neq 0$, alors $v_n \neq 0$ à partir d'un certain rang ; ainsi, $\left(\dfrac{u_n}{v_n}\right)$ est définie à partir d'un certain rang, est convergente, et $\lim \dfrac{u_n}{v_n} = \dfrac{\lim u_n}{\lim v_n}$.

Théorème 1.2.7 – Opérations impliquant des limites infinies

Dans le cas où $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont des suites réelles, les résultats ci-dessus restent vrais si la limite de $(u_n)_{n\in\mathbb{N}}$ et/ou de $(v_n)_{n\in\mathbb{N}}$ est infinie, avec les règles arithmétiques suivantes (règles usuelles dans $\overline{\mathbb{R}}$) :

$$a + \infty = +\infty; \qquad +\infty + \infty = +\infty; \qquad a \cdot (+\infty) = (\mathrm{sg}(a))\infty \text{ (pour } a \neq 0);$$

$$(\pm\infty)\cdot(\pm\infty) = \pm\infty; \qquad \frac{1}{\pm\infty} = 0, \qquad \frac{1}{0^+} = +\infty, \qquad \frac{1}{0^-} = -\infty.$$

Avertissement 1.2.8 – Formes indéterminées

En revanche, les opérations arithmétiques suivantes ne sont pas définies, et donnent des formes indéterminées (ayez des exemples en tête) :

$$\infty - \infty; \qquad 0 \cdot \infty; \qquad \frac{\infty}{\infty}; \qquad \frac{0}{0}$$

Aux règles précédentes, nous ajoutons la suivante :

Proposition 1.2.9 – Produit $0 \times$ borné

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles ou complexes telles que $(u_n)_{n\in\mathbb{N}}$ tende vers 0 et $(v_n)_{n\in\mathbb{N}}$ soit bornée. Alors $(u_n v_n)_{n\in\mathbb{N}}$ tend vers 0.

La continuité de l'exponentielle et du logarithme (ou une propriété de composition des limites, pour les cas infinis non indéterminés) permet d'obtenir les règles d'exponentiation :

Proposition 1.2.10 – Passage à la limite pour les puissances

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles, (u_n) étant de plus strictement positive. On suppose que $u_n \to \ell$ et $v_n \to \ell'$, et que $(\ell, \ell') \notin \{(+\infty, 0), (0, 0), (1, +\infty), (1, -\infty)\}$. Alors :

$$u_n^{v_n} \longrightarrow \ell^{\ell'}.$$

Dans cette proposition, les règles d'exponentiation ont été étendues par les opérations suivantes dans $\overline{\mathbb{R}}$:

- $a^{+\infty} = 0$ si $a \in [0, 1[$,
- $a^{+\infty} = +\infty$ si $a \in]1, +\infty]$
- $a^{-\infty} = +\infty$ si $a \in [0, 1[$,
- $a^{-\infty} = 0$ si $a \in]1, +\infty]$,
- $0^b = 0$ si $b \in]0, +\infty]$.

Avertissement 1.2.11 – Formes indéterminées pour les puissances

On notera les formes indéterminées relatives aux exponentiations :

$$\infty^0, \qquad 0^0 \qquad \text{et} \qquad 1^\infty.$$

II.3 Comportement des limites vis-à-vis des inégalités

Les suites considérées dans cette section sont réelles. Pour les suites complexes, l'utilisation de résultats de ce type nécessite de séparer les études de la partie réelle et de la partie imaginaire.

Théorème 1.2.12 – Conservation des inégalités larges

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles, et $N \in \mathbb{N}$ tels que : $\forall n \geqslant N,\ u_n \leqslant v_n$. Alors, si $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ admettent des limites dans $\overline{\mathbb{R}}$, $\lim\limits_{n\to+\infty} u_n \leqslant \lim\limits_{n\to+\infty} v_n$.

$\lhd$ **Éléments de preuve.**

Par l'absurde. $\rhd$

Avertissement 1.2.13

Les inégalités strictes ne se conservent pas !

Théorème 1.2.14 – Théorème de convergence par encadrement

Soit $(u_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$, trois suites réelles, et $N \in \mathbb{N}$, tels que : $\forall n \geqslant N,\ u_n \leqslant v_n \leqslant w_n$. Si $(u_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ convergent toutes deux vers une même limite finie, alors $(v_n)_{n\in\mathbb{N}}$ converge aussi, et

$$\lim v_n = \lim u_n = \lim w_n.$$

Remarque 1.2.15

C'est avant tout un théorème d'existence de la limite. La valeur vient en bonus. Ne pas le rédiger comme un double passage à la limite dans les deux inégalités, qui ne justifie pas correctement cette existence.

Théorème 1.2.16 – Théorème de divergence par minoration ou majoration

Soit $(u_n)n \in \mathbb{N}$, $(v_n)_{n\in\mathbb{N}}$ et $N \in \mathbb{N}$ tels que pour tout $n \geqslant N$, $u_n \leqslant v_n$.

1. Si $(u_n)_{n\in\mathbb{N}}$ tend vers $+\infty$, alors $(v_n)_{n\in\mathbb{N}}$ tend vers $+\infty$.
2. Si $(v_n)_{n\in\mathbb{N}}$ tend vers $-\infty$, alors $(u_n)_{n\in\mathbb{N}}$ tend vers $-\infty$.

Le théorème de minoration s'utilise parfois en comparant à des suites dont on connaît bien la nature, par exemple les suites géométriques. La divergence ou la convergence des suites géométriques étant rapide, cela ne s'applique toutefois que pour des suites ayant un comportement assez marqué.
Voici une façon assez commode de faire une comparaison à une suite géométrique. Elle permet également de comparer les séries associées (critère de d'Alembert).

Méthode 1.2.17 – Comparaison à une suite géométrique

- Étudier l'existence et le cas échéant la valeur de $\lim \left|\frac{u_{n+1}}{u_n}\right|$.
- En cas d'existence, notons $\ell = \lim \left|\dfrac{u_{n+1}}{u_n}\right|$.
  - Si $\ell < 1$, on peut majorer à partir d'un certain rang $(|u_n|)_{n\in\mathbb{N}}$ par une suite géométrique de raison $r \in]\ell, 1[$, et on en déduit que $(u_n)_{n\in\mathbb{N}}$ tend vers 0.
  - Si $\ell > 1$, on peut minorer à partir d'un certain rang $(|u_n|)_{n\in\mathbb{N}}$ par une suite géométrique de raison $r \in]1, \ell[$, donc $(|u_n|)_{n\in\mathbb{N}}$ tend vers $+\infty$.
  - Si $\ell = 1$, on ne peut pas conclure.

Exemples 1.2.18

Limite de $\frac{a^n}{n!}$, $a \in \mathbb{C}$.

II.4 Propriétés des suites monotones

Théorème 1.2.19 – Théorème de la convergence monotone

Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle.

1. Si $(u_n)_{n\in\mathbb{N}}$ est croissante et majorée, alors $(u_n)_{n\in\mathbb{N}}$ converge dans $\mathbb{R}$.
2. Si $(u_n)_{n\in\mathbb{N}}$ est croissante et non majorée, alors $(u_n)_{n\in\mathbb{N}}$ diverge vers $+\infty$.
3. Si $(u_n)_{n\in\mathbb{N}}$ est décroissante et minorée, alors $(u_n)_{n\in\mathbb{N}}$ converge dans $\mathbb{R}$.
4. Si $(u_n)_{n\in\mathbb{N}}$ est décroissante et non minorée, alors $(u_n)_{n\in\mathbb{N}}$ diverge vers $-\infty$

De plus, dans le cas croissant, on a alors $\lim\limits_{n\to+\infty} u_n = \sup\limits_{n\in\mathbb{N}} u_n$.

Définition 1.2.20 – Suites adjacentes

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles. On dit qu'elles sont adjacentes si et seulement si :

1. l'une est croissante et l'autre décroissante ;
2. $(v_n - u_n)_{n\in\mathbb{N}}$ tend vers 0.

Théorème 1.2.21 – Théorème des suites adjacentes

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles adjacentes. Alors :

(i) $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ convergent, et $\lim\limits_{n\to+\infty} u_n = \lim\limits_{n\to+\infty} v_n$,

(ii) Pour tout $n \in \mathbb{N}$, $\quad |v_n - \ell| \leqslant |v_n - u_n|$.

III Comparaisons asymptotiques

Proposition/Définition 1.3.1 – Domination

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles. Les propositions suivantes sont équivalentes:

(i) $\exists M \in \mathbb{R},\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ |u_n| \leqslant M|v_n|$

(ii) il existe une suite (μ_n) et $n_0 \in \mathbb{N}$ telle que pour tout $n \geqslant n_0$, $u_n = \mu_n v_n$, et (μ_n) est bornée

Si ces propriétés sont satisfaites, on dit que (u_n) est dominée par (v_n), et on note $u_n = O(v_n)$.

Proposition/Définition 1.3.2 – Négligeabilité

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles. Les propositions suivantes sont équivalentes:

(i) $\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ |u_n| \leqslant \varepsilon|v_n|$;

(ii) il existe une suite $(\varepsilon_n)_{n\in\mathbb{N}}$ et un entier n_0 tels que :

$$\forall n \geqslant n_0,\ \ u_n = \varepsilon_n v_n \qquad \text{et} \qquad \lim \varepsilon_n = 0.$$

Si ces propriétés sont satisfaites, on dit que (u_n) est négligeable devant (v_n), et on note $u_n = o(v_n)$.

Définition 1.3.3 – Équivalence

Deux suites réelles (u_n) et (v_n) sont dites *équivalentes* s'il existe une suite (α_n) et un entier n_0 tels que

$$\forall n \geqslant n_0,\ u_n = \alpha_n v_n \qquad \text{et} \qquad \lim \alpha_n = 1.$$

On note $u_n \underset{+\infty}{\sim} v_n$ ou $u_n \sim v_n$.

La propriété suivante permet d'écrire une équivalence sous forme d'un développement asymptotique, ce qui s'avère très souvent pratique d'un point de vue technique, pour contourner certains problèmes techniques liés aux équivalents (sommes, compositions)

Proposition 1.3.4 – Caractérisation de l'équivalence par la négligeabilité

$u_n \underset{+\infty}{\sim} v_n \iff u_n = v_n + o(v_n)$.

Remarque 1.3.5

Si (v_n) ne s'annule pas à partir d'un certain rang, ces propriétés peuvent être caractérisées par l'étude de $\dfrac{v_n}{u_n}$:

- $\left(\dfrac{u_n}{v_n}\right)$ bornée $\iff$ $u_n = O(v_n)$;
- $\dfrac{u_n}{v_n} \longrightarrow 0 \iff u_n = o(v_n)$;
- $\dfrac{u_n}{v_n} \longrightarrow 1 \iff u_n \sim v_n$.

On rappelle les règles de manipulation suivantes, qui se démontrent facilement avec le point de vue séquentiel (la deuxième caractérisation).

Propriétés 1.3.6 – Transitivités strictes et larges de o et O

1. Si $u_n = O(v_n)$ et $v_n = O(w_n)$, alors $u_n = O(w_n)$.
2. Si $u_n = o(v_n)$ et $v_n = o(w_n)$, alors $u_n = o(w_n)$.
3. Si $u_n = o(v_n)$ et $v_n = O(w_n)$, alors $u_n = o(w_n)$.

4. Si $u_n = O(v_n)$ et $v_n = o(w_n)$, alors $u_n = o(w_n)$.

Propriétés 1.3.7 – sommes de o et O

1. Si $u_n = o(w_n)$ et $v_n = o(w_n)$, alors $u_n + v_n = o(w_n)$.
2. Si $u_n = O(w_n)$ et $v_n = O(w_n)$, alors $u_n + v_n = O(w_n)$.
3. Si $u_n = o(w_n)$ et $v_n = O(w_n)$, alors $u_n + v_n = O(w_n)$.
4. Si $u_n = O(w_n)$ et $v_n = o(w_n)$, alors $u_n + v_n = O(w_n)$.

Propriétés 1.3.8 – produits de o et O

1. Si $u_n = o(w_n)$ et $v_n = o(x_n)$, alors $u_n v_n = o(w_n x_n)$.
2. Si $u_n = O(w_n)$ et $v_n = o(x_n)$, alors $u_n v_n = o(w_n x_n)$.
3. Si $u_n = o(w_n)$ et $v_n = O(x_n)$, alors $u_n v_n = o(w_n x_n)$.
4. Si $u_n = O(w_n)$ et $v_n = O(x_n)$, alors $u_n v_n = O(w_n x_n)$.
5. En particulier, $u_n = w_n o(x_n) \iff u_n = o(w_n x_n)$ et $u_n = w_n O(x_n) \iff u_n = O(w_n x_n)$.
   (Les o et O sont multiplicatifs : on peut rentrer ou sortir un facteur multiplicatif).

Proposition 1.3.9 – Équivalents de produits, quotients, puissances

Soit (u_n), (v_n), (u'_n) et (v'_n) quatre suites réelles.

1. Si $u_n \sim u'_n$ et $v_n \sim v'_n$, alors $u_n v_n \sim u'_n v'_n$.
2. Si de plus v_n est non nulle à partir d'un certain rang, alors $\dfrac{u_n}{v_n} \sim \dfrac{u'_n}{v'_n}$.
3. Si $u_n \sim u'_n$ et si a est un réel fixé, alors $(u_n)^a \sim (u'_n)^a$.

Avertissement 1.3.10

- Ne pas sommer des équivalents (problème de compensation de la partie principal)
- Ne pas composer des équivalents par une fonction.

Exemples 1.3.11

Donner des contre-exemples pour ces deux avertissements.

Pour conourner ces problèmes, revenir à des techniques de développements limités ou asymptotiques. Enfin, deux propriétés importantes des équivalents :

Théorème 1.3.12 – Conservation des limites par équivalence

Si $u_n \underset{+\infty}{\sim} v_n$, et si $(u_n)_{n\in\mathbb{N}}$ converge vers ℓ dans $\overline{\mathbb{R}}$, alors $(v_n)_{n\in\mathbb{N}}$ converge, et sa limite est ℓ.

Proposition 1.3.13 – Conservation du signe

Soit (u_n) et (v_n) deux suites équivalentes. Alors il existe n_0 tels que pour tout $n \geqslant n_0$, u_n et v_n soient de même signe.

Et pour terminer, la piqûre de rappel des équivalents classiques. Je les donne sous leur version fonctionnelle, ce qui signifie ici qu'on peut remplacer x par n'importe quelle suite convergeant vers 0 (les équivalents classiques étant tous donnés au voisinage de 0).
Ne pas hésiter sur ces équivalents !

Proposition 1.3.14 – Équivalents classiques

1. $\ln(1+x) \underset{0}{\sim} x$;
2. $e^x - 1 \underset{0}{\sim} x$;
3. $(1+x)^a - 1 \underset{0}{\sim} ax \quad (a \neq 0)$;
4. $\sin(x) \underset{0}{\sim} x$;
5. $\cos(x) - 1 \underset{0}{\sim} -\dfrac{x^2}{2}$;
6. $\tan(x) \underset{0}{\sim} x$;
7. $\mathrm{sh}(x) \underset{0}{\sim} x$;
8. $\mathrm{ch}(x) - 1 \underset{0}{\sim} \dfrac{x^2}{2}$;
9. $\mathrm{th}(x) \underset{0}{\sim} x$;
10. $\mathrm{Arcsin}(x) \underset{0}{\sim} x$;
11. $\mathrm{Arctan}(x) \underset{0}{\sim} x$;

Et un petit dernier qui a un statut un peu particulier, car spécifique aux suites et ne découlant pas de techniques de dérivation :

Théorème 1.3.15 – Formule de Stirling

On a : $n! \underset{+\infty}{\sim} \sqrt{2\pi n}\left(\dfrac{n}{e}\right)^n$.

IV Suites récurrentes

Nous terminons ce chapitre par quelques rappels sur les propriétés et l'étude de certaines suites récurrentes.

IV.1 Suites récurrentes linéaires

Définition 1.4.1 – Suites récurrentes linéaires

Soit $(u_n)_{n\in\mathbb{N}}$. On dit que (u_n) est une *suite récurrente linéaire d'ordre k* si (u_n) vérifie une relation du type :

$$\forall n \in \mathbb{N}, \quad u_{n+k} = a_{k-1,n}u_{n+k-1} + \cdots + a_{1,n}u_{n+1} + a_{0,n}u_n.$$

où $((a_{0,n}), \dots, (a_{(k-1,n)})) \in (\mathbb{C}^{\mathbb{N}})^k$.

Une suite récurrente d'ordre k est entièrement déterminée par sa relation de récurrence et la donnée de ses k premiers termes. Plus précisément :

Théorème 1.4.2 – Structure et dimension de l'ensemble des SRL

Soit, pour $i \in [\![0, k-1]\!]$, $(a_{i,k}) \in \mathbb{C}^{\mathbb{N}}$. On pose $\mathcal{S}$ l'ensemble des suites $(u_n)_{n\in\mathbb{N}}$ vérifiant la relation de récurrence :

$$\forall n \in \mathbb{N}, \quad u_{n+k} = a_{k-1,n}u_{n+k-1} + \cdots + a_{1,n}u_{n+1} + a_{0,n}u_n.$$

Alors $\mathcal{S}$ est un $\mathbb{C}$-espace vectoriel, et l'application

$$\Phi : \begin{cases} \mathcal{S} & \longrightarrow \mathbb{C}^k \\ (u_n) & \longmapsto (u_0, \dots, u_{k-1}) \end{cases}$$

est un isomorphisme. En particulier, $\mathcal{S}$ est de dimension k.

Dans le cas où les suites $(a_{i,n})$ sont constantes, on parle de suite récurrente linéaire d'ordre k à coefficients constants. On dispose dans ce cas de techniques d'explicitation, grâce aux racines du polynôme caractéristique.

Définition 1.4.3 – Polynôme caractéristique d'une récurrence linéaire

Soit $(u_n)_{n\in\mathbb{N}}$ une suite récurrente linéaire à coefficients constants, de relation

$$u_{n+k} = a_{k-1}u_{n+k-1} + \cdots + a_1u_{n+1} + a_0u_n \quad i.e. \quad u_{n+k} - a_{k-1}u_{n+k-1} - \cdots - a_1u_{n+1} - a_0u_n = 0$$

Le *polynôme caractéristique* associé à la suite récurrente $(u_n)_{n\in\mathbb{N}}$ est le polynôme

$$\chi(X) = X^k - a_{k-1}X^{k-1} - \cdots - a_1X - a_0.$$

Théorème 1.4.4 – Explicitation des SRL, HP si $n \geqslant 2$

Soit $\mathcal{S}$ l'ensemble des suites (u_n) vérifiant une relation de récurrence linéaire dont le polynôme caractéristique est χ. Soit $\lambda_1, \dots, \lambda_p$ les racines de χ (dans $\mathbb{C}$), de multiplicité $\alpha_1, \dots, \alpha_p$ respectivement. Alors $\mathcal{S}$ est l'ensemble des suites telles que

$$\forall n \in \mathbb{N}, \quad u_n = \sum_{i=1}^{p} P_i(n)\lambda_i^n,$$

où pour tout $i \in [\![1, p]\!]$, $P_i \in \mathbb{C}_{\alpha_i - 1}[X]$.

◁ **Éléments de preuve.**

Hors-programme dans le cas général. Voir cours de MPSI pour le cas particulier $k \leqslant 2$ (ou adapter celle-ci). La preuve utilise des polynômes d'endomorphisme. Reprenez-là lorsque vous serez un peu plus à l'aise sur le sujet.

- Considérer $\chi(D)$ où D est l'opérateur qui à (u_n) associe (u_{n+1}), et le factoriser. On est ramené à l'étude de $(D - \lambda_i \mathrm{Id})^{\alpha_i}((P_i(n)\lambda_i^n))$, et on remarque qu'à chaque itération, le degré du polynôme diminue strictement. Cela prouve que ces suites sont bien solutions.
- La réciproque se fait par un argument de dimension, en montrant que les $(n^j\lambda_i^n)_{n\in\mathbb{N}}$, $j < \alpha_i$, forment une famille libre (par exemple en exploitant un équivalent en $+\infty$).

▷

IV.2 Suites récurrentes linéaires perturbées

Une suite récurrente linéaire perturbée est une suite définie par une récurrence qui ressemble à une récurrence linéaire, à laquelle on ajoute une perturbation sous forme d'une suite (b_n). Ainsi, elle vérifie une relation du type :

$$(E) \ : \ \forall n \in \mathbb{N}, \quad u_{n+k} - \sum_{i=0}^{k-1} a_{i,n}u_{n+i} = b_n.$$

On note également (EH) la relation de récurrence homogène associée (obtenue en remplaçant b_n par 0).

Théorème 1.4.5 – Structure de l'ensemble des SRL perturbées

L'ensemble $\mathcal{S}$ des suites vérifiant la relation de récurrence (E) est un espace affine dirigé par l'espace $\mathcal{S}_H$ des suites vérifiant (EH). Ainsi, si $(u_n)_{n\in\mathbb{N}}$ est une solution particulière, toute autre suite (v_n) vérifiant (E) vérifiera

$$\forall n \in \mathbb{N}, \quad v_n = u_n + w_n, \quad (w_n) \in \mathcal{S}_H.$$

◁ **Éléments de preuve.**

Il s'agit d'une fibre de l'endomorphisme de $\mathbb{C}^{\mathbb{N}}$ défini par $(u_n) \mapsto u_{n+k} - \sum_{i=0}^{k-1} a_{i,n}u_{n+i}$ ▷

Proposition 1.4.6 – Explicitation d'une SRL perturbée par $(P(n)\lambda^n)$

Soit (EH) une relation de récurrence linéaire à coefficients constants, de polynôme caractéristique χ, et (E) la relation correspondante, perturbée par $(b_n) = P(n)\lambda^n$. Alors il existe une solution particulière de (E) de la forme

$$u_n = n^\alpha Q(n)\lambda^n,$$

où $\alpha = \mathrm{mult}_\chi(\lambda)$ et $\deg(Q)) \leqslant \deg(P)$.

◅ **Éléments de preuve.**
Hors-programme. Encore une fois, utiliser $\chi(D)$. ▻

Remarque 1.4.7

1. La résolution des suites arithmético-géométriques $u_{n+1} = au_n + b$, $a \neq 1$, est un cas particulier de ces deux résultats plus général : il existe une solution constante c, vérifiant $c = ac + b$. C'est donc le point fixe. La solution générale s'écrit alors sous la forme $\lambda a^n + c$.
2. Avant cela, la résolution des suites arithmétiques $u_{n+1} = u_n + a$ en est aussi un cas particulier : la proposition donne l'existence d'une solution particulière sous la forme cn (il faut multiplier par n car 1 est racine de multiplicité 1 du polynôme caractéristique). En remplaçant, on trouve facilement $c = a$. La solution générale est alors $\lambda 1^n + an$, et on ajuste avec les termes initiaux.

Exemple 1.4.8

Un exemple un peu moins trivial. Expliciter (u_n) tel que

$$u_0 = 2, \quad u_1 = 1, \quad u_{n+2} = 5u_{n+1} - 6u_n + (n+1)2^n.$$

IV.3 Suites récurrentes $u_{n+1} = f(u_n)$

On rappelle ici les grands principes d'étude des systèmes dynamiques discrets, c'est-à-dire des suites récurrentes non linéaires d'ordre 1, du type $u_{n+1} = f(u_n)$, où f sera supposée continue sur une partie D de $\mathbb{R}$.

Méthode 1.4.9 – Étude d'une suite récurrente

- Justifier la bonne définition de la suite, par exemple en exhibant un intervalle stable (voir ci-dessous)
- Étudier l'existence de la limite : étudier la monotonie, et obtenir l'existence de la limite à l'aide par exemple du théorème de la limite monotone.
- Une fois assurée l'existence de la limite, déterminer les valeurs possibles de cette limite ℓ en passant à la limite dans la relation de récurrence :

  Si f est continue, soit ℓ vérifie $f(\ell) = \ell$, soit ℓ est un bord ouvert du domaine de f.
- Déterminer, parmi l'ensemble des valeurs possibles de ℓ laquelle est la bonne.
- Déterminer les points fixes de f peut déjà s'avérer utile pour la première étape (aide à la recherche d'intervalles stables, de majorants et de minorants). Commencer par cela peut être efficace.

Un objet incournable dans l'étude des suites récurrentes est :

Définition 1.4.10 – Intervalle stable

On dit qu'un intervalle I est stable par f si f est définie sur I (i.e. $I \subset D_f$) et $f(I) \subset I$.

Théorème 1.4.11 – CS pour que (u_n) soit bien définie

Si I est un intervalle stable par f, et si $u_0 \in I$, alors $(u_n)_{n\in\mathbb{N}}$ est bien définie.

◁ **Éléments de preuve.**
Montrer par récurrence que u_n est défini, et $u_n \in I$. Comme I est dans le domaine, on peut alors continuer. ▷

Évidemment, si on parvient à trouver un indice N tel que u_N soit dans un intervalle stable, on parvient à la même conclusion.
Les intervalles stables permettent aussi souvent d'étudier les variations de $(u_n)_{n\in\mathbb{N}}$. La situation la plus simple est la suivante (et assez facile à détecter graphiquement) :

Proposition 1.4.12 – Étude des variations lorsque $f - \mathrm{id}$ est de signe constant

Supposons $f - \mathrm{id}$ de signe constant sur un intervalle stable I, et $u_0 \in I$. Alors (u_n) est monotone, de sens de monotonie déterminé par le signe de $f - \mathrm{id}$.

◁ **Éléments de preuve.**
D'après ce qui précède, pour tout n, $u_n \in I$. Appliquer f pour comparer u_{n+1} à u_n. ▷

Proposition 1.4.13 – Étude des variations lorsque f est croissante

Supposons f croissante sur un intervalle stable I, et $u_0 \in I$ (peut s'adapter si on n'a pas cette inclusion dès le rang initial). Alors (u_n) est monotone. On trouve le sens de monotonie en comparant u_0 et u_1.

◁ **Éléments de preuve.**
Tout d'abord, tous les termes u_n sont dans I. On peut alors propager par récurrence l'inégalité initiale entre u_0 et u_1, en appliquant la fonction croissante f. ▷

Proposition 1.4.14 – Étude des variations lorsque f est décroissante

Supposons f décroissante sur un intervalle stable I, et $u_0 \in I$ (peut s'adapter si on n'a pas cette inclusion dès le rang initial). Alors (u_{2n}) et (u_{2n+1}) sont monotones, de sens de monotonie opposé. On trouve les sens de monotonie en comparant u_0 et u_2.

◁ **Éléments de preuve.**
Appliquer le résultat précédent à $f \circ f$, décrivant la récurrence associée aux suites (u_{2n}) et (u_{2n+1}). ▷

Remarque 1.4.15

Ainsi, si f est décroissante sur un intervalle stable, les deux suites extraites (u_{2n}) et (u_{2n+1}) sont monotones de sens de variation opposés. Cela ne suffit pas pour prouver que $(u_n)_{n\in\mathbb{N}}$ converge. Pour cela, il faut encore s'assurer que (u_{2n}) et (u_{2n+1}) ont même limite. Les limites de ces deux suites sont à rechercher parmi les points fixes de $f \circ f$. Si $f \circ f$ a deux points fixes distincts, on peut très bien avoir convergence de (u_{2n}) vers l'un des deux et de (u_{2n+1}) vers l'autre. Essayez de construire un exemple graphique.

Avertissement 1.4.16

Les points fixes de f sont point fixes de $f \circ f$, mais il peut exister des points fixes de $f \circ f$ qui ne sont pas points fixes de f.

Exemple 1.4.17

Étude de la suite $u_0 \in \mathbb{R}$, $\forall n \in \mathbb{N}, \quad u_{n+1} = u_n^2 - u_n$.

Pour terminer, nous donner deux résultats hors-programme, mais à savoir refaire, permettant l'étude de la vitesse de convergence (c'est-à-dire estimant le comportement asymptotique de la suite, soit sous forme d'une domination, soit sous forme d'un équivalent).

Proposition 1.4.18 – Vitesse de convergence en un point fixe attractif, HP

Soit f une fonction définie et de classe $\mathcal{C}^1$ sur un intervalle I, et soit $a \in I$ un point fixe de f tel que $|f'(a)| < 1$.

1. Alors il existe ε tel que s'il existe n_0 tel que $|u_{n_0} - a| < \varepsilon$, alors $u_n \to a$. On dit que a est un point fixe attractif.
2. Si tel est le cas, on a alors, pour tout $k \in]|f'(a)|, 1[$, $u_n = O(k^n)$.

Le deuxième résultat s'applique lorsque $f'(a) = 1$. Dans ce cas, il est clair graphiquement que la convergence sera beaucoup plus lente. On peut, sous réserve quelques hypothèses supplémentaires, trouver un équivalent de (u_n). Plutôt que de donner un résultat tout finalisé, nous donnons ici plutôt une méthode. Nous avons besoin pour cela d'un lemme archi-classique, probablement vu en MPSI déjà.

Lemme 1.4.19 – Théorème de la moyenne de Cesàro, cas convergent

Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle ou complexe, et

$$\forall n \geqslant 1, \quad M_n = \frac{1}{n} \sum_{k=0}^{n-1} u_n$$

la suite de ses moyennes. Alors si (u_n) converge vers ℓ, (M_n) aussi.

Méthode 1.4.20 – Equivalent de (u_n) convergeant vers a tel que $f'(a) = 1$

- Sous des hypothèses permettant de mener les calculs à leurs termes (hypothèses que je ne donne pas explicitement), trouver α tel que
$$u_{n+1}^\alpha - u_n^\alpha = u_n^\alpha \left(\left(\frac{f(u_n)}{u_n} \right)^\alpha - 1 \right)$$
admette une limite finie non nulle. On pourra pour cela faire un développement limité de f au voisinage de a, ou utiliser des équivalents classiques.
- Le théorème de Cesàro permet alors d'obtenir, par télescopage, un équivalent de $u_n^\alpha - u_0^\alpha$, duquel on déduit un équivalent de (u_n).

Exemple 1.4.21

Donner un équivalent de la suite définie par $u_0 = 1$ et $\forall n \in \mathbb{N}, \; u_{n+1} = \sin(u_n)$.

2

Cardinaux et dénombrabilité

Nous faisons dans ce chapitre quelques rappels sur les cardinaux finis, et leur utilisation dans des questions de dénombrement. Dans un second temps, nous quittons les révisions pour aborder le problème des cardinaux infinis, beaucoup moins intuitif, l'objectif étant l'étude des ensembles de même cardinal que $\mathbb{N}$, qu'on appelle ensembles dénombrables.

I Rappels sur les cardinaux finis

I.1 Ensembles finis, cardinaux

Nous rappelons la définition du cardinal d'un ensemble fini :

Définition 2.1.1 – Cardinal d'un ensemble fini

Soit E un ensemble.

- On dit que E est un ensemble fini s'il existe $n \in \mathbb{N}$ et une bijection $[\![1, n]\!] \to E$.
- L'entier n est alors unique, est appelé cardinal de E, et est noté $|E|$ ou $\mathrm{Card}(E)$ ou $\sharp E$.

◁ **Éléments de preuve.**
L'unicité de n provient du fait qu'il ne peut pas exister de bijection de $[\![1, n]\!]$ dans $[\![1, m]\!]$ si $n \neq m$, ce qui n'est pas si évident que cela à montrer si on ne dispose d'aucun outil de cardinalité. Cela peut se faire en montrant par récurrence sur n que si $m \neq n$ et $f : [\![1, n]\!] \to [\![1, m]\!]$, alors f n'est ni injective ni surjective. Pour passer de n à $n+1$, considérer $f(n+1)$. Quitte à composer par une transposition bijective, on peut considérer que $f(n+1) = m+1$, ce qui permet d'enlever n d'un côté, m de l'autre (avec une discussion à faire), pour se ramener à l'hypothèse de récurrence. ▷

Lemme 2.1.2 – Finitude des sous-ensembles de $[\![1, n]\!]$

Tout sous-ensemble F de $[\![1, n]\!]$ peut être mis en bijection avec un ensemble $[\![1, m]\!]$, avec $m \leqslant n$.

◁ **Éléments de preuve.**
On peut ranger les éléments de F dans l'ordre : $F = \{x_1 < x_2 < \cdots < x_m\}$. ▷

Proposition 2.1.3 – Sous-ensemble d'un ensemble fini

Soit F un sous-ensemble de E. Si E est fini, alors F aussi.

◁ **Éléments de preuve.**
Restreindre et corestreindre une bijection $E \to [\![1, n]\!]$ et utiliser le lemme. ▷

I.2 Rappel des propriétés calculatoires sur les cardinaux

Lemme 2.1.4 – Réexpression du cardinal

Soit E un ensemble fini, et A un sous-ensemble de E. Alors

$$|A| = \sum_{k \in E} \mathbb{1}_A(k).$$

◁ **Éléments de preuve.**

C'est évident intuitivement. Pour une bonne formalisation, par une bijection, on peut se ramener au cas où $A = [\![1, p]\!]$ et $E = [\![1, n]\!]$. ▷

Proposition 2.1.5 – Cardinal d'une union disjointe

Soit A, B, $A_1, \dots, A_n$ des ensembles finis.

1. Si $A \cap B = \varnothing$, alors $|A \uplus B| = |A| + |B|$.
2. Plus généralement, si pour tout $(i, j) \in [\![1, n]\!]^2$ tel que $i \neq j$, $A_i \cap A_j = \varnothing$, alors

$$|A_1 \uplus \cdots \uplus A_n| = |A_1| + \cdots + |A_n|.$$

◁ **Éléments de preuve.**

Le point 2 s'obtient du point 1 par récurrence. Le point 1 résulte du lemme et des propriétés des sommes ▷

Proposition 2.1.6 – Cardinal d'un complémentaire

Si $A \subset B$, alors $|\complement_B A| = |B| - |A|$.

◁ **Éléments de preuve.**

Écrire l'un des trois ensembles en jeu comme union disjointe des deux autres. ▷

Corollaire 2.1.7 – Cardinal d'un sous-ensemble

Si $A \subset B$, alors $|A| \leqslant |B|$, avec égalité si et seulement si $A = B$.

◁ **Éléments de preuve.**

Quel est le cardinal de $\complement_B A$ lorsqu'on est dans le cas d'égalité ? ▷

Proposition 2.1.8 – Cardinal d'une union quelconque

Soit A et B des ensembles finis. On a :

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

◁ **Éléments de preuve.**

On peut décomposer en $A \cup B = A \uplus (B \backslash (A \cap B))$ ▷

Proposition 2.1.9 – Cardinal d'un produit cartésien

1. Soit A et B deux ensembles finis. Alors $|A \times B| = |A| \times |B|$.
2. Plus généralement, soit $A_1, \dots, A_n$ des ensembles finis. Alors

$$|A_1 \times \cdots \times A_n| = \prod_{i=1}^{n} |A_i|.$$

◁ **Éléments de preuve.**
Le point 2 s'obtient du 1 par récurrence.
Pour le point 1, découper en tranche, ou utiliser une bijection de $[\![0, nm-1]\!]$ dans $[\![0, n-1]\!] \times [\![0, m-1]\!]$ définie par une opération arithmétique classique. ▷

I.3 Grands principes combinatoires

Par définition même :

Proposition 2.1.10 – Injectivité, surjectivité, bijectivité et cardinal

Soit E et F deux ensembles finis, et soit $f : E \longrightarrow F$ une application. Alors :

1. Si f est bijective, $\mathrm{Card}(E) = \mathrm{Card}(F)$.
2. Si f est injective, $\mathrm{Card}(E) \leqslant \mathrm{Card}(F)$
3. Si f est surjective, $\mathrm{Card}(E) \geqslant \mathrm{Card}(F)$

◁ **Éléments de preuve.**
Restreindre ou corestreindre de façon adéquate, de sorte à se ramener à une bijection. ▷

Cela incite, pour des dénombrements (ou calculs de cardinaux, à se ramener par une bijection, à des modèles simples et connus.

Méthode 2.1.11 – Combinatoire bijective

Pour dénombrer un ensemble fini E (c'est-à-dire pour déterminer son cardinal), une méthode classique consiste à le mettre en bijection avec un ensemble dont le cardinal est connu.

Exemple 2.1.12

Par définition des coefficients binomiaux, le cardinal de $\mathcal{P}_k([\![1, n]\!])$ est $\binom{n}{k}$. En déduire le nombre de l'ensemble des chemins monotones de $(0, 0)$ à (a, b) dans $\mathbb{N}^2$, chaque pas étant de longueur 1, et se faisant vers la droite ou vers le haut.

Méthode 2.1.13 – Tri, partitionnement, principe additif

Si on peut trier les objets de E en plusieurs paquets disjoints $E_1, \dots, E_n$ de cardinaux respectifs $a_1, \dots, a_n$, alors le cardinal de E est

$$E = \sum_{k=1}^{n} a_k.$$

Exemple 2.1.14

Formule sommatoire, ou lemme de la chaussette : $\displaystyle\sum_{k=O}^{p} \binom{n+k}{n} = \binom{n+k+1}{n+1}$.

Méthode 2.1.15 – Choix successifs, principe multiplicatif

Dénombrer un ensemble se fait souvent en s'intéressant à la façon de construire un élément, en décomposant sa construction de façon élémentaire en plusieurs étapes. Si un élément se construit en 2 étapes, et que :

- la première étape peut se faire de a façons différentes ;
- pour chaque choix de première étapes, la deuxième étape peut se faire de b façons différentes (le nombre b ne dépendant pas de la première étape) ;
- toutes les constructions faites ainsi sont distinctes

alors le nombre total d'éléments est ab

◁ **Éléments de preuve.**

Une formalisation passe par un partionnement de l'ensemble E dénombré en $\biguplus_{k\in[\![1,a]\!]} E_k$, où on a rangé les éléments à dénombrer dans des ensembles E_k, en les triant suivant le résultat de la première étape. La troisième hypothèse certifie que l'union est disjointe (ce qui est représenté par le signe + dans l'union), et la deuxième hypothèse assure que les E_k sont tous de cardinal b.

Cette formalisation est équivalente au lemme du berger, stipulant que si $f : E \to F$ est une surjection telle que les images réciproques $f^{-1}(\{y\})$ aient toutes même cardinal b, alors $|E| = b|F|$. ▷

Exemple 2.1.16

Dénombrement des injections, cardinal de $\mathfrak{S}_n$.

Remarque 2.1.17

Les deux méthodes peuvent être combinées, par exemple un tri, et pour chaque part du tri, une succession de choix.

Exemple 2.1.18

Formule de Vandermonde : $\displaystyle\sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k} = \binom{a+b}{n}$

Méthode 2.1.19 – Formalisation par couple, et lecture inverse

Une formalisation mathématique de deux choix successifs peut se faire sous forme d'un couple, en mettant E en bijection avec un ensemble F de couples (x, y) où x est le résultat du premier choix, y le résultat du deuxième choix. Il y a souvent des contraintes entre x et y (liés à la façon de procéder au second choix). Ces contraintes définissent une propriété satisfaite par x et y, qu'on note $\mathcal{P}(x, y)$. Ainsi

$$F = \{(x, y) \mid \mathcal{P}(x, y)\}$$

. Sous cette forme on peut voir que la contrainte peut souvent être retournée en lisant le couple dans l'autre sens : F et clariement en bijection avec

$$G = \{(y, x) \mid \mathcal{P}(x, y)\}$$

Cela revient à intervertir le rôle des deux étapes, donc l'ordre dans lequel on effectue la construction.

Exemple 2.1.20

La formule du capitaine, ou comité-président.

Méthode 2.1.21 – Démonstration ou interprétation combinatoire d'une formule

- Interpréter combinatoirement une formule permet d'en avoir une compréhension intuitive et naturelle, qui aide à la retenir ou a la retrouver.
- Cela permet également éventuellement de la démontrer, comme la formule de Vandermonde ci-dessus.
- Pour démontrer ou interpréter combinatoirement une formule, chercher un modèle adéquat, et dénombrer de deux façons différentes, par exemple de façon directe et en effectuant un tri, ou bien en effectuant des choix successifs dans des ordres différents etc.

Exemple 2.1.22

Interpréter combinatoirement la formule $\displaystyle\sum_{k=0}^{n}\binom{n}{k} = 2^n$.

Méthode 2.1.23 – Signes, principe de l'interrupteur

- Une formule combinatoire ne contient *a priori* que des termes positifs. Certaines formules contenant aussi des signes négatifs peuvent quand même se démontrer combinatoirement, en réarrageant l'égalité de sorte à passer les termes négatifs de l'autre côté. On a alors deux quantités à comparer, ce qu'on peut faire combinatoirement, par exemple en établissant des bijections.
- Une situation classique : des signes qui alternent devant une expression avec des coefficients binomiaux : il s'agit de comparer des ensembles ayant des cardinaux de parité différente : pour cela, faire entrer ou sortir un élément (interrupteur qu'on allume ou éteint). On illustre cela dans l'exemple qui suit, qui est à peut près le plus simple du genre.

Exemple 2.1.24

Montrer combinatoirement que $\displaystyle\sum_{k=0}^{n}(-1)^k\binom{n}{k} = 0$ si $n > 0$.

II Dénombrabilité

II.1 Cardinaux infinis

On ne peut pas définir d'emblée (en tout cas pas simplement) de cardinal d'un ensemble fini. Cantor fait une construction (compliquée) qui permet de définir des cardinaux (ou plutôt des classes cardinales) de façon explicite, mais ce n'est pas l'objet de ce cours. Nous nous contentons de définir la condition d'égalité des cardinaux.

Définition 2.2.1 – Ensembles de même cardinal

Soit E et F deux ensembles.

- On dit que E et F ont même cardinal s'il existe une bijection $f : E \to F$. On dit aussi que E est équipotent à F.
- On dit que E est subpotent à F (ou que le cardinal de E est strictement inférieur à celui de F) s'il existe une injection de E dans F.

Remarque 2.2.2 – (HP)

- On peut montrer que la relation de subpotence vérifie les propriétés d'une relation d'ordre sur les classes cardinales. L'antisymétrie provient d'un théorème classique dû à Cantor et Berstein, affirmant que s'il existe une injection $f : E \to F$, et une injection $g : F \to E$, alors il existe une bijection $E \to F$.
- Avec l'axiome du choix, on peut montrer que cette « relation d'ordre » est totale.

Attention, certaines intuitions sur les comparaisons de cardinaux d'ensembles finis peuvent être fausses dans ce contexte.

Avertissement 2.2.3

Deux ensembles vérifiant une inclusion stricte $E \subsetneq F$ peuvent être de même cardinal. Pourtant F contient strictement plus d'éléments que E.

Exemple 2.2.4 – Hôtel de Hilbert

$f : \mathbb{N} \to \mathbb{N}^*$ tel que $f(n) = n + 1$ est une bijection. Donc $\mathbb{N}$ et $\mathbb{N}^*$ ont même cardinal.

II.2 Dénombrabilité

Définition 2.2.5 – Ensemble dénombrable

Soit E un ensemble. On dit que :

- E est dénombrable s'il est de même cardinal que $\mathbb{N}$;
- au plus dénombrable s'il est fini ou dénombrable.

Ainsi, E est dénombrable si et seulement s'il existe une bijection $f : \mathbb{N} \to E$.

Proposition 2.2.6 – Dénombrabilité de $\mathbb{N}^2$

Le produit cartésien $\mathbb{N}^2$ est dénombrable.

◁ **Éléments de preuve.**

Construire une bijection en diagonales. Ou bien utiliser une bijection classique entre $\mathbb{N}^2$ et $\mathbb{N}^*$, basée sur l'étude des facteurs pairs de la décomposition de n en produit de facteurs premiers. ▷

Nous souhaitons caractériser les ensembles au plus dénombrables à l'aide des sous-ensembles de $\mathbb{N}$. Pour cela, nous commençons pas établir un lemme sur le cardinal des sous-ensembles de $\mathbb{N}$.

Lemme 2.2.7

Tout sous-ensemble de $\mathbb{N}$ est au plus dénombrable.

◁ **Éléments de preuve.**

Considérer un sous-ensemble $E \in \mathcal{P}(\mathbb{N})$, non fini. Construire par récurrence une application injective de $[\![0, n]\!]$ dans E, en rajoutant à chaque étape le minimum des éléments non encore utilisés. Cela définit une bijection $\mathbb{N} \to E$. ▷

Proposition 2.2.8 – Première caractérisation des ensembles au plus dénombrables

Un ensemble E est au plus dénombrable si et seulement s'il peut être mis en bijection avec une partie de $\mathbb{N}$.

◁ **Éléments de preuve.**

Un sens est évident par les définitions de la finitude ou de la dénombrabilité. L'autre provient du lemme. ▷

Nous donnons une deuxième caractérisation des ensembles au plus dénombrables en terme de surjection ou d'injection. Cette caractérisation est plus souple pour prouver les différentes propriétés de dénombrabilité liées aux constrructions ensemblistes. Cette caractérisation nécessite un petit lemme.

Lemme 2.2.9 – Inversibilité à gauche ou droite des injections et surjections

Soit $f : E \to F$.

1. Si f est injective, il existe $g : F \to E$ telle que $g \circ f = \mathrm{id}_E$, et g est alors surjective.

2. Si f est surjective, il existe $g : E \to F$ telle que $f \circ g = \mathrm{id}_F$, et g est alors injective.

◁ **Éléments de preuve.**
On construit g associant à y un antécédent de ses antécédents. Dans le premier cas, on n'a pas le choix quand on est dans l'image de f, et sinon, on définit g comme on veut. Dans le deuxième cas, on doit choisir parmi les antécédents de y. ▷

Remarque 2.2.10

- Ce sont en fait des équivalences.
- On utilise l'axiome du choix (qui affirme la possibilité de pouvoir faire simultanément une infinité de choix).

Proposition 2.2.11 – Deuxième caractérisation des ensembles au plus dénombrables

Soit E un ensemble non vide. Les trois propriétés suivantes sont équivalentes :

(i) E est au plus dénombrable

(ii) il existe une fonction injective $f : E \longrightarrow \mathbb{N}$.

(iii) il existe une fonction surjective $f : \mathbb{N} \longrightarrow E$.

◁ **Éléments de preuve.**

- $(i) \Longrightarrow (ii)$: considérer une bijection $E \to F \subset \mathbb{N}$, et composer par l'injection canonique.
- $(ii) \Longrightarrow (iii)$: c'est l'inversibilité à gauche d'une injection
- $(iii) \Longrightarrow (ii)$: c'est l'inversibilité à droite d'une surjection. Pourquoi l'axiome du choix n'est-il pas nécessaire ici ?
- $(ii) \Longrightarrow (i)$: corestriction à l'image.

▷

II.3 Règles opératoires sur les ensembles dénombrables

Proposition 2.2.12 – Construction d'ensembles dénombrables

1. Si E et F sont au plus dénombrables, alors $E \times F$ est au plus dénombrable.
2. Plus généralement, un produit $E_1 \times \cdots \times E_n$ d'un nombre fini d'ensembles au plus dénombrables est au plus dénombrable. Si l'un d'eux est dénombrable et les autres non vides, alors le produit est dénombrable.
3. Une union d'un nombre au plus dénombrable d'ensembles au plus dénombrables est au plus dénombrable. Si l'un des ensembles est dénombrable, l'union est dénombrable.

◁ **Éléments de preuve.**

1. Construire une surjection par composition : $\mathbb{N} \to \mathbb{N} \times \mathbb{N} \to E \times F$
2. Récurrence
3. Supposons les ensembles indexés sur I dénombrable. Construire une surjection

$$\mathbb{N} \to \mathbb{N} \times \mathbb{N} \to I \times \mathbb{N} \to \bigcup_{i \in I} A_i.$$

▷

Corollaire 2.2.13

1. L'ensemble $\mathbb{Z}$ est dénombrable.
2. L'ensemble $\mathbb{N}^p$ est dénombrable.
3. L'ensemble $\mathbb{Q}$ des rationnels est dénombrable.

Un résultat important sur les familles sommables découle de ces propriétés. On rappelle la définition suivante.

Définition 2.2.14 – Support d'une famille de scalaires

Soit $a = (a_i)_{i \in I}$ une famille de réels ou complexes. Le support de a est l'ensemble

$$J = \{i \in I \mid a_i \neq 0\}.$$

Proposition 2.2.15 – Support d'une famille sommable

Soit $a = (a_i)_{i \in I}$ une famille de réels ou complexes. Si a est sommable, alors le support de a est au plus dénombrable.

◁ **Éléments de preuve.**

Se ramener au cas d'une famille réelle positive, en contrôlant le support de a en fonction de celui des parties positives et négatives de ses parties réelles et imaginaires.

Dans le cas positif, exprimer le support en fonction des

$$I_\varepsilon = \{i \in I \mid a_i > \varepsilon\}.$$

Que peut-on dire du cardinal de I_ε ? ▷

II.4 Des ensembles infinis non dénombrables

Comme on l'a vu, des ensembles qui semblent bien plus gros que $\mathbb{N}$ sont dénombrables, donc ont dans les faits « autant » d'éléments que $\mathbb{N}$. Une question qui peut se poser alors est de savoir si tous les ensembles infinis ont même cardinal, ou s'il existe d'autres cardinaux (plus gros). Un premier élément de réponse provient de l'étude du cardinal de $\mathbb{R}$, à partir de l'écriture décimale d'un nombre réel. On rappelle qu'un nombre réel admet un et un seul développement décimal, à condition d'exclure les développements terminant uniquement par une infinité de 9.

Théorème 2.2.16 – Cardinal de $\mathbb{R}$, Cantor

L'ensemble des réels $\mathbb{R}$ est non dénombrable.

◁ **Éléments de preuve.**

(démonstration non exigible) Sinon, considérer une énumération $(x_n)_{n \in \mathbb{N}^*}$ de tous les réels, et considérer un y tel que le n-ième chiffre après la virgule de y soit distinct de 0, 9 et du n-ième chiffre de x_n. Le réel y peut-il être dans l'énumération ? Pourquoi prendre la précaution de prendre les chiffres distincts de 0 et 9 ? ▷

Cette preuve est connue sous le nom d'« argument diagonal de Cantor », dont une généralisation donne un deuxième élément de réponse à la question initiale :

Théorème 2.2.17 – Cantor, HP

Soit X un ensemble infini. Alors $\mathrm{Card}(X) < \mathrm{Card}(\mathcal{P}(X))$ (*i.e.* X est subpotent à $\mathcal{P}(X)$, mais pas équipotent)

◁ **Éléments de preuve.**
S'il existe une surjection $f : X \to \mathcal{P}(X)$, considérer $Y \subset X$ constitué des éléments x tels que $x \notin f(x)$, et obtenir une contradiction en considérant l'appartenance de c à Y, c étant un élément de X tel que $f(c) = Y$. ▷

On peut donc construire une chaîne infini d'ensembles ayant des cardinaux strictement croissants. Il y a donc une infinité de cardinaux infinis distincts.

Remarque 2.2.18

En quoi le résultat précédent empêche-t-il l'existence d'un « ensemble des ensembles » ?

3

Espaces vectoriels normés

Le but de ce chapitre est d'introduire un cadre qui permettra de généraliser l'analyse réelle de première année au cas de fonctions définies d'un espace vectoriel vers un autre. En particulier, la notion importante à la base de la notion de limite (et à partir de là, de continuité et de dérivabilité) est la notion de convergence d'une suite.
On aimerait donc donner un cadre rigoureux à l'étude de la convergence de suites dans un espace vectoriel E sur $\mathbb{R}$ ou $\mathbb{C}$. Si cet espace est de dimension finie, il suffit d'étudier la convergence coordonnée par coordonnée relativement à une base fixée. On verra dans un chapitre ultérieur que la propriété de convergence ne dépend alors pas du choix de la base. Ainsi, dans ce contexte, toute notion « raisonnable » de convergence équivaut à la convergence coordonnée par coordonnée.
Lorsque E n'est pas de dimension finie, les choses sont beaucoup moins claires, et nécessitent de revenir à l'intuition et la définition de la convergence, en terme d'approximation : une suite réelle ou complexe $(u_n)_{n\in\mathbb{N}}$ converge vers ℓ, si et seulement si, pour toute qualité d'approximation $\varepsilon > 0$, à partir d'un certain rang, u_n est presque égal à ℓ, à ε près. Cette définition nécessite de mesurer une distance de u_n à ℓ.
Pour généraliser cette définition, nous introduisons la notion de norme sur un espace vectoriel, permettant de définir ensuite une distance. Beaucoup de notions d'analyse réelle (notamment dans un premier temps d'analyse séquentielle) se généralisent alors dans le cadre des espaces vectoriels normés (c'est-à-dire munis d'une norme). On montrera par exemple dans le chapitre suivant des généralisations de certaines propriétés des fonctions continues, rencontrées dans le cadre de fonctions réelles : le théorème des bornes atteintes, le théorème des valeurs intermédiaires et le théorème de Heine par exemple.
Nous ferons remarquer dans la dernière partie (HP) qu'il est possible de définir directement une notion plus générale de distance, sans passer par une norme, ce qui est suffisant pour pouvoir définir de façon satisfaisante la notion de convergence.

I Normes sur un espace vectoriel

Dans toute cette partie $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$.

I.1 Définitions et propriétés élémentaires

Définition 3.1.1 – Norme sur un $\mathbb{K}$-espace vectoriel

Soit E un espace vectoriel sur $\mathbb{K}$, et $N : E \to \mathbb{R}_+$ une application à valeurs réelles positives. On dit que N est une norme si et seulement si N vérifie les propriétés suivantes :

(i) (séparation) pour tout $x \in E$, $N(x) = 0 \iff x = 0$;

(ii) (homogénéité) pour tout $x \in E$ et $\lambda \in \mathbb{K}$, $N(\lambda x) = |\lambda| N(x)$;

(iii) (inégalité triangulaire) pour tout $(x, y) \in E^2$, $N(x + y) \leqslant N(x) + N(y)$.

La propriété de positivité est en fait redondante, mais ce n'est généralement pas la plus dure à montrer dans les situations concrètes.

Proposition 3.1.2

Soit $N : E \to \mathbb{R}$ vérifiant les points *(i)*, *(ii)* et *(iii)* de la propriété précédente. Alors pour tout $x \in \mathbb{R}$, $N(x) \geqslant 0$ (et par conséquent, N est une norme).

◁ **Éléments de preuve.**

Considérer $N(x + (-x))$ et appliquer les trois points. ▻

Définition 3.1.3 – Espace vectoriel normé

- Un espace vectoriel normé E est un $\mathbb{K}$ espace vectoriel E muni d'une norme $N : E \to \mathbb{R}_+$.
- Cette norme sera souvent notée $\|\cdot\|$. Ainsi, pour $x \in E$, $\|x\|$ désignera la norme de x.

Exemples 3.1.4

- Si $E = \mathbb{K}$, $|\cdot|$ est une norme.
- Si $E = \mathbb{R}^2$, $(x, y) \mapsto \sqrt{x^2 + y^2}$ est une norme.

Le deuxième exemple est un cas particulier d'une situation plus générale déjà croisée en première année.

Proposition 3.1.5 – Norme euclidienne

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. Pour tout x, on note

$$\|x\| = \sqrt{\langle x, x\rangle}.$$

Alors $\|\cdot\|$ est une norme sur E, appelée norme euclidienne associée au produit scalaire $\langle\cdot,\cdot\rangle$.

◁ **Éléments de preuve.**

Pour l'inégalité triangulaire, développer d'un côté $\langle x + y, x + y\rangle$, d'un autre $(\|x\| + \|y\|)^2$ et comparer les deux. ▻

Proposition 3.1.6 – Inégalités triangulaires

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Alors :

1. Pour tout $(x_1, \dots, x_n) \in E^n$, $\left\| \sum_{k=1}^{n} x_k \right\| \leqslant \sum_{k=1}^{n} \|x_k\|$.
2. Pour tout $(x, y) \in E^2$, $|\|x\| - \|y\|| \leqslant \|x - y\|$

Lemme 3.1.7 – Transfert d'une norme

Soit $\varphi : E \to F$ une application linéaire injective entre deux espaces vectoriels, et N une norme sur F. Alors l'application N' définie sur E par

$$N'(x) = N(\varphi(x))$$

est une norme sur E.

◁ **Éléments de preuve.**

Vérifications faciles par linéarité, et caractérisation de l'injectivité par le noyau. ▻

Corollaire 3.1.8 – Restriction d'une norme

Soit N une norme sur E, et soit F un sous-espace vectoriel de E. Alors la restriction N_F de N à F est une norme sur F.
Ainsi, tout sous-espace vectoriel d'un e.v.n. est muni d'une structure d'e.v.n. par restriction de la norme de E.

I.2 Normes sur des espaces de dimension finie

Proposition/Définition 3.1.9 – Normes usuelles sur $\mathbb{K}^n$

Soit, pour tout $X = (x_1, \dots, x_n) \in \mathbb{K}^n$:

$$\|X\|_1 = \sum_{k=1}^{n} |x_k|, \qquad \|X\|_2 = \sqrt{\sum_{k=1}^{n} |x_k|^2}, \qquad \|X\|_\infty = \max_{i \in [\![1,n]\!]} |x_i|.$$

Alors $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sont des normes.

◁ **Éléments de preuve.**
Le seul point un peu délicat est l'inégalité triangulaire pour $\|\cdot\|_2$ dans le cas où $\mathbb{K} = \mathbb{C}$. On peut se ramener au cas réel en considérant le vecteur $\tilde{X} = (|x_1|, \dots, |x_n|)$. On peut aussi adapter la démonstration du cas réel, en définissant un produit « hermitien » :

$$\langle X, Y \rangle = \sum_{i=1}^{n} x_i \overline{y_i},$$

où $X = (x_1, \dots, x_n)$ et $Y = (y_1, \dots, y_n)$. L'inégalité de Cauchy-Schwarz se généralise à cette situation, et l'inégalité triangulaire en découle, comme dans le cas réel. ▷

Remarque 3.1.10

1. Si $\mathbb{K} = \mathbb{R}$, $\|\cdot\|_2$ est la norme euclidienne associée au produit scalaire canonique.
2. Un même espace vectoriel peut être muni de plusieurs normes ! Comme la notion de convergence est définie à l'aide des normes, elle peut être fortement dépendant du choix de la norme. Ainsi, une suite peut très bien converger pour une norme et pas pour une autre. On verra des exemples en exercice.

Par transfert, on en déduit :

Proposition/Définition 3.1.11 – Normes usuelles sur un espace de dimension finie

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, muni d'une base $\mathcal{B} = (b_1, \dots, b_n)$. Pour $X \in E$ tel que $[X]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, on définit :

$$\|X\|_{\mathcal{B},1} = \sum_{k=1}^{n} |x_k|, \qquad \|X\|_{\mathcal{B},2} = \sqrt{\sum_{k=1}^{n} |x_k|^2}, \qquad \|X\|_{\mathcal{B},\infty} = \max_{i \in [\![1,n]\!]} |x_i|.$$

S'il n'y a pas d'ambiguïté sur la base, on notera simplement $\|\cdot\|_1$, $\|\cdot\|_2$ ou $\|\cdot\|_\infty$.

Avertissement 3.1.12

Toute norme sur un $\mathbb{R}$-espace vectoriel n'est pas euclidienne. Par exemple, $\|\cdot\|_\infty$ n'est pas euclidienne.

◁ **Éléments de preuve.**
La formule de polarisation donne l'unique candidat à être le produit scalaire associé. Trouver un défaut de bilinéarité. ▷

I.3 Normes sur des espaces de fonctions

On rappelle que si E est un espace vectoriel sur $\mathbb{K}$, alors tout ensemble de fonctions E^A défini sur un ensemble quelconque A et à valeurs dans E est un espace vectoriel sur $\mathbb{K}$. On peut chercher alors à définir des normes sur cet espace ou sur des sous-espaces vectoriel de cet espace. Dans un premier temps, nous nous intéressons au sous-espace $\mathcal{C}^0([a,b],\mathbb{K})$ de $\mathbb{K}^{[a,b]}$, toujours avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

L'intégration correspondant à une sommation continue, les exemples précédents s'adaptent bien dans le cadre de fonctions continues.

Proposition/Définition 3.1.13 – Normes sur $\mathcal{C}^0([a,b],\mathbb{K})$

Pour tout $f \in \mathcal{C}^0([a,b],\mathbb{K})$, on définit

$$\|f\|_1 = \int_a^b |f(t)|\,\mathrm{d}t, \qquad \|f\|_2 = \sqrt{\int_a^b |f(t)|^2\,\mathrm{d}t}, \qquad \|f\|_\infty = \sup_{x\in[a,b]} f(t).$$

Alors $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sont des normes, appelées respectivement :

- norme de la convergence en moyenne pour $\|\cdot\|_1$;
- norme de la convergence en moyenne quadratique pour $\|\cdot\|_2$;
- Norme de la conevrgence uniforme, ou norme norme infinie pour $\|\cdot\|_\infty$.

◁ **Éléments de preuve.**
Même type de démonstration que dans le cas de la dimension finie ▷

Remarque 3.1.14

1. Ces normes s'étendent-elle au cas des fonctions continues par morceaux ? (la réponse n'est pas forcément la même suivant les normes).
2. Dans la définition de $\|\cdot\|_\infty$, le sup est-il un max ?

La norme $\|.\|_\infty$ peut être définie dans un contexte beaucoup plus général.

Définition 3.1.15 – Parties bornées, applications bornées

Soit A un ensemble quelconque, $(E, \|.\|)$ un e.v.n., et B une partie quelconque de E.

(i) On dit que B est bornée s'il existe un réel $M \in \mathbb{R}_+$ tel que pour tout $x \in B$, $\|x\| \leqslant M$.

(ii) On dit qu'une application $f \in E^A$ est bornée si son image $\mathrm{Im}(f)$ est un sous-ensemble borné de E, c'est-à-dire s'il existe $M \in \mathbb{R}_+$ tel que

$$\forall x \in A, \quad \|f(x\| \leqslant M.$$

On note $\mathcal{B}(A,E)$ l'ensemble des applications bornées de A dans E.

Proposition 3.1.16 – Structure de $\mathcal{B}(A,E)$

L'ensemble $\mathcal{B}(A,E)$ est un espace vectoriel sur $\mathbb{K}$.

Proposition/Définition 3.1.17 – Norme de la convergence uniforme sur $\mathcal{B}(A,E)$

Soit A un ensemble quelconque et $(E, \|\cdot\|)$ un e.v.n.. Pour tout $f \in \mathcal{B}(A,E)$, on définit

$$\|f\|_\infty = \sup_{a \in A} \|f(a)\|.$$

L'application $\|\cdot\|_\infty$ est une norme sur $\mathcal{B}(A,E)$, appelée norme de la convergence uniforme sur $\mathcal{B}(A,E)$.

$\lhd$ **Éléments de preuve.**
Même preuve que dans le cas $\mathcal{C}^0([a,b],\mathbb{K})$. $\rhd$

I.4 Normes sur $\mathbb{K}[X]$

Suivant qu'on s'intéresse à des propriétés algébriques liées aux coefficients, ou analytiques liées aux valeurs de la fonction polynomiale associée, on peut adopter différentes normes sur l'espace des polynômes. En particulier, les normes définies sur $\mathbb{K}^n$ s'adaptent bien, du fait que, même si $\mathbb{K}[X]$ n'est pas de dimension finie, les polynômes ont un nombre fini de coefficients non nuls.

Ainsi, pour $P = \sum_{k=0}^{+\infty} a_k X^k$ (somme en réalité finie), on peut définir

$$\|P\|_1 = \sum_{n=0}^{+\infty} |a_n|, \qquad \|P\|_2 = \sqrt{\sum_{n=0}^{+\infty} |a_n|^2}, \qquad \|P\|_\infty = \sup_{n \in \mathbb{N}} |a_n|.$$

Et dans un contexte plus analytique, on peut définit les normes $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$ associées aux fonctions polynomiales définies par les polynômes formels.

Remarque 3.1.18

La définition des normes séquentielles ci-dessus s'étend à certains sous-espaces de l'espace des suites $\mathbb{K}^{\mathbb{N}}$ (les sous-espaces permettant d'assurer la bonne définition de ces normes).

I.5 Normes sur $\mathcal{M}_n(\mathbb{K})$

L'espace $\mathcal{M}_n(\mathbb{K})$ est isomorphe à $\mathbb{K}^{n^2}$ (par exemple en mettant les unes au dessus des autres les colonnes de $\mathcal{M}_n(\mathbb{K})$). Via cet isomophisme, les normes usuelles de $\mathbb{K}^{n^2}$ se transfèrent à $\mathcal{M}_n(\mathbb{K})$: pour une matrice $A = (a_{i,j})$, on obtient les normes

$$\|A\|_1 = \sum_{1 \leqslant i,j \leqslant n} |a_{i,j}|, \qquad \|A\|_2 = \sqrt{\sum_{1 \leqslant i,j \leqslant n} |a_{i,j}|^2}, \qquad \|A\|_\infty = \max_{1 \leqslant i,j \leqslant n} |a_{i,j}|.$$

Remarque 3.1.19

1. Si $\mathbb{K} = \mathbb{R}$, la norme $\|\cdot\|_2$ est associée au produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$, défini par

$$\langle A, B \rangle = \operatorname{tr}(A^\top B) = \sum_{1 \leqslant i,j \leqslant n} a_{i,j} b_{i,j}.$$

2. Ces 3 définitions s'étendent aux matrices rectangulaires $\mathcal{M}_{n,p}(\mathbb{K})$.

3. Dans le cas de matrices carrées, comparer $\|AB\|$ et $\|A\| \cdot \|B\|$ pour ces différentes normes

Définition 3.1.20 – Norme matricielle sur $\mathcal{M}_n(\mathbb{K})$

Une norme $\|\cdot\|$ sur $\mathcal{M}_n(\mathbb{K})$ est appelée *norme matricielle* si elle est sous-multiplicative, *i.e.* :

$$\forall (A,B) \in \mathcal{M}_n(\mathbb{K})^2, \quad \|AB\| \leqslant \|A\| \cdot \|B\|.$$

Proposition 3.1.21 – Existence d'une norme matricielle, HP

Les normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sont matricielles, la normes $\|\cdot\|_\infty$ ne l'est pas.

◁ **Éléments de preuve.**

- La matrice constituée uniquement de 1 est un contre-exemple pour $\|\cdot\|_1$ et $\|\cdot\|_\infty$
- Le cas de $\|\cdot\|_2$ résulte de l'inégalité de CS. Ce résultat est classique (nombreux exercices) mais pas au programme. La démonstration doit donc être connue.

▷

L'existence d'une norme matricielle est d'une grande utilité dans l'étude de convergences de séries matricielles par exemple, notamment des séries faisant intervenir des puissances d'une matrice A. C'est le cas par exemple de la série exponentielle, permettant de définir l'exponentielle d'une matrice carrée :

$$\exp(A) = \sum_{n=0}^{+\infty} \frac{A^n}{n!}.$$

Nous justifierons ultérieurement que cette série est convergente (indépendamment de la norme choisie). L'un des points clés (outre l'équivalence des normes en dimension finie) sera de pouvoir majorer $\|A^n\|$ par $\|A\|^n$, pour pouvoir se ramener à la convergence d'une série numérique.
Nous définirons dans le chapitre suivant une autre norme matricielle sur $\mathcal{M}_n(\mathbb{K})$, notée $|||A|||$, et appelée *norme triple*, ou *norme d'opérateur*.

I.6 Produits d'e.v.n.

Les normes $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sur des espaces de dimension finie ont été définies à partir des modules des coordonnées, c'est à dire d'une norme particulière sur $\mathbb{K}$. On peut alors généraliser ces constructions en remplaçant des coordonnées sur $\mathbb{K}$ par des composantes vectorielles, à condition de disposer de normes sur chacune des composantes. On obtient de la sorte des normes sur des produits cartésiens d'e.v.n..

Proposition/Définition 3.1.22 – Produit cartésien d'e.v.n.

Soit $((E_i, N_i))_{i \in [\![1,n]\!]}$ une famille finie d'e.v.n.. Pour $X = (x_1, \dots, x_n) \in E_1 \times \cdots \times E_n$, on définit

$$\|X\|_1 = \sum_{i=1}^{n} N_i(x_i), \qquad \|X\|_2 = \sqrt{\sum_{i=1}^{n} N_i(x_i)^2} \qquad \text{et} \qquad \|X\|_\infty = \max_{i \in [\![1,n]\!]} N_i(x_i).$$

Ce sont des normes. La norme $\|\cdot\|_\infty$ ainsi définie est appelée norme produit des normes N_i.

◁ **Éléments de preuve.**

Adaptation facile du cas de $\mathbb{K}^n$. ▷

Remarque 3.1.23

La notation $\|\cdot\|_\infty$ n'indique que la dernière couche de la définition de la norme. Les normes internes utilisées sur chaque composante peuvent ne pas être des normes infinies.

II Topologie d'un espace vectoriel normé

II.1 Distance et boules

Définition 3.2.1 – Distance associée à une norme

Soit $(E, \|\cdot\|)$ un e.v.n., et $(x, y) \in E$. On définit la distance de x à y par :

$$d(x, y) = \|y - x\|.$$

Proposition 3.2.2 – Propriétés de la distance

La distance $d : E^2 \to \mathbb{R}$ vérifie :

(i) (séparation) pour tout $(x, y) \in E^2$, $d(x, y) = 0 \iff x = y$;

(ii) (symétrie) pour tout $(x, y) \in E^2$, $d(x, y) = d(y, x)$;

(iii) (inégalité triangulaire) pour tout $(x, y, z) \in E^3$, $d(x, z) \leqslant d(x, y) + d(y, z)$;

(iv) (positivité) pour tout $(x, y) \in E^2$, $d(x, y) \geqslant 0$.

Remarque 3.2.3

- La propriété (iv) découle des 3 autres.
- Les propriétés (i) à (iii) ci-dessus (et donc aussi (iv)) permettent de définir des distances sur des ensembles beaucoup plus généraux (qui n'ont pas besoin d'être des espaces vectoriels).

Dans ce qui suit, on se place dans un e.v.n. E, dont la norme sera notée $\|\cdot\|$, et la distance associée sera notée d.

Définition 3.2.4 – Boules ouvertes, boules fermées, sphère

Soit $a \in E$, et $r \in \mathbb{R}_+$

(i) La boule ouverte de centre a et de rayon r est définie par $B(a, r) = \{x \in E \mid d(x, a) < r\}$;

(ii) La boule fermée de centre a et de rayon r est définie par $\overline{B}(a, r) = \{x \in E \mid d(x, a) \leqslant r\}$;

(iii) La sphère de centre a et de rayon r est définie par $S(a, r) = \{y \in E \quad d(x, a) = r\}$.

Pour lever une éventuelle ambiguïté, cela arrive d'utiliser la notation $\mathring{B}(a, r)$ pour désigner la boule ouverte $B(a, r)$.

Exemples 3.2.5

1. Les boules dans $\mathbb{R}^2$ au sens de la norme $\|\cdot\|_2$ sont représentées dans la figure 3.1. Essayez de représenter de la même façon les boules pour la norme $\|\cdot\|_1$ et la norme $\|\cdot\|_\infty$, et observez qu'un boule n'est pas toujours ronde !
2. À quoi ressemble la boule $B(f, \varepsilon)$, pour $f \in \mathcal{C}([a, b], \mathbb{R})$, muni de la norme $\|\cdot\|_\infty$?

Exemple 3.2.6

Dans $\mathbb{R}$, les boules sont des intervalles (voir figure 3.2) :

- $B(x, r) =]x - r, x + r[$
- $\overline{B}(x, r) = [x - r, x + r]$

En fait, tout intervalle borné ouvert est une boule ouverte, tout intervalle borné fermé est une boule fermée :

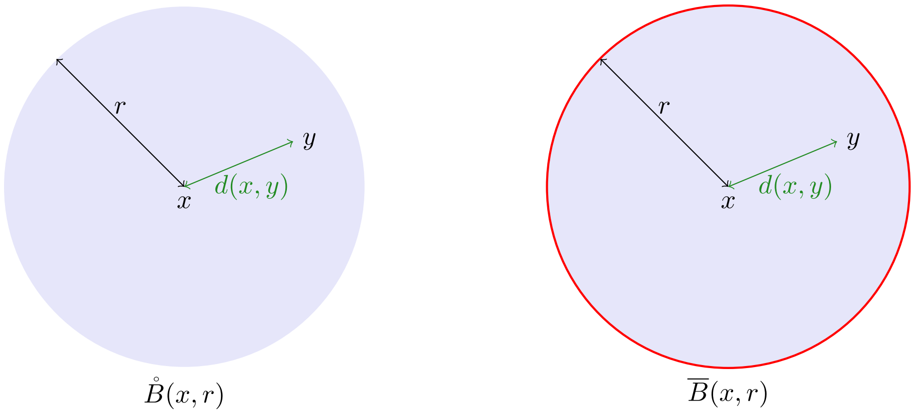

FIGURE 3.1 – Boules dans $(\mathbb{R}^2, \|\cdot\|_2)$

- $]a, b[= B\left(\frac{a+b}{2}, \frac{b-a}{2}\right)$,
- $[a, b] = \overline{B}\left(\frac{a+b}{2}, \frac{b-a}{2}\right)$

FIGURE 3.2 – Boule ouverte, boule fermée dans $\mathbb{R}$

Proposition 3.2.7 – Caractérisation des parties bornées

(i) Une boule est bornée.

(ii) Une partie A de E est bornée si et seulement s'il existe $(a, r) \in E \times \mathbb{R}_+$ tel que $A \subset B(a, r)$.

◁ **Éléments de preuve.**
Utiliser l'inégalité triangulaire. Faire un dessin ! ▷

L'inégalité triangulaire implique également une autre propriété des boules, bien visible sur tous les exemples étudiés : la convexité (au sens géoémtrique).

Définition 3.2.8 – Ensemble convexe, figure 3.3

Soit F une partie E. On dit que F est convexe si et seulement si pour tout couple de points A et B de E, le segment $[AB] = \{(1-t)A + tB,\ t \in [0, 1]\}$ est entièrement inclus dans E.

Exemple 3.2.9

Par définition, les intervalles sont les sous-ensembles convexes de $\mathbb{R}$.

Théorème 3.2.10 – Convexité des boules

Pour tout $a \in E$ et $r > 0$, $\mathring{B}(a, r)$ et $\overline{B}(a, r)$ sont des parties convexes de E.

◁ **Éléments de preuve.**
Encore un petit coup d'inégalité triangulaire. ▷

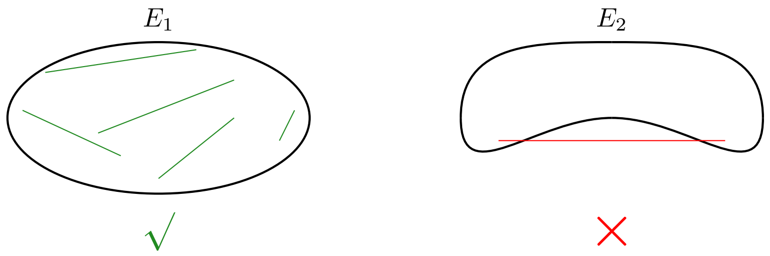

Figure 3.3 – Un sous-ensemble convexe E_1 et un sous-ensemble non convexe E_2 de $\mathbb{R}^2$

II.2 Ouverts, fermés, voisinages

La notion de boule est très mesurée. Nous définissons maintenant des objets un peu similaires, permettant d'assurer la présence de « matière » de toute part d'un point a, mais sans être aussi mesuré. Ainsi, il s'agit d'une notion plus qualitative que la notion très quantitative de boule. L'ensemble E est toujours un e.v.n., muni de sa norme $\|\cdot\|$ et de sa distance d.

Définition 3.2.11 – Voisinage, figure 3.4

Soit $a \in E$. Un *voisinage* V de a est un sous-ensemble V de E tel qu'il existe une boule ouverte centrée en a entièrement contenue dans V :

$$\exists \varepsilon > 0,\ B(a, \varepsilon) \subset V, \qquad \text{i.e.} \qquad \exists \varepsilon > 0,\ \forall x \in E,\ d(x, a) < \varepsilon \Longrightarrow x \in V.$$

On note $\mathcal{V}(a) \subset \mathcal{P}(E)$ l'ensemble des voisinages de a.

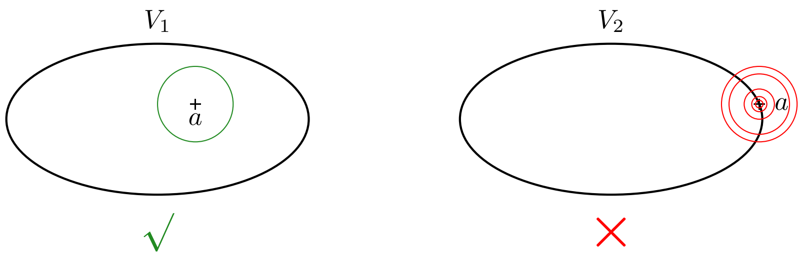

Figure 3.4 – Voisinage de a

En gros, V est un voisinage de x si x est « à l'intérieur de V », et non sur un bord. En s'éloignant un peu de x, on ne sort pas de V.

Exemple 3.2.12 – Voisinages dans $\mathbb{R}$

- Dans $\mathbb{R}$, un voisinage de x est un ensemble contenant un intervalle $]a, b[$ tel que $x \in]a, b[$.
- Par extension et commodité, on dit parfois qu'un ensemble contenant un intervalle $]a, +\infty[$ est un voisinage de $+\infty$. Version symétrique pour $-\infty$. On notera $\mathcal{V}(+\infty)$ l'ensemble des voisinages de $+\infty$ dans $\mathbb{R}$.

Proposition 3.2.13 – Intersection et union de voisinages

Soit $(V_i)_{i\in I} \in \mathcal{V}(a)^I$ une famille de voisinages de a.

(i) Soit $V \in \mathcal{V}(a)$ et $W \in \mathcal{P}(E)$ tel que $V \subset W$. Alors $W \in \mathcal{V}(a)$.

(ii) En particulier, si $I \neq \varnothing$, $\displaystyle\bigcup_{i\in I} V_i \in \mathcal{V}(a)$.

(iii) Si I est fini, $\bigcap_{i \in I} V_i \in \mathcal{V}(a)$.

Exemple 3.2.14

Trouver dans $\mathbb{R}$ un exemple de famille infinie de voisinages de $a \in \mathbb{R}$ dont l'intersection n'est pas un voisinage de a.

Définition 3.2.15 – Partie ouverte

- Un *ouvert* U de E est une partie U de E qui est voisinage de tous ses points
- De manière équivalente, $U \subset E$ est un ouvert ssi :

$$\forall x \in U,\ \exists \varepsilon > 0,\ B(x, \varepsilon) \subset U.$$

Intuitivement, un ouvert est un ensemble dont le « bord » est flou : on peut s'en approcher, mais jamais l'atteindre en restant dans U. Ainsi, l'image qu'il faut en garder est qu'un ouvert est un ensemble ne contenant pas son bord. Évidemment, n'ayant pas défini la notion de bord, ceci reste une image.

Définition 3.2.16 – Partie fermée

Un sous-ensemble F de E est *fermé* si son complémentaire $\complement_E F$ est ouvert.

Cette fois, intuitivement, c'est le complémentaire qui ne contient pas son bord, donc F, lui contient tout son bord.

Exemples 3.2.17

1. Les intervalles ouverts sont des sous-ensembles ouverts de $\mathbb{R}$.
2. Les intervalles fermés sont des sous-ensembles fermés de $\mathbb{R}$.
3. Les intervalles semi-ouverts ne sont ni ouverts ni fermés.
4. $\mathbb{R}$ et $\varnothing$ sont des sous-ensembles à la fois fermés et ouverts de $\mathbb{R}$.
5. On peut montrer que les sous-ensembles ouverts de $\mathbb{R}$ sont les unions disjointes d'intervalles ouverts (voir exercices)

Définition 3.2.18 – Topologie

La *topologie* d'un e.v.n. E est la donnée de l'ensemble $\mathcal{O}$ de tous ses ouverts.

La propriété suivante généralise l'observation faite sur les intervalles.

Proposition 3.2.19 – Propriétés topologiques des boules

1. Les boules ouvertes $\mathring{B}(a, r)$ sont des parties ouvertes de E.
2. Les boules fermées $\overline{B}(a, r)$ sont des parties fermées de E.
3. Les sphères $S(a, r)$ sont des parties fermées de E.

Proposition 3.2.20 – union, intersection d'ouverts et de fermés

1. Toute union quelconque d'ouverts est un ouvert ;
2. Toute intersection d'un nombre *fini* d'ouverts est un ouvert ;
3. Toute intersection quelconque de fermés est un fermé ;
4. Toute union d'un nombre *fini* de fermés est un fermé.

◁ **Éléments de preuve.**

1. Si x est dans l'union, il existe une boule centrée en x restant dans l'un des ouverts, donc dans leur union.
2. Prendre le minimum des rayons des boules centrées en x restant dans chaque ouvert. Pourquoi doit-on se limiter au cas fini ?
3. Par complémentation
4. De même.

▷

Exemples 3.2.21

Voici deux contre-exemples à bien garder en tête :

1. Contre-exemple pour une intersection infinie d'ouverts : $\displaystyle\bigcap_{n=1}^{+\infty} \left]-\frac{1}{n}, 1\right[= [0, 1[$.
2. Contre-exemple pour une union infinie de fermés : $\displaystyle\bigcup_{n=1}^{+\infty} \left[\frac{1}{n}, 1\right] =]0, 1]$.

Proposition 3.2.22 – Propriétés topologiques d'un produit cartésien

Soit $((E_i, N_i))_{i\in[\![1,n]\!]}$ une famille finie d'e.v.n.. On munit le produit cartésien $E = E_1 \times \cdots \times E_n$ de la norme produit.

1. Soit, pour tout $i \in [\![1, n]\!]$, U_i un ouvert de E_i. Alors $U_1 \times \cdots \times U_n$ est un ouvert de E.
2. Soit, pour tout $i \in [\![1, n]\!]$, F_i un fermé de E_i. Alors $F_1 \times \cdots \times F_n$ est un fermé de E.

II.3 Intérieur, adhérence, frontière

Définition 3.2.23 – Intérieur

Soit A une partie de l'e.v.n. E.

(i) Un point $a \in E$ est *intérieur* à A s'il existe $\varepsilon > 0$ tel que $B(a, \varepsilon) \subset A$.

(ii) L'intérieur de A, noté $\mathring{A}$, est la partie de E constituée de tous les points intérieurs à A :

$$\mathring{A} = \{a \in E \mid \exists\varepsilon > 0, \quad B(a, \varepsilon) \subset A\}.$$

Proposition 3.2.24 – Caractérisation topologique de l'intérieur

Soit A une partie de E ;

(i) $\mathring{A} \subset A$;

(ii) $\mathring{A}$ est un ouvert de E ;

(iii) Si U est un ouvert de A tel que $U \subset A$, alors $U \subset \mathring{A}$.

Ces trois propriétés peuvent être résumées en disant que $\mathring{A}$ est le plus grand ouvert de E inclus dans A.

Corollaire 3.2.25 – Intérieur d'un ouvert

(i) Si A est ouvert, $\mathring{A} = A$. C'est même une équivalence.

(ii) $\mathring{\mathring{A}} = \mathring{A}$.

Exemples 3.2.26

1. Décrire l'intérieur d'un intervalle.
2. Décrire $\overset{\circ}{\widehat{\overline{B}(a,r)}}$?
3. Justifier que $\mathring{\mathbb{Q}} = \varnothing$. A-t-on $\overset{\circ}{\widehat{A \cup B}} = \mathring{A} \cup \mathring{B}$?

Définition 3.2.27 – Adhérence

Soit A une partie de E.

(i) Un point a est adhérent à A si pour tout $\varepsilon > 0$, $B(a,\varepsilon) \cap A \neq \varnothing$.

(ii) L'adhérence $\overline{A}$ est l'ensemble des points adhérents à A :

$$\overline{A} = \{a \in E \mid \forall \varepsilon > 0,\ B(a,\varepsilon) \cap A \neq \varnothing\}.$$

L'adhérence est donc l'ensemble des points qui « collent » à A, dans le sens où on peut en trouver des points de A arbitrairement proches.

Exemples 3.2.28

1. Décrire $\overline{\mathbb{Q}}$. A-t-on $\overline{A \cap B} = \overline{A} \cap \overline{B}$?
2. Que dire de l'adhérence d'un intervalle de $\mathbb{R}$?
3. Décrire $\overline{\mathring{B}(a,r)}$
4. Soit $F = \{x \mapsto \mathrm{e}^{-nx},\quad n \in \mathbb{N}^*\} \subset \mathcal{C}^0([0,1],\mathbb{R})$. Montrer que la fonction nulle est adhérente à F pour la norme $\|\cdot\|_1$ mais pas pour la norme $\|\cdot\|_\infty$.
5. Soit A une partie non vide et bornée de $\mathbb{R}$. Montrer que $\sup(A)$ et $\inf(A)$ sont adhérents à A.

Lemme 3.2.29 – Lien avec l'intérieur

Soit $A \subset E$. Alors

- $\overline{A} = (\mathring{\widehat{A^c}})^c$, où X^c désigne le complémentaire de X dans E ;
- $(\overline{A})^c = \overset{\circ}{\widehat{A^c}}$
- $(\overline{A})^c = \overset{\circ}{\widehat{A^c}}$

Proposition 3.2.30 – Caractérisation topologique de l'adhérence

Soit $A \subset E$. Alors $\overline{A}$ est le plus petit fermé (au sens de l'inclusion) tel que $A \subset \overline{A}$. Autrement dit :

(i) $\overline{A}$ est un fermé

(ii) $A \subset \overline{A}$

(iii) Si F est un fermé tel que $A \subset F$, alors $\overline{A} \subset F$.

On donne une autre caractérisation en terme de distance. Pour cela, on définit dans un premier temps la distance d'un point à une partie.

Définition 3.2.31 – Distance à une partie

Soit A une partie non vide de E. Soit $x \in E$. On définit la distance de x à A par :

$$d(x,A) = \inf_{y \in A} \|y - x\|_E$$

Proposition 3.2.32 – Caractérisation de l'adhérence par la distance

Soit $A \subset E$. Alors
$$\overline{A} = \{x \in E \mid d(x, A) = 0\}.$$

Propriétés 3.2.33 – Propriétés relatives aux intersections et unions

Ces propriétés ne sont pas explicitement marquées au programme, mais assez utiles. Soit $A \subset E$.

1. $\overset{\circ}{\widehat{A \cap B}} = \mathring{A} \cap \mathring{B}$.
2. $\overline{A \cup B} = \overline{A} \cup \overline{B}$.
3. $\mathring{A} \cup \mathring{B} \subset \overset{\circ}{\widehat{A \cup B}}$, mais l'inclusion réciproque peut être fausse.
4. $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$, mais l'inclusion réciproque peut être fausse.

Définition 3.2.34 – Frontière

Soit $A \subset E$. La frontière de A, notée $\mathrm{Fr}(A)$, est
$$\mathrm{Fr}(A) = \overline{A} \backslash \mathring{A} = \overline{A} \cap \overline{A^c}.$$

Il s'agit donc des points arbitrairement proche de points de A et de points qui ne sont pas dans A.

Exemples 3.2.35

1. Quelle est la frontière de $[a, b[$ dans $\mathbb{R}$?
2. Quelle est la frontière de $\mathbb{Q}$ dans $\mathbb{R}$?
3. Déterminer la frontière de $\mathring{B}(a, r)$. Quelle est la frontière de $\overline{B}(a, r)$?
4. De façon générale, étant donné $A \subset E$, a-t-on $\mathrm{Fr}(A) = \mathrm{Fr}(\overline{A})$? A-t-on $\mathrm{Fr}(A) = \mathrm{Fr}(\mathring{A})$?

Définition 3.2.36 – Parties denses

(i) Soit A et B deux parties de E telles que $A \subset B$. On dit que A est dense dans B si $B \subset \overline{A}$. Autrement dit, A est dense dans B ssi :
$$\forall b \in B,\ \forall \varepsilon > 0,\ \exists a \in A,\ d(a, b) < \varepsilon.$$

(ii) Soit A une partie de E. On dit que A est partout dense si $\overline{A} = E$ (ce qui équivaut donc à dire que A est dense dans E.

Proposition 3.2.37 – Lien avec la densité dans $(\mathbb{R}, \leqslant)$

Soit $A \subset \mathbb{R}$. Les propositions suivantes sont équivalentes :

(i) A est (partout) dense dans $\mathbb{R}$;

(ii) pour tout $(a, b) \in \mathbb{R}^2$ tel que $a < b$, il existe $x \in A$ tel que $a < x < b$.

Exemples 3.2.38

1. $\mathbb{Q} \subset \mathbb{R}$ est partout dense (dans $\mathbb{R}$). Il est dense dans $[0, 1]$.
2. $\mathrm{Esc}([a, b], \mathbb{R})$ est dense dans $\mathcal{C}^0([a, b], \mathbb{R})$, ainsi que dans $\mathcal{C}^0_m([a, b], \mathbb{R})$ (espace des fonctions continues par morceaux).
3. $\mathrm{GL}_n(\mathbb{C})$ est dense dans $\mathcal{M}_n(\mathbb{C})$.

II.4 Topologie relative

Définition 3.2.39 – Voisinages relatifs

Soit $A \subset E$, et $a \in A$. Un voisinage relatif de a dans A est une partie $V \subset A$ telle qu'il existe $V_0 \in \mathcal{V}(a)$ tel que $V = V_0 \cap A$.

Exemple 3.2.40

1. On se place dans l'e.v.n. $\mathbb{C}$. Soit $a \in \mathbb{R}$. Alors $]a-1, a+1[$ est un voisinage relatif de a dans $\mathbb{R}$, mais ce n'est pas un voisinage de a dans $\mathbb{C}$
2. On se place dans l'e.v.n. $\mathbb{R}$. Soit $a \in \mathbb{R}$. Alors $[a, a+1[$ est un voisinage relatif de a dans $[a, +\infty[$.

Terminologie 3.2.41 – Voisinage à droite, voisinage à gauche

Dans l'e.v.n. $(\mathbb{R}, |\cdot|)$, un voisinage relatif de a dans $[a, +\infty[$ est parfois appelé voisinage à droite de a.
De même, un voisinage à gauche de a est un voisinage relatif de a dans $]-\infty, a]$.

Proposition 3.2.42 – Caractérisation des voisinages relatifs

Soit $V \subset A \subset E$. Les propriétés suivantes sont équivalentes :
- V est un voisinage relatif de a dans A ;
- il existe $\varepsilon > 0$ tel que $A \cap B(a, \varepsilon) \subset V$;
- il existe $\varepsilon > 0$ tel que pour tout $x \in A$, $d(x, a) < \varepsilon \Longrightarrow a \in V$.

Proposition/Définition 3.2.43 – Ouverts relatifs

Soit $A \subset E$, et $U \subset A$. Les propriétés suivantes sont équivalentes :

(i) U est un voisinage relatif de tous ses points dans A ;

(ii) pour tout $a \in U$, il existe $\varepsilon > 0$ tel que $B(x, \varepsilon) \cap A \subset U$;

(iii) il existe un ouvert U_0 de E tel que $U = A \cap U_0$.

Si ces propriétés sont vérifiées, on dit que U est un ouvert relatif de A.

Exemples 3.2.44

1. Décrire les ouverts relatifs convexes de $[0, 1]$ dans $\mathbb{R}$.
2. Que dire des ouverts relatifs de A si A est ouvert ?
3. Montrer que A et $\varnothing$ sont toujours des ouverts relatifs de A.

Proposition/Définition 3.2.45 – Fermés relatifs

Soit $A \subset E$, et $F \subset A$. Les propriétés suivantes sont équivalentes :

(i) $A \backslash F$ est un ouvert relatif de A ;

(ii) il existe F_0 un fermé de E tel que $F = A \cap F_0$.

Lorsque ces propriétés sont vérifiées, on dit que F est un fermé relatif de A.

Exemples 3.2.46

1. Que dire des fermés relatifs de A lorsque A est fermé ?
2. Montrer que A et $\varnothing$ sont toujours des fermés relatifs de A.

Théorème 3.2.47 – Union, intersection d'ouverts, fermés relatifs

1. Une union quelconque d'ouverts relatifs de A est un ouvert relatif de A.
2. Une intersection finie d'ouverts relatifs de A est un ouvert relatif de A.
3. Une union finie de fermés relatifs de A est un fermé relatif de A.
4. Une intersection quelconque de fermés relatifs de A est un fermé relatif de A

III Convergences

III.1 Suites convergentes

Les notions de convergence vues pour les suites réelles et complexes se généralisent sans peine au cas des suites à valeurs dans un evn.

Proposition/Définition 3.3.1 – Convergence d'une suite d'éléments d'un e.v.n.

Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un e.v.n. E, et $\ell \in E$. Les propriétés suivantes sont équivalentes :

(i) (point de vue métrique) $\forall \varepsilon > 0,\quad \exists N \in \mathbb{N},\quad \forall n \geqslant N,\quad \|u_n - \ell\| \leqslant \varepsilon$.

(ii) (réexpression du point de vue métrique) $\forall \varepsilon > 0,\quad \exists N \in \mathbb{N},\quad \forall n \geqslant N,\quad u_n \in B(\ell, \varepsilon)$.

(iii) (point de vue topologique) $\forall V \in \mathcal{V}(\ell),\quad \exists N \in \mathbb{N},\quad \forall n \geqslant N,\quad u_n \in V$.

Si ces propriétés sont vérifiées, on dit que (u_n) converge vers ℓ, ou que (u_n) admet ℓ comme limite.

Remarque 3.3.2

Dans le point (i), l'inégalité $\|u_n - \ell\| \leqslant \varepsilon$ peut être exprimée indifféremment avec une inégalité large ou stricte, sans que cela ne modifie le sens de la propriété. C'est ce que suggère l'équivalence avec le point (ii).

Théorème 3.3.3 – Unicité de la limite

Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un e.v.n. E. Si (u_n) admet une limite, alors cette limite est unique.

Cette unicité permet de définir sans ambiguïté LA limite d'une suite (en cas d'existence).

Notation 3.3.4 – Limite d'une suite

Si (u_n) converge vers ℓ, on écrira $\ell = \lim\limits_{n\to+\infty} u_n$, ce qui est défini sans ambiguïté en vertu du théorème précédent.

Exemples 3.3.5

1. Quelle est la limite, pour la norme $\|\cdot\|_\infty$, de la suite $(A_n) \in \mathcal{M}_2(\mathbb{R})^n$ par
$$A_n = \begin{pmatrix} \frac{1}{2^n} & 2 + \frac{1}{n+1} \\ \left(1 + \frac{1}{n}\right)^n & 1 \end{pmatrix}.$$
2. Quelle est la limite dans $\mathbb{R}[X]$ muni de la norme $\|\cdot\|_\infty$ sur les coefficients, de la suite définie

pour $n \geqslant 0$ par

$$P_n = \sum_{k=0}^{n} \frac{1}{(k+1)\binom{n}{k}} X^k$$

3. Quelle est la limite pour la norme de la convergence en moyenne sur $[0,1]$ de la suite

$$f_n : x \mapsto \mathrm{e}^{-nx}.$$

Cette suite admet-elle une limite pour la norme de la convergence uniforme ?

Ainsi, la notion de convergence est fortement dépendante de la norme (en tout cas en dimension infinie).

Proposition 3.3.6 – Convergent implique borné

Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'éléments de E. Si (u_n) est convergente, alors elle est bornée.

Proposition 3.3.7 – Caractérisation de la convergence dans un produit cartésien

Soit $(E_1, N_1), \dots, (E_p, N_p)$ des e.v.n., et $E = E_1 \times \cdots \times E_p$ le produit cartésien, muni de la norme produit. Soit $(X_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$, et pour tout $n \in \mathbb{N}$,

$$X_n = (x_{1,n}, \dots, x_{p,n}).$$

Alors $(X_n)_{n\in\mathbb{N}}$ converge vers $L = (\ell_1, \dots, \ell_p)$ si et seulement si pour tout $k \in [\![1,p]\!]$, $(x_{k,n})_{n\in\mathbb{N}}$ converge vers ℓ_k.

En particulier, on retrouve l'équivalence entre la convergence dans $\mathbb{R}^n$ et la convergence coordonnée par coordonnée pour $\|\cdot\|_\infty$.

Proposition 3.3.8 – Caractérisation par la limite des normes

Soit $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$, $\ell \in E$. La suite (u_n) converge vers ℓ si et seulement si la suite de réels $(\|u_n - \ell\|)_{n\in\mathbb{N}}$ converge vers 0.

En particulier, cela permet d'exploiter un certain nombre de techniques de majorations vues dans le cadre de réels.

Méthode 3.3.9

Pour montrer que $u_n \to \ell$, il suffit de trouver une majoration

$$\forall n \in \mathbb{N}, \quad \|u_n - \ell\| \leqslant M_n,$$

où (M_n) est une suite de réels de limite nulle.

Comme il n'y a *a priori* pas dans E d'autres opérations que l'addition et la multiplication par un scalaire.

Proposition 3.3.10 – Règles opératoires générales

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites d'éléments d'un e.v.n. E, et $(\lambda_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$. Si (u_n), (v_n) et λ_n sont convergentes, alors :

(i) la suite $(u_n + v_n)$ est convergente, et

$$\lim_{n\to+\infty} (u_n + v_n) = \lim_{n\to+\infty} u_n + \lim_{n\to+\infty} v_n;$$

(ii) la suite $(\lambda_n u_n)$ est convergente, et

$$\lim_{n\to+\infty} \lambda_n u_n = \lim_{n\to+\infty \lambda_n} \lim_{n\to+\infty} u_n.$$

◅ **Éléments de preuve.**
Majorer la norme de la différence avec la limite présumée en utilisant l'inégalité triangulaire. ▻

Nous donnons dès maintenant une règle opératoire concernant les produits, mais elle sera démontrée plus tard, nécessitant quelques propriétés liées à la continuité des applications linéaires et multilinéaires.

Proposition 3.3.11 – Limite d'un produit

Soit $(E, \|\cdot\|)$ un e.v.n., muni d'un produit tel que :

(i) $(x,y) \mapsto x \cdot y$ soit bilinéaire ;

(ii) il existe $k \in \mathbb{R}_+$ tel que pour tout $(x,y) \in E$, $\|xy\| \leqslant k \cdot \|x\| \cdot \|y\|$.

Alors, si $(u_n) \in \mathrm{E}^{\mathbb{N}}$ et $(v_n) \in \mathrm{E}^{\mathbb{N}}$ sont deux suites convergentes, il en est de même de $(u_n v_n)$, et de plus,

$$\lim_{n \to +\infty} u_n v_n = \lim_{n \to +\infty} u_n \cdot \lim_{n \to +\infty} v_n.$$

Exemple 3.3.12

Calculer la limite lorsque n tend vers $+\infty$ de $\begin{pmatrix} \cos\left(\frac{1}{n}\right) & \mathrm{e}^{\frac{1}{\sqrt{n}}} \\ \frac{1}{\ln(n)} & 1 + \frac{1}{n^2} \end{pmatrix}^{42}$.

III.2 Caractérisations séquentielles

La convergence des suites permet d'obtenir des caractérisations souvent assez pratiques d'un certain nombre des notions topologiques étudiées jusqu'ici.

Proposition 3.3.13 – Caractérisation séquentielle de l'adhérence

Soit $A \subset E$, et $a \in E$. Les propositions suivantes sont équivalentes:

(i) $a \in \overline{A}$

(ii) il existe une suite $(x_n) \in A^{\mathbb{N}}$ telle que $x_n \to a$.

Corollaire 3.3.14

En particulier, toute suite convergente d'éléments de A converge dans $\overline{A}$.

On en déduit les autres caractérisations qui suivent.

Proposition 3.3.15 – Caractérisation séquentielle de la densité

Soit A et B des parties de E.

(i) A est dense dans B si et seulement si pour tout $b \in B$, il existe $(x_n)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ tel que $x_n \to b$.

(ii) A est partout dense si et seulement si pour tout $b \in E$, il existe $(x_n)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ tel que $x_n \to b$.

Proposition 3.3.16 – Caractérisation des fermés

Soit $F \subset E$. Les propositions suivantes sont équivalentes:

(i) F est fermé ;

(ii) toute suite convergente $(x_n)_{n \in \mathbb{N}} \in F^{\mathbb{N}}$ admet sa limite dans F

Autrement dit, F est fermé si et seulement si F est stable par passage à la limite.
Par restriction à A, on obtient de même une caractérisation des fermés relatifs.

Proposition 3.3.17 – Caractérisation des fermés relatifs

Soit $A \subset E$, et $F \subset A$. Les propositions suivantes sont équivalentes:

(i) F est un fermé relatif de A ;

(ii) toute suite $(x_n)_{n\in\mathbb{N}} \in F^{\mathbb{N}}$ convergente dans A admet sa limite dans F.

IV Valeurs d'adhérence et compacité

IV.1 Valeur d'adhérence d'une suite

Définition 3.4.1 – Suite extraite, fonction extractrice

1. Soit $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. Une *suite extraite* de $(u_n)_{n\in\mathbb{N}}$ est une suite $(v_n)_{n\in\mathbb{N}}$ telle qu'il existe $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$ strictement croissante telle que pour tout $n \in \mathbb{N}$, $v_n = u_{\varphi(n)}$.
2. La fonction φ est appelée *fonction extractrice* de la suite extraite $(v_n)_{n\in\mathbb{N}}$.

Ainsi formellement, une suite extraite de (u_n) est une composée de $u_n : \mathbb{N} \longrightarrow \mathbb{R}$ par une fonction strictement croissante $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$. En pratique, cela signifie que (v_n) est constitué de termes de (u_n), dans l'ordre, et sans répétition d'indice.
Le comportement des suites extraites à l'infini donne des indications quant au comportement de la suite initiale. Si le comportement de la suite initiale détermine le comportement d'une suite extraite, il est beaucoup plus délicat de faire chemin arrière, une suite extraite ne pouvant fournir qu'une information partielle sur la suite totale.

Théorème 3.4.2 – Théorème de convergence des suites extraites

Soit $(u_n)_{n\in\mathbb{N}}$ une suite convergente dans E. Alors toutes les suites extraites de $(u_n)_{n\in\mathbb{N}}$ sont convergentes, de même limite que (u_n).

$\lhd$ **Éléments de preuve.**

La stricte croissance de l'extractrice φ montre que pour tout n, $\varphi(n) \geqslant n$. Le résultat est alors immédiat par la définition de la limite. $\rhd$

Proposition 3.4.3

Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle ou complexe. Alors $(u_n)_{n\in\mathbb{N}}$ converge dans $\overline{\mathbb{R}}$ ou $\mathbb{C}$ si et seulement si $(u_{2n})_{n\in\mathbb{N}}$ et $(u_{2n+1})_{n\in\mathbb{N}}$ convergent vers une même limite ℓ, et dans ce cas, $\lim u_n = \ell$.

C'est un cas particulier de la situation plus générale suivante :

Théorème 3.4.4 – Convergence par étude de suites extraites couvrantes

Soit $(u_n)_{n\in\mathbb{N}}$, et $(\varphi_i)_{i\in I}$ une famille **finie** d'extractrices, telle que $\bigcup_{i\in I} \varphi_i(\mathbb{N}) = \mathbb{N}$. Alors (u_n) converge vers ℓ si et seulement si pour tout $i \in I$, $(u_{\varphi_i(n)})$ converge vers ℓ.

$\lhd$ **Éléments de preuve.**

Sens direct déjà acquis. Pour le sens réciproque, considérer $\varepsilon_i > 0$, associé à chaque φ_i, ainsi qu'un rang de validité N_i, puis se placer au delà du plus grand des N_i. $\rhd$

La notion de suite extraite est intimement liée à celle de valeur d'adhérence :

Définition 3.4.5 – Valeur d'adhérence

Soit $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$, et $a \in E$. On dit que a est une *valeur d'adhérence* de la suite $(u_n)_{n\in\mathbb{N}}$ s'il existe une suite extraite $(u_{\varphi(n)})$ de (u_n) telle que $\lim u_{\varphi(n)} = x$.

Ainsi, l'ensemble des valeurs d'adhérence d'une suite $(u_n)_{n\in\mathbb{N}}$ est l'ensemble de toutes les limites (finies) des suites extraites de $(u_n)_{n\in\mathbb{N}}$.

Exemples 3.4.6

1. Décrire les valeurs d'adhérence de la suite $(-1)^n$.
2. Décrire les valeurs d'adhérence de la suite $\left(\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}^n\right)_{n\in\mathbb{N}}$ dans $\mathcal{M}_n(\mathbb{K})$ muni de la norme $\|\cdot\|_1$.
3. Décrire une suite réelle n'ayant pas de valeur d'adhérence.

Proposition 3.4.7 – Valeurs d'adhérences d'une suite convergente

Soit $(u_n) \in E^{\mathbb{N}}$. Si $u_n \to \ell$, alors ℓ est valeur d'adhérence de (u_n) et c'est la seule.

◁ **Éléments de preuve.**

C'est une réexpression du théorème de convergence des suites extraites. ▷

Ainsi, une suite ayant au moins 2 valeurs d'adhérences est divergente.

Avertissement 3.4.8

La réciproque est fausse !

Exemples 3.4.9

Construire une suite réelle non convergente et admettant une seule valeur d'adhérence.

Proposition 3.4.10 – Valeurs d'adhérence d'une suite extraite

Soit (v_n) une suite extraite de (u_n). L'ensemble des valeurs d'adhérence de (v_n) est inclus dans l'ensemble des valeurs d'adhérence de (u_n).

◁ **Éléments de preuve.**

Faire une double-extraction, ce qui revient à composer les extractrices. Attention au sens dans lequel écrire cette composée. ▷

Proposition 3.4.11 – Caractérisation des valeurs d'adhérence

Soit $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ et $a \in E$. Les propositions suivantes sont équivalentes:

(i) a est valeur d'adhérence de $(u_n)_{n\in\mathbb{N}}$;

(ii) pour tout voisinage V de a, il existe $I \subset \mathbb{N}$ infini tel que pour tout $n \in I$, $u_n \in V$;

(iii) pour tout voisinage V de a, pour tout $N \in \mathbb{N}$, il existe $n \geqslant N$ tel que $u_n \in V$;

(iv) pour tout $\varepsilon > 0$, il existe $I \subset \mathbb{N}$ infini tel que pour tout $n \in I$, $u_n \in B(a, \varepsilon)$;

(v) pour tout $\varepsilon > 0$, pour tout $N \in \mathbb{N}$, il existe $n \geqslant N$ tel que $u_n \in B(a, \varepsilon)$.

◁ **Éléments de preuve.**

- $(i) \Longrightarrow (ii)$: utiliser la définition topologique de la convergence d'une suite extraite vers a ;
- $(ii) \Longrightarrow (iv)$: les boules centrées en a sont des voisinages de a ;
- $(ii) \Longleftrightarrow (iii)$ et $(iv) \Longleftrightarrow (v)$ car un sous-ensemble de $\mathbb{N}$ est infini si et seulement si il est non borné ;
- $(v) \Longrightarrow (i)$ en construisant une suite $(u_{\varphi(n)})$ associée à une suite (ε_n) convergent vers 0.

▷

IV.2 Théorème de Bolzano-Weierstass sur $\mathbb{R}$

Comme l'a montré un exemple précédent, une suite n'admet pas nécessairement de valeur d'adhérence. Le théorème de Bolzano-Weierstrass, que vous avez vu en première année, donne un résultat d'existence pour les suites réelles, sous certaines conditions.

Théorème 3.4.12 – Théorème de Bolzano-Weierstrass réel

De toute suite réelle bornée on peut extraire une suite convergente. En d'autres termes, toute suite réelle bornée admet au moins une valeur d'adhérence.

◁ **Éléments de preuve.**

On peut faire une dichotomie, en gardant toujours une moitié de l'intervalle contenant une infinité de termes de la suite (u_n). ▷

Tout ce qui précède reste évidemment valide, mais on peut également s'intéresser dans ce cadre à la convergence vers $+\infty$ ou $-\infty$, et donc définir des valeurs d'adhérence infinies.

Proposition/Définition 3.4.13 – Valeur d'adhérence infinie

Soit $(u_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. Les propositions suivantes sont équivalentes:

(i) il existe une suite extraite $(u_{\varphi(n)})$ convergent vers $+\infty$;

(ii) pour tout $A \in \mathbb{R}$, il existe une infinité d'indices $n \in \mathbb{N}$ tels que $u_n \geqslant A$;

(iii) pour tout $A \in \mathbb{R}$ et tout $N \in \mathbb{N}$, il existe $n \geqslant N$ tel que une $u_n \geqslant A$;

(iv) (u_n) n'est pas majorée.

On dit dans ce cas que $+\infty$ est une valeur d'adhérence de (u_n) dans $\overline{\mathbb{R}}$.

Proposition 3.4.14 – Existence d'une valeur d'adhérence dans $\overline{\mathbb{R}}$, HP

Toute suite réelle (u_n) admet une valeur d'adhérence au moins dans $\overline{\mathbb{R}}$.

◁ **Éléments de preuve.**

Discuter suivant que (u_n) est bornée ou non, pour pouvoir appliquer le théorème de Bolzano-Weierstrass. ▷

Théorème 3.4.15 – Caractérisation de la convergence par les valeurs d'adhérence, HP

Une suite (u_n) converge dans $\overline{\mathbb{R}}$ si et seulement si elle admet une unique valeur d'adhérence dans $\overline{\mathbb{R}}$.

◁ **Éléments de preuve.**

Sens direct déjà étudié. Sens réciproque : pour tout voisinage V de a (unique valeur d'adhérence), seul un nombre fini de termes de u_n est hors de V (sinon on peut en extraire une suite convergente d'après la proposition précédente, et on obtient une deuxième valeur d'adhérence). ▷

Avertissement 3.4.16

Attention à bien considérer les éventuelles valeurs d'adhérences infinies. Comme on l'a vu sur un exemple, l'existence et l'unicité d'une valeur d'adhérence réelle n'est pas suffisante pour assurer la convergence.

IV.3 Sous-ensembles compacts

Une autre situation permettant d'assurer la caractérisation de la convergence par l'unicité de la valeur d'adhérence est le cas où les valeurs prises par (u_n) se retrouvent toutes dans un ensemble sur lequel on peut utiliser une propriété similaire à la conclusion du théorème de Bolzano-Weierstrass. Cela motive la définition suivante.

Définition 3.4.17 – Sous-ensemble compact

Soit E un e.v.n., et K une partie de E. On dit que K est compact (ou séquentiellement compacte, ou compact au sens de Bolzano-Weierstrass) si et seulement si de toute suite $(u_n)_{n\in\mathbb{N}} \in K$ d'éléments de K, on peut extraire une suite convergeant vers un élément de K.

La propriété définissant la compacité est appelée propriété de Bolzano-Weierstrass.

Exemples 3.4.18

1. Les intervalles $[a, b]$ fermés bornés sont des compacts de $\mathbb{R}$.
2. $\mathbb{U}$ est un compact de $\mathbb{C}$.
3. Un sous-ensemble fini de E est compact.

Propriétés 3.4.19

Soit E un e.v.n., et K un compact.

1. K est fermé et borné.
2. Soit F un fermé relatif de K. Alors F est compact.

Avertissement 3.4.20

On prendra garde que (i) ne permet pas de caractériser les compacts. On montrera que c'est le cas en dimension finie. En revanche, un théorème célèbre (théorème de Riesz) affirme que la boule fermée $\overline{B}(0, 1)$ est compacte si et seulement si E est de dimension finie, ce qui donne un contre-exemple en dimension infinie.

Théorème 3.4.21 – Caractérisation de la convergence par les v.a. sur un compact

Soit E un e.v.n., et K un compact de E. Soit $(u_n)_{n\in\mathbb{N}} \in K^{\mathbb{N}}$. Alors (u_n) converge si et seulement si (u_n) admet une unique valeur d'adhérence.

◁ **Éléments de preuve.**

Si (u_n) a une unique v.a. ℓ et ne converge pas, on peut extraire une suite ne s'approchant pas trop de ℓ. Cette suite doit avoir une v.a. qui ne peut pas être ℓ. ▷

Remarque 3.4.22

Du fait même de sa définition (équivalente à l'existence d'une valeur d'adhérence), la notion de compacité joue un rôle central dans de nombreux problèmes existentiels. L'hypothèse de compacité permet par exemple d'obtenir l'existence d'extrema de fonctions continues.

Théorème 3.4.23 – Produit de compacts

Soit $(E_1, N_1), \dots, (E_p, N_p)$ des e.v.n., et pour tout $i \in [\![1, p]\!]$, K_i un compact de E_i. Alors $K_1 \times \cdots \times K_i$ est un sous-ensemble compact de $E_1 \times \cdots \times E_p$ (muni de la norme produit $\|\cdot\|_\infty$.

◁ **Éléments de preuve.**
Par extractions successives, on en déduit une version plus générale. ▷

Théorème 3.4.24 – Bolzano-Weierstrass pour $\|\cdot\|_\infty$ en dimension finie

Soit E un e.v.n. de dimension finie, de base $\mathcal{B}$, et muni de la norme $\|\cdot\|_{\mathcal{B},\infty}$. Alors toute suite bornée $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ admet au moins une valeur d'adhérence.

◁ **Éléments de preuve.**
C'est un cas particulier du théorème précédent :
- tout d'abord pour récupérer BW sur $\mathbb{C}$
- puis en décomposant E en produit de droites.

À chaque étape, le caractère borné permet de se restreindre à un intervalle fermé borné de la droite considérée, donc compact. ▷

Une variante de ce résultat est le très important résultat suivant :

Théorème 3.4.25 – Compacité de $\overline{B}(a, r)$ pour $\|\cdot\|_{\mathcal{B},\infty}$ en dimension finie

Soit E un espace de dimension finie, et $\mathcal{B}$ une base de E. Dans l'e.v.n. $(E, \|\cdot\|_{\mathcal{B},+\infty})$, $\overline{B}(a, r)$ est compact.

◁ **Éléments de preuve.**
C'est une réexpression du résultat précédent, à quoi on ajoute le caractère fermé de la boule. ▷

V Comparaison de normes

Comme on l'a vu dans plusieurs énoncés (par exemple le théorème de Bolzano-Weierstrass), il est souvent plus pratique de travailler avec une norme particulière. En général, ce n'est pas suffisant pour pouvoir généraliser le résultat pour une norme quelconque. On s'intéresse ici à des conditions permettant de nous assurer que pour l'étude d'un problème particulier, deux normes sont interchangeables (*i.e.* que travailler avec l'une plutôt que l'autre est équivalent).

V.1 Domination

Définition 3.5.1 – Domination de normes

Soit E un espace vectoriel sur $\mathbb{K}$, et N_1 et N_2 deux normes sur E. On dit que N_1 est dominée par N_2 s'il existe $\alpha \in \mathbb{R}_+^*$ tel que

$$\forall x \in E, \quad N_1(x) \leqslant \alpha N_2(x).$$

Proposition 3.5.2 – Caractérisation par l'image de la sphère unité fermée

Soit E un $\mathbb{K}$-espace vectoriel et N_1 et N_2 deux normes. Les propositions suivantes sont équivalentes:

(i) N_1 est dominée par N_2

(ii) Toute partie bornée pour N_2 est également bornée pour N_1

(iii) il existe α tel que pour tout $x \in E$, $N_2(x) \leqslant 1 \Longrightarrow N_1(x) \leqslant \alpha$.
En d'autres termes, la boule fermé unité pour N_2 est bornée pour N_1 ;

(iv) il existe α tel que pour tout $x \in E$, $N_2(x) = 1 \Longrightarrow N_1(x) \leqslant \alpha$.
En d'autres termes, la sphère unité pour N_2 est bornée pour N_1 ;

Proposition 3.5.3 – Transitivité de la relation de domination

Si N_1 est dominée par N_2 et N_2 dominée par N_3, alors N_1 est dominée par N_3.

Exemple 3.5.4

Dans $\mathbb{R}^2$:

- $\|\cdot\|_1$ est dominée par $\|\cdot\|_2$
- $\|\cdot\|_2$ est dominée par $\|\cdot\|_\infty$
- $\|\cdot\|_\infty$ est dominée par $\|\cdot\|_1$.

Plus généralement :

Théorème 3.5.5 – Comparaison des normes 1, 2 et ∞

Soit $(E_1, N_1), \dots, (E_p, N_p)$ des e.v.n., et $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ les trois normes usuelles définies sur le produit cartésien $E_1 \times \cdots \times E_n$. Alors

- $\|\cdot\|_1$ est dominée par $\|\cdot\|_2$
- $\|\cdot\|_2$ est dominée par $\|\cdot\|_\infty$
- $\|\cdot\|_\infty$ est dominée par $\|\cdot\|_1$.

Soit alors E un e.v.n. muni d'une base $\mathcal{B}$. En considérant la décomposition de E comme produit cartésien de droites, on en déduit une comparaison similaire entre $\|\cdot\|_{\mathcal{B},1}$, $\|\cdot\|_{\mathcal{B},2}$ et $\|\cdot\|_{\mathcal{B},\infty}$.

Proposition 3.5.6 – Comparaison des topologies

Soit E un $\mathbb{K}$-ev, et N_1 et N_2 deux normes. On note $\mathcal{V}_i(a)$ l'ensemble des voisinages de a au sens de la norme N_i et $\mathcal{O}_i$ l'ensemble des ouverts au sens de la norme i. Les propositions suivantes sont équivalentes:

(i) N_1 est dominée par N_2

(ii) $\forall a \in E$, $\mathcal{V}_1(a) \subset \mathcal{V}_2(a)$

(iii) $\mathcal{O}_1 \subset \mathcal{O}_2$.

Proposition 3.5.7 – Comparaison des convergences

Soit E un $\mathbb{K}$-ev, et N_1 et N_2 deux normes. Si N_1 est dominée par N_2, alors toute suite convergeant au sens de N_2 converge aussi au sens de N_1, vers la même limite.

◁ **Éléments de preuve.**
Peut se faire soit par majoration soit par les voisinages en utilisant la propriété précédente. ▷

Méthode 3.5.8 – Montrer que N_1 n'est pas dominée par N_2

- Il suffit de trouver une suite bornée pour N_2 qui ne le soit pas pour N_1
- Il suffit de trouver une suite convergente pour N_2 qui ne le soit pas pour N_1.

Exemples 3.5.9

1. Dans $\mathcal{C}^0([0,1])$ $\|\cdot\|_1$ est dominée par $\|\cdot\|_\infty$, mais $\|\cdot\|_\infty$ n'est pas dominée par $\|\cdot\|_1$
2. Montrer que dans $\mathbb{K}[X]$, $\|\cdot\|_\infty$ est dominée par $\|\cdot\|_1$ mais $\|\cdot\|_1$ n'est pas dominée par $\|\cdot\|_\infty$.

V.2 Normes équivalentes

Définition 3.5.10 – Normes équivalentes

On dit que deux normes N_1 et N_2 sur E sont équivalentes si N_1 est dominée par N_2 et N_2 est dominée par N_1.
De manière équivalente, cela revient à dire qu'il existe $(\alpha, \beta) \in \mathbb{R}_+^*$ tels que

$$\forall x \in E, \quad \alpha N_2(x) \leqslant N_1(x) \leqslant \beta N_2(x).$$

Proposition 3.5.11

Cela définit une relation d'équivalence sur l'ensemble des normes sur E.

Proposition 3.5.12

Les comparaisons faites dans le paragraphe précédent amènent :
- les normes $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sont équivalentes sur un produit fini d'e.v.n. ;
- les normes $\|\cdot\|_{\mathcal{B},1}$, $\|\cdot\|_{\mathcal{B},2}$ et $\|\cdot\|_{\mathcal{B},\infty}$ sont équivalentes sur E de dimension finie, muni d'une base $\mathcal{B}$.

Les comparaisons des topologies et propriétés de convergence établies dans le paragraphe précédent amènent :

Théorème 3.5.13 – Comparaison des convergences pour 2 normes équivalentes

Soit N_1 et N_2 deux normes équivalentes sur E et $(u_n) \in E^{\mathbb{N}}$ une suite d'éléments de E. Alors (u_n) est convergente au sens de la norme N_1 si et seulement si elle est convergente au sens de la norme N_2, la valeur de la limite étant la même pour les deux normes.

En particulier, les objets et propriétés topologiques et séquentiels (définis soit avec les ouverts ou voisinages, soit avec des propriétés de convergence) sont les mêmes pour deux normes équivalentes. Notamment, deux normes équivalentes définissent les mêmes compacts.
Ainsi, la compacité, montrée pour un produit cartésien de compacts au sens de la topologie produit (définie avec $\|\cdot\|_\infty$) reste donc aussi valide avec les normes $\|\cdot\|_1$ et $\|\cdot\|_2$.
L'équivalence des normes permet donc de choisir, parmi celles équivalentes à la norme donnée initialement, celle qui est le plus adaptée à la situation. Comme on l'a vu à plusieurs reprises, la norme $\|\cdot\|_\infty$ est souvent assez pratique à manipuler. Mais parfois, cela peut être intéressant de « panacher » les normes au gré des étapes du raisonnement.
Attention toutefois à ne faire cela qu'avec des normes équivalentes !
On verra dans le chapitre suivant que c'est toujours le cas lorsque E est de dimension finie.

VI Espaces métriques (HP)

Comme on l'a évoqué plus haut, on peut définir des distances de façon beaucoup plus générale

Définition 3.6.1 – Distance

Soit A un ensemble quelconque. Une distance d sur A est une application $d : A^2 \to \mathbb{R}$ telle que
(i) (séparation) pour tout $(x, y) \in A^2$, $d(x, y) = 0 \iff x = y$;
(ii) (symétrie) pour tout $(x, y) \in A^2$, $d(x, y) = d(y, x)$
(iii) (inégalité triangulaire) pour tout $(x, y, z) \in A^3$, $d(x, z) \leqslant d(x, y) + d(y, z)$.

Comme on l'a vu, cela implique en particulier la positivité. La propriété 3.2.2 pourquoi la réf ne marche pas bien ? affirme alors que l'application $(x, y) \mapsto \|y - x\|$ est une distance.

Exemples 3.6.2

- Soit A une partie d'un e.v.n.. La distance associée à la norme de E se restreint en une distance sur A.
- Distance entre 2 sommets d'un graphe non orienté connexe (en un seul morceau), comme étant le nombre minimal d'arêtes à parcourir pour aller d'un sommet à l'autre.

Définition 3.6.3 – Espace métrique

Un espace métrique (E, d) est un ensemble E (pas forcément un espace vectoriel), muni d'une distance d sur E.

On peut alors définir dans un espace métrique (E, d), de même que dans un *e.v.n.*, des boules, des voisinages, des ouverts, et des fermés.
Ce point de vue permet de clarifier la définition de la topologie relative (les ouverts et fermés relatifs) vue en cours de chapitre.

Proposition 3.6.4 – Caractérisation de la topologie relative par restriction

Soit $(E, \|\cdot\|)$ un e.v.n., et A une partie de E. Alors la distance d définie par $\|\cdot\|$ se restreint à A en une distance d_A. La topologie associée à l'espace métrique (A, d_A) est alors la topologie relative de A définie en II-4.

On peut également définir une notion de convergence.

Définition 3.6.5 – Convergence dans un espace métrique

Soit (E, d) un espace métrique, et $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. On dit que $(u_n)_{n\in\mathbb{N}}$ converge vers $\ell \in E$ si l'une des propriétés équivalentes suivantes est satisfaite (et donc toutes) :

(i) $\forall \varepsilon > 0, \quad \exists N \in \mathbb{N}, \quad \forall n \in \mathbb{N}, \quad n \geqslant N \Longrightarrow d(u_n, \ell) \leqslant \varepsilon$;

(ii) $\forall \varepsilon > 0, \quad \exists N \in \mathbb{N}, \quad \forall n \in \mathbb{N}, \quad n \geqslant N \Longrightarrow u_n \in B(\ell, \varepsilon)$;

(iii) $\forall V \in \mathcal{V}(\ell), \quad \exists N \in \mathbb{N}, \quad \forall n \in \mathbb{N}, \quad n \geqslant N \Longrightarrow u_n \in V$;

Les caractérisations séquentielles restent alors vraies.

4

Intégrales généralisées

Le but de ce chapitre est de généraliser la définition des intégrales au cas d'intervalles non bornés, ou d'intervalles aux bornes desquels la fonction f n'est pas définie, et n'admet pas de limite finie.
Le procédé est le même que celui qui permet de passer des sommes finies aux séries, à savoir passer à la limite sur les bornes. D'ailleurs, comme vous vous en rendrez compte, l'étude des intégrales généralisées possède des similarités très fortes avec l'étude des séries.
Pour cela, afin de pouvoir définir l'intégrale de f sur un intervalle I quelconque, il est déjà nécessaire de pouvoir définir son intégrale sur tout segment $[a,b] \subset I$. Le programme de première année nous donne une condition suffisante simple à énoncer pour l'intégrabilité sur un segment : la continuité par morceaux de f. Même si ce n'est qu'un condition nécessaire et pas nécessaire, c'est ce cadre que nous adoptons.
Nous rappelons donc la définition importante suivante :

Définition 4.0.1 – Fonction continue par morceaux sur un intervalle I

1. Soit $I = [a,b]$ un intervalle fermé borné. On dit que f est continue par morceaux s'il existe une subdivision (finie)
$$a = \sigma_0 < \sigma_1 < \cdots < \sigma_{n-1} < \sigma_n = b$$
de $[a,b]$ telle que pour tout $k \in [\![0, n-1]\!]$, $f_{|]\sigma_k,\sigma_{k+1}[}$ soit continue et prolongeable par continuité sur $[\sigma_k, \sigma_{k+1}]$.
2. Soit I un intervalle quelconque. On dit que f est continue par morceaux sur I si f est continue par morceaux sur tout segment $[a,b]$ tel que $[a,b] \subset I$.

On notera $\mathcal{C}^0_m([a,b],\mathbb{K})$ l'espace des fonctions continues par morceaux. La notation n'étant pas complètement standard, la redéfinir si vous l'utilisez.

Exemples 4.0.2

1. $x \mapsto \lfloor x \rfloor$ est continue par morceaux sur $\mathbb{R}$.
2. $x \mapsto \begin{cases} \left\lfloor \frac{1}{x} \right\rfloor & \text{si } x > 0 \\ 0 & \text{si } x = 0 \end{cases}$ n'est pas continue par morceaux sur $[0,+\infty[$.
Sa restriction à $]0,+\infty[$ est continue par morceaux.

I Intégrales sur un intervalle quelconque

Dans toute cette partie, on considère un intervalle $I = [a,b[$, où $a \in \mathbb{R}$ et $b \in \mathbb{R} \cup \{+\infty\}$, tels que $a < b$.
Tout ce qu'on fait s'adapte facilement, par symétrie, à un intervalle $]b,a]$.
Le corps $\mathbb{K}$ est égal à $\mathbb{R}$ ou $\mathbb{C}$.

I.1 Intégrale généralisée sur un intervalle $[a, b[$.

Par définition, une fonction continue par morceaux sur I est en particulier continue par morceaux sur tout segment $[c, d]$ inclus dans I, notamment sur des segments $[a, x]$. Elle est donc intégrable (au sens de Riemann) sur ces intervalles fermés bornés. Cela motive la définition suivante.

Définition 4.1.1 – Intégrale généralisée sur $[a, b[$

- Soit $f : [a, b[\to \mathbb{K}$ une application continue par morceaux (c.p.m) sur $[a, b[$. On dit que l'intégrale $\displaystyle\int_a^b f$ est convergente si la fonction

$$x \mapsto \int_a^x f(t)\, \mathrm{d}t,$$

définie sur $[a, b[$, admet une limite finie (*i.e.* dans $\mathbb{K}$) lorsque x tend vers b.

- Lorsque $\displaystyle\int_a^b f$ est convergente, sa valeur est notée d'une des façons suivantes, et définie par :

$$\int_I f = \int_I f(t)\, \mathrm{d}t = \int_a^b f = \int_a^b f(t)\, \mathrm{d}t = \lim_{x \to b^-} \int_a^x f(t)\, \mathrm{d}t.$$

- Si $\displaystyle\int_a^b f$ n'est pas convergente, on dit qu'elle est divergente.

On parle aussi parfois d'intégrale impropre et de borne impropre pour désigner la borne b qui est exclue du domaine.

Exemples 4.1.2

1. $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}t}{t}$ est divergente.
2. $\displaystyle\int_0^{+\infty} \mathrm{e}^{-t}\, \mathrm{d}t$ est convergente.
3. $\displaystyle\int_0^1 \ln(t)\, \mathrm{d}t$ est convergente.

Remarque 4.1.3

La forte ressemblance avec les séries pourrait faire croire que si $\displaystyle\int_a^{+\infty} f$ converge, alors $f(x) \to 0$. Est-ce le cas ?

Une situation où la convergence est simple à obtenir est la suivante :

Proposition 4.1.4 – Intégrale faussement généralisée

On suppose que $b \neq +\infty$. Soit $f : [a, b[\to \mathbb{K}$ une application admettant un prolongement $\tilde{f}$ c.p.m sur $[a, b]$. Alors $\displaystyle\int_a^b f$ est convergente, et

$$\int_a^b f = \int_a^b \tilde{f}.$$

Une telle intégrale est parfois appelée intégrale faussement généralisée.

Exemple 4.1.5

Justifier la convergence de $\displaystyle\int_0^1 \frac{\sin(t)}{t}\, \mathrm{d}t$.

Remarque 4.1.6

En particulier, si f est c.p.m sur $[a,b]$, considérer l'intégrale $\displaystyle\int_a^b f(t)\, \mathrm{d}t$ au sens de Riemann, ou au sens de l'intégrale généralisée de $f_{|[a,b[}$ revient au même. Cela donne une certaine cohérence à la notation, et permet de s'assurer que la relation

$$\int_a^b f(t)\, \mathrm{d}t = \lim_{x\to b^-} \int_a^x f(t)\, \mathrm{d}t$$

reste aussi vrai pour des fonctions déjà c.p.m. (donc Riemann-intégrables) sur $[a,b]$.

Dans la suite, on reprend $b \in \mathbb{R} \cup \{+\infty\}$.

Proposition 4.1.7 – Linéarité

Soit $f, g \in \mathcal{C}^0_m([a,b[,\mathbb{K})$ et $\lambda \in \mathbb{K}$.

1. Si $\displaystyle\int_a^b f$ et $\displaystyle\int_a^b g$ sont convergentes, alors $\displaystyle\int_a^b f + \lambda g$ aussi, et
$$\int_a^b f + \lambda g = \int_a^b f + \lambda \int_a^b g.$$
2. Si $\displaystyle\int_a^b f$ est convergente et $\displaystyle\int_a^b g$ est divergente, alors $\displaystyle\int_a^b (f+g)$ est divergente.
3. Si $\displaystyle\int_a^b f$ et $\displaystyle\int_a^b g$ sont divergentes, on ne peut rien dire en toute généralité de $\displaystyle\int_a^b (f+g)$.

Proposition 4.1.8 – Positivité et croissance

Soit $f, g : [a,b[\to \mathbb{K}$ deux applications c.p.m. sur $[a,b[$.

1. (Positivité) Si f est positive sur $I = [a,b[$, et $\displaystyle\int_a^b f$ est convergente, alors $\displaystyle\int_a^b f(t)\, \mathrm{d}t \geqslant 0$.
2. (Croissance) Si $f \leqslant g$ sur I, et si $\displaystyle\int_a^b f$ et $\displaystyle\int_a^b g$ sont convergentes, alors
$$\int_a^b f \leqslant \int_a^b g.$$
3. (Stricte positivité) Si f est continue et positive sur $[a,b[$, non partout nulle, et d'intégrale convergente, alors
$$\int_a^b f > 0.$$

Remarque 4.1.9

La propriété de stricte positivité s'utilise souvent dans sa version contraposée : si f est continue et positive sur I, d'intégrale (convergente) nulle, alors f est identiquement nulle.

Proposition 4.1.10 – Chasles

Soit $f : [a, b[\to \mathbb{K}$ une application c.p.m. sur $[a, b[$, et $c \in [a, b[$. Alors :

1. L'intégrale $\displaystyle\int_a^b f$ est convergente si et seulement si $\displaystyle\int_c^b f$ est convergente.
2. Le cas échéant, on a alors

$$\int_a^b f(t)\ \mathrm{d}t = \int_a^c f(t)\ \mathrm{d}t + \int_c^b f(t)\ \mathrm{d}t.$$

Proposition 4.1.11 – Intégrale dépendant de sa borne non impropre

Soit $f : [a, b[\to \mathbb{K}$ une fonction continue telle que $\displaystyle\int_a^b f$ soit convergente. Alors la fonction

$$F : x \mapsto \int_x^b f(t)\ \mathrm{d}t$$

est bien définie sur $[a, b[$. De plus, F est dérivable sur $[a, b[$, et

$$\forall x \in [a, b[, \quad F'(x) = -f(x).$$

Attention au signe moins, provenant du fait que la dépendance se fait sur la borne inférieure de l'intégrale.

Remarques 4.1.12

1. Attention à l'hypothèse de continuité, la c.p.m ne suffit pas ici !
2. Par composition, on en déduit des formules de dérivation pour $x \mapsto \displaystyle\int_{\alpha(x)}^b f(t)\ \mathrm{d}t$, comme dans le cas d'intégrales définies.
3. Le signe disparaît lorsqu'on adapte à l'intervalle $]b, a]$: la dérivée de $F : x \mapsto \displaystyle\int_b^x f$ est dans ce cas $f(x)$.

I.2 Intégrales de référence

Comme pour le cas des séries, on pourra souvent se ramener à l'étude de la convergence d'intégrales de fonctions positives. Il est donc intéressant de pouvoir disposer de techniques efficaces d'études de convergence dans ce contexte.
Souvent ces techniques seront non calculatoires, dans le sens où elles ne passent pas par le calcul des intégrales partielles et de leurs limites. On privilégiera souvent des méthodes de comparaison, comme dans le cas de séries.
Pour cela, il est nécessaire de disposer d'un certain nombre de séries de référence dont on connaît les propriétés de convergence, afin de pouvoir utiliser pour mener nos comparaisons.

Proposition 4.1.13 – Intégrales de référence en la borne $+\infty$

1. Intégrale de Riemann : pour $\alpha \in \mathbb{R}$, $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t^\alpha}$ converge si et seulement si $\alpha > 1$.
2. Intégrale exponentielle : pour $a \in \mathbb{R}$, $\displaystyle\int_0^{+\infty} \mathrm{e}^{-at}\ \mathrm{d}t$ converge si et seulement si $a > 0$.

On adapte ces exemples pour la borne $-\infty$:

Proposition 4.1.14 – Intégrales de référence en la borne $-\infty$

1. Intégrale de Riemann : pour $\alpha \in \mathbb{R}$, $\displaystyle\int_{-\infty}^{-1} \frac{\mathrm{d}t}{|t|^\alpha}$ converge si et seulement si $\alpha > 1$.
2. Intégrale exponentielle : pour $a \in \mathbb{R}$, $\displaystyle\int_{-\infty}^{0} \mathrm{e}^{-at}\,\mathrm{d}t$ converge si et seulement si $a < 0$.

Proposition 4.1.15 – Intégrale de Riemann en une borne finie b

1. Pour $\alpha \in \mathbb{R}$, l'intégrale de Riemann $\displaystyle\int_{a}^{b} \frac{1}{|t-b|^\alpha}\,\mathrm{d}t$ converge si et seulement si $\alpha < 1$.
2. Pour $\alpha \in \mathbb{R}$, l'intégrale de Riemann $\displaystyle\int_{b}^{a} \frac{1}{|t-b|^\alpha}\,\mathrm{d}t$ converge si et seulement si $\alpha < 1$.

Exemples 4.1.16

- $\displaystyle\int_0^1 \frac{\mathrm{d}t}{\sqrt{t}}$ est convergente, alors que $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{\sqrt{t}}$ est divergente.
- $\displaystyle\int_0^1 \frac{\mathrm{d}t}{t^2}$ est divergente, alors que $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t^2}$ est convergente.
- $\displaystyle\int_0^1 \frac{\mathrm{d}t}{t}$ et $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t}$ sont toutes deux divergentes.

Nous verrons un peu plus tard d'autres intégrales de référence, appelées intégrales de Bertrand, de la forme $\displaystyle\int_1^{+\infty} \frac{1}{t^\alpha \ln(t)^\beta}\,\mathrm{d}t$, qui peuvent être utiles pour affiner un peu les comparaisons (notamment quand α devient proche de 1, ou égal à 1). Mais ces intégrales ne sont pas explicitement au programme, et il faudra savoir redonner les justifications idoines au cas par cas.
Un autre exemple qu'on peut obtenir par un calcul explicite, mais qui sert assez rarement d'intégrale de référence du fait de son comportement peu marqué, est l'intégrale du logarithme en 0. Comme ce n'est pas une intégrale de référence officielle du programme, elle n'a que le statut d'exemple.

Exemple 4.1.17

L'intégrale $\displaystyle\int_0^1 \ln(t)\,\mathrm{d}t$ est convergente.

I.3 Théorèmes de comparaison pour les intégrales de fonctions positives

Les séries de référence ont été étudiées de façon purement calculatoire en revenant à la définition : le calcul explicite de l'intégrale pertielle a permis dans chacun des cas d'étudier l'existence d'une limite.
Assez rapidement, cette méthode devient insuffisante pour l'étude de la convergence des intégrales, soit parce que les calculs deviennent trop complexes, soit (cela arrive), parce qu'on n'est pas en mesure de calculer explicitement l'intégrale à étudier, en se servant des fonctions usuelles.
Les techniques de comparaisons nous permettrons de contourner ce point de vue calculatoire, exactement comme lors de l'étude des séries.
On travaille toujours sur un intervalle $I = [a, b[$, $b \in \mathbb{R} \cup \{+\infty\}$.

Proposition 4.1.18 – Bonne définition dans $\overline{\mathbb{R}}_+$

Soit $f : [a, b[\to \mathbb{R}_+$ c.p.m. et positive. Alors $x \mapsto \displaystyle\int_a^x f(t)\,\mathrm{d}t$ admet toujours une limite, finie ou infinie. Dans le cas oà la limite est infinie, on s'autorisera la notation

$$\int_a^b f(t)\ \mathrm{d}t = +\infty.$$

Remarque 4.1.19

1. Dans le cas positif, on peut donc considérer $\displaystyle\int_a^b f(t)\ \mathrm{d}t$ avant toute étude de convergence
2. Le programme stipule expressément qu' « un calcul aboutissant à un résultat fini vaut preuve de convergence ». Autrement dit, cela autorise à mener une méthode calculatoire sans se préoccuper de la convergence.
3. Cela permet aussi de justifier la convergence par majoration en obtenant après calculs une majoration stricte par $+\infty$ (ce qui utilise implicitement les théorèmes de comparaison qu'on va établir ci-dessous).

Lemme 4.1.20 – CNS de convergence d'une intégrale de fonction positive

Soit $f : [a, b[\to \mathbb{R}$ une application c.p.m. sur $[a, b[$, et positive. Les propositions suivantes sont équivalentes:

(i) $\displaystyle\int_a^b f$ converge

(ii) $\displaystyle x \mapsto \int_a^x f(t)\ \mathrm{d}t$ est bornée.

Théorème 4.1.21 – Théorème de comparaison des intégrales de fonctions positives

Soit $f, g : [a, b[\to \mathbb{K}$ c.p.m. sur $[a, b[$ et telles que $0 \leqslant f \leqslant g$ sur $[a, b[$. Alors :

1. Si $\displaystyle\int_a^b g$ est convergente, il en est de même de $\displaystyle\int_a^b f$.
2. Si $\displaystyle\int_a^b f$ est divergente, il en est de même de $\displaystyle\int_a^b g$.

Remarque 4.1.22

Il suffit d'avoir la comparaison $0 \leqslant f \leqslant g$ au voisinage à gauche de b seulement. Par ailleurs, les intervalles de définition de f et g n'ont pas besoin d'avoir la même borne inférieure.

Corollaire 4.1.23 – Théorème de comparaison par domination ou négligeabilité

1. Soit $f, g : [a, b[\to \mathbb{K}$ c.p.m. sur $[a, b[$, positives, et telles que $f(x) \underset{x\to b}{=} O(g(x))$.

   Alors la convergence de $\displaystyle\int_a^b g$ entraîne celle de $\displaystyle\int_a^b f$.
2. Cela reste vrai si on remplace l'hypothèse de domination par $f(x) \underset{x\to b}{=} o(g(x))$

Exemple 4.1.24 – Intégrale de Gauss

Justifier la convergence de $\displaystyle\int_0^{+\infty} \mathrm{e}^{-t^2}\ \mathrm{d}t$ et de $\displaystyle\int_{-\infty}^0 \mathrm{e}^{-t^2}\ \mathrm{d}t$ (les deux moitiés de l'intégrale de Gauss $\displaystyle\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2}\ \mathrm{d}t$).

Corollaire 4.1.25 – Théorème de comparaison par équivalences

Soit $f, g : [a, b[\to \mathbb{K}$ c.p.m. sur $[a, b[$, positives, et telles que $f(x) \underset{x\to b}{\sim} g(x)$.

Alors $\displaystyle\int_a^b f$ et $\displaystyle\int_a^b g$ ont même nature.

Exemple 4.1.26

Pour quelles valeurs de x l'intégrale

$$\int_0^1 t^{x-1}\mathrm{e}^{-t}\ \mathrm{d}t$$

est-elle convergente ?

I.4 Extension aux intégrales sur un intervalle quelconque I

Avant de voir comment se ramener à l'étude d'intégrales positives via l'étude de la convergence absolue, on étend les différents concepts et résultats vus jusqu'à présent au cas d'intégrales de fonctions définies sur un intervalle quelconque.

Comme on l'a dit le cas des intervalles semi-ouverts $]a, b]$ est similaire à celui exposé ci-dessus, en remplaçant le passage à la limite à gauche en b par le passage à la limite à droite en a sur la borne inférieure de l'intégrale.

Comme le cas des intervalles fermés bornés $[a, b]$ est du ressort du programme de MPSI (ce ne sont pas des intégrales généralisées, et il n'y a pas de passage à la limite nécessaire, même si ce ne serait pas fautif d'après la remarque 4.1.6), il reste à étudier le cas des intervalles ouverts $]a, b[$, avec $a < b$ dans $\overline{\mathbb{R}}$.

Proposition/Définition 4.1.27 – Intégrale généralisée sur un intervalle ouvert

Soit $(a, b) \in \overline{\mathbb{R}}^2$ tels que $a < b$, et soit $f :]a, b[$ une application c.p.m. sur $]a, b[$.

- On dit que $\displaystyle\int_a^b f(t)\ \mathrm{d}t$ est convergente si et seulement si il existe $c \in]a, b[$ tel que $\displaystyle\int_a^c f(t)\ \mathrm{d}t$ soit convergente (en la borne a) et $\displaystyle\int_c^b f(t)\ \mathrm{d}t$ soit convergente (en la borne b).
- Cette propriété est alors indépendante du choix de $c \in]a, b[$
- On définit alors l'intégale de f sur $]a, b[$ par

$$\int_a^b f(t)\ \mathrm{d}t = \int_a^c f(t)\ \mathrm{d}t + \int_c^b f(t)\ \mathrm{d}t.$$

- Cette définition est indépendante de c, et peut se réécrire

$$\int_a^b f(t)\ \mathrm{d}t = \lim_{x\to a^+} \lim_{y\to b^-} \int_x^y f(t)\ \mathrm{d}t.$$

Exemples 4.1.28 – Gauss, fonction Γ d'Euler

1. Intégrale de Gauss : $\displaystyle\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2}\ \mathrm{d}t$ est convergente.
2. Fonction Γ d'Euler : $\Gamma(x) = \displaystyle\int_0^{+\infty} t^{x-1}\mathrm{e}^{-t}\ \mathrm{d}t$.

   Quel est le domaine de définition de Γ ?

Avertissement 4.1.29

Dans l'expression sous forme de double limite, les deux bornes doivent bien être indépendante. Si on lie les bornes, il peut y avoir des compensations qui faussent l'étude.

Exemple 4.1.30

L'intégrale $\displaystyle\int_{-\infty}^{+\infty} \sin(t)\ \mathrm{d}t$ est divergente, mais $\displaystyle\lim_{x\to+\infty}\int_{-x}^{x} \sin(t)\ \mathrm{d}t = 0$.

Les propriétés vues dans le cadre des intégrales généralisées en une seule borne s'étendent facilement.

Théorème 4.1.31 – Chasles

Soit I un intervalle quelconque, d'extrémités $a < b$ dans $\overline{\mathbb{R}}$, et c tel que $a < c < b$. Alors :

1. $\displaystyle\int_a^b f(t)\ \mathrm{d}t$ est convergente si eu seulement si $\displaystyle\int_a^c f(t)\ \mathrm{d}t$ et $\displaystyle\int_c^b f(t)\ \mathrm{d}t$ le sont.
2. Dans ce cas,
$$\int_a^b f(t)\ \mathrm{d}t = \int_a^c f(t)\ \mathrm{d}t + \int_c^b f(t)\ \mathrm{d}t.$$

◁ **Éléments de preuve.**

C'est juste un théorème bilan regroupant tous les cas qu'on a déjà vus.

- Si I est fermé borné, c'est la relation de Chasles usuelle pour les fonctions Riemann-intégrables sur un segment
- Si I est semi-ouvert, c'est le théorème 4.1.10.
- Si I est ouvert, c'est la définition de la convergence 4.1.27

▷

Proposition 4.1.32 – Linéarité

Soit I un intervalle quelconque de bornes $a < b$ dans $\overline{\mathbb{R}}$. Soit $f, g \in \mathcal{C}^0_m(I, \mathbb{K})$ et $\lambda \in \mathbb{K}$.

1. Si $\displaystyle\int_a^b f$ et $\displaystyle\int_a^b g$ sont convergentes, alors $\displaystyle\int_a^b f + \lambda g$ aussi, et
$$\int_a^b f + \lambda g = \int_a^b f + \lambda \int_a^b g.$$
2. Si $\displaystyle\int_a^b f$ est convergente et $\displaystyle\int_a^b g$ est divergente, alors $\displaystyle\int_a^b (f+g)$ est divergente.

Proposition 4.1.33 – Positivité et croissance

Soit I un intervalle quelconque de bornes $a < b$ dans $\overline{\mathbb{R}}$. Soit $f, g : I \to \mathbb{R}$ deux applications c.p.m. sur I.

1. (Positivité) Si f est positive sur I, et $\displaystyle\int_a^b f$ est convergente, alors $\displaystyle\int_a^b f(t)\ \mathrm{d}t \geqslant 0$.
2. (Croissance) Si $f \leqslant g$ sur I, et si $\displaystyle\int_a^b f$ et $\displaystyle\int_a^b g$ sont convergentes, alors $\displaystyle\int_a^b f \leqslant \int_a^b g$.
3. (Stricte positivité) Si f est continue et positive sur I, non partout nulle, et d'intégrale convergente, alors
$$\int_a^b f > 0.$$

I.5 Intégrabilité

Comme pour les séries, l'étude de la convergence d'une intégrale peut parfois s'étudier en étudiant d'abord la convergence de l'intégrale de son module. Cela permet de ramener l'étude au cas d'une intégrale d'une fonction positive, et donc d'exploiter les méthodes du paragraphe précédent, notamment les techniques de comparaison.

Définition 4.1.34 – Intégrabilité

1. Soit $f : I \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). On dit que f est intégrable ou que $\displaystyle\int_I f(t)\, dt$ est absolument convergente, si :
   (i) f est continue par morceaux sur I ;
   (ii) $\displaystyle\int_I |f(t)|\, dt$ est convergente.
2. Soit f définie sur un intervalle I tel que $\sup I = b$, et $b \notin I$. On dit que f est intégrable en b s'il existe $c \in I$ tel que $f_{|[c,b[}$ soit intégrable.
3. De même si $a = \inf(I)$, on dit que f est intégrable en a s'il existe $c \in I$ tel que $f_{]a,c]}$ soit intégrable.

Avertissement 4.1.35

- Attention à bien avoir la bonne terminologie, l'intégrabilité d'une fonction se définit par la convergence de $\displaystyle\int_I |f|$ et non de $\displaystyle\int_I f$.
- De plus, la notion d'intégrabilité porte sur la fonction, et non sur l'intégrale
- Faire le parallèle avec la notion de sommabilité des familles $(a_i)_{i\in I}$!

Remarques 4.1.36

1. Si f est c.p.m sur $[a, b[$ et de signe constant, la convergence de $\displaystyle\int_I f$ équivaut à l'intégrabilité de f.
   Cela reste en particulier vrai si f est de signe constant au voisinage de b.
2. La notion d'intégrabilité en b est une notion locale, qui permet d'évacuer momentanément les problèmes qu'il pourrait y avoir à un autre endroit de l'intervalle I.

Proposition 4.1.37 – Intégrabilité aux bornes vs intégrabilité globale

Soit $f : I \to \mathbb{K}$, où I est un intervalle de bornes a et b. On suppose que f est c.p.m sur I. Les propositions suivantes sont équivalentes:

- f est intégrable sur I
- f est intégrable en a et en b.

Remarque 4.1.38

Avec les notations de la proposition, si $a \in I$, l'intégrabilité en a est acquise, et il ne reste dans le second point que l'intégrabilité en b à étudier.

La propriété importante permettant de ramener dans beaucoup de situations l'étude de la convergence d'une intégrale à l'étude de l'intégrabilité de l'intégrande est similaire à la propriété reliant la convergence absolue et la convergence des séries.

Proposition 4.1.39 – CVA implique CV

Soit $f : I \to \mathbb{K}$. Si f est intégrable, alors $\displaystyle\int_I f(t)\ \mathrm{d}t$ est convergente.

◁ **Éléments de preuve.**

- On commencer par étudier le cas $\mathbb{K} = \mathbb{R}$, en décomposant $f = f^+ - f^-$, et en utilisant le fait que $0 \leqslant f^+, f^- \leqslant |f|$. Le théorème de comparaison et la linéarité permettent de conclure.
- On passe de même de $\mathbb{R}$ à $\mathbb{C}$ en décomposant f en partie réelle et partie imaginaire.

▷

Exemple 4.1.40

Montrer la convergence de $\displaystyle\int_1^{+\infty} \frac{\cos(t)}{t^2}\ \mathrm{d}t$

Avertissement 4.1.41

Attention, comme pour les séries, la réciproque est fausse : il peut exister des fonctions f telles que $\displaystyle\int_a^b f(t)\ \mathrm{d}t$ converge sans que f soit intégrable. On dira dans ce cas que $\displaystyle\int_a^b f(t)\ \mathrm{d}t$ est semi-convergente.

Exemple 4.1.42

- En minorant $\displaystyle\int_{n\pi+\frac{\pi}{6}}^{(n+1)\pi-\frac{\pi}{6}} \frac{|\sin(x)|}{x}\ \mathrm{d}x$, montrer que $f : x \mapsto \dfrac{\sin(x)}{x}$ n'est pas intégrable sur $[1, +\infty[$
- En effectuant une IPP sur l'intégrale partielle, montrer que $\displaystyle\int_1^{+\infty} \frac{\sin(x)}{x}\ \mathrm{d}x$ est semi-convergente.

Ainsi, l'étude d'une intégale généralisée passe souvent par l'étude d'une intégrale de fonction positive. Pour cette raison, il peut être commode de remarquer qu'on peut toujours donner une valeur dans $\overline{\mathbb{R}}$ à l'intégrale d'une fonction positive :

Proposition/Définition 4.1.43 – Intégrale d'une fonction positive non intégrable

Soit I un intervalle quelconque, de bornes $a < b$ dans $\overline{\mathbb{R}}$. Si f est c.p.m. sur I, positive, et non intégrable, alors

$$\lim_{x\to a^+} \lim_{y\to b^-} f(t)\ \mathrm{d}t = +\infty.$$

On s'autorise à écrire dans ce cas :

$$\int_a^b f(t)\ \mathrm{d}t = +\infty.$$

Remarques 4.1.44

1. Selon les termes du programme, « pour une fonction à valeurs dans $\mathbb{R}_+$, un calcul aboutissant à un résultat fini vaut preuve de convergence ».
2. Cela signifie concrètement que si f est positive, on peut commencer à manipuler (calculer ou majorer) $\displaystyle\int_a^b f$ sans s'être préoccupé au préalable de la convergence de cette intégrale
3. Cette remarque vaut notamment pour l'étude de l'intégrabilité de f. On peut toujours considérer $\displaystyle\int_a^b |f|$, sans justification préalable de sa convergence. Dans ce contexte, selon les termes du programme officiel, « un calcul montrant que $\displaystyle\int_I |f| < +\infty$ vaut preuve d'intégrabilité ».

4. Dans ce cadre, les manipulations liées à la linéarité et aux majorations restent valides dans $\overline{\mathbb{R}}$ quelle que soit la nature des intégrales, ce qui découle de la proposition suivante.

Proposition 4.1.45 – Validité des manipulations usuelles dans $\overline{\mathbb{R}}$

Soit $f, g : I \to \mathbb{R}$ des fonctions c.p.m et positives, et $\lambda \in \mathbb{R}_+^*$.

1. Dans $\overline{\mathbb{R}}_+$, l'égalité suivante est toujours valide :
$$\int_I f(t) + \lambda g(t)\ \mathrm{d}t = \int_I f(t) + \lambda \int_I g(t)\ \mathrm{d}t.$$
2. Si $f \leqslant g$ sur I, alors dans $\overline{\mathbb{R}}_+$, $\displaystyle\int_I f(t)\ \mathrm{d}t \leqslant \int_I g(t)\ \mathrm{d}t$.

◁ **Éléments de preuve.**

Dans le cas où tout converge, c'est déjà vu, et dans les autres cas, il suffit d'exprimer le fait que les divergences se font toujours vers $+\infty$ dans le cadre positif.

- Pour la linéarité, il s'agit aussi de constater que si les deux intégrales divergent, puisque $\lambda > 0$, il ne peut pas y avoir de compensation entre les deux termes, et $\displaystyle\int_I f + \lambda g$ est nécessairement divergente aussi.
- Pour la comparaison, utiliser les théorèmes de comparaison pour l'étude des convergences.

▷

Les théorèmes de comparaison s'adaptent pour l'étude de l'intégrabilité en une borne.

Théorème 4.1.46 – Théorèmes de comparaison pour l'intégrabilité

Soit $f, g : I \to \mathbb{K}$ continues par morceaux, et $b = \sup I \in \overline{\mathbb{R}}$.

1. Si $f(x) \underset{b}{=} O(g(x))$, alors l'intégrabilité de g en b implique celle de f en b
2. Si $f(x) \underset{b}{=} o(g(x))$, alors l'intégrabilité de g en b implique celle de f en b.
3. Si $f(x) \underset{b}{\sim} g(x)$, l'intégrabilité de f en b équivaut à celle de g.

Ce théorème s'adapte bien entendu à la borne inférieure $a = \inf(I)$.

Exemple 4.1.47

Montrer l'intégrabilité sur $]0, +\infty[$ de $f : x \mapsto \dfrac{(\mathrm{e}^{-\frac{1}{t}} - 1)\sin(t)}{t}$.

Définition 4.1.48 – Ensemble $L^1(I, \mathbb{K})$

On note $L^1(I, \mathbb{K})$ l'ensemble des fonctions intégrables sur l'intervalle I. C'est donc une partie de $\mathcal{C}_m^0(I, \mathbb{K})$.

Proposition 4.1.49 – Espace $L^1(I, \mathbb{K})$

$L^1(I, \mathbb{K})$ est un sous-espace vectoriel de $\mathcal{C}_m^0(I, \mathbb{K})$.

Proposition 4.1.50 – Inégalité triangulaire dans $L^1(I, \mathbb{K})$

Soit $f \in L^1(I, \mathbb{K})$. Alors
$$\left|\int_I f\right| \leqslant \int_I |f|.$$

On note dans la suite de ce paragraphe $L_c^1(I,\mathbb{K})$ le sous-espace de $L^1(I,\mathbb{K})$ constitué des fonctions intégrables sur I et continues. La notation n'est pas standard et n'apparaît pas dans le programme officiel. Il faut la réintroduire si vous l'utilisez.
Le résultat suivant a déjà été vu lorsque $I = [a,b[$. Le cas général s'en déduit en recollant deux intervalles semi-ouverts.

Proposition 4.1.51 – Positivité, et stricte positivité

Soit $f \in L_c^1(I,\mathbb{R})$ telle que $f \geqslant 0$. Alors $\displaystyle\int_I f \geqslant 0$, le cas d'égalité étant obtenu si et seulement si $f = 0$.

Corollaire 4.1.52 – E.v.n $(L_c^1(I,\mathbb{K}), \|\cdot\|_1)$

On définit, pour tout $f \in L_c^1(I,\mathbb{K})$,

$$\|f\|_1 = \int_I |f|.$$

Alors $\|\cdot\|_1$ est une norme sur $L_c^1(I,\mathbb{K})$.

I.6 Bilan méthodologique

Méthode 4.1.53 – Étudier la convergence d'une intégrale

Soit $f :]a,b[\to \mathbb{K}$

- Étudier la c.p.m. sur $I =]a,b[$, voir si éventuellement f est (ou peut se prolonger en) une fonction c.p.m. sur $[a,b]$ (dans ce cas, il n'y a rien à faire) ou sur $[a,b[$ (ce qui évite d'avoir à couper en 2)
- S'il y a un problème de c.p.m. au milieu de l'intervalle, voir si on peut se ramener à des fonctions c.p.m. en subdivisant davantage l'intervalle.
- Commencer par l'étude de la borne en laquelle la convergence vous paraît la plus douteuse. S'il y a divergence en cette borne, il ne sera pas nécessaire d'étudier l'autre borne.
- Commencer par étudier l'intégrabilité en cette borne, notamment par comparaisons, puis en l'autre.
- Si la fonction n'est pas intégrable en l'une des 2 bornes, étudier la semi-convergence (il peut être intéressant de se ramener à un intervalle $[c,b[$ ou $]a,c]$, pour isoler les 2 bornes à étudier). Il n'y a pas de méthode standard, mais on pourra penser à l'une de ces possibilités :
  - ∗ faire un calcul explicite ;
  - ∗ faire une (ou plusieurs) intégration(s) par parties de sorte à diminuer l'ordre de grandeur de l'intégrande, dans l'espoir de se ramener à une fonction intégrable. C'est ce qu'on a fait pour $\displaystyle\int_1^{+\infty} \frac{\sin(t)}{t}\,\mathrm{d}t$;
  - ∗ couper le domaine en tranches sur lesquelles f reste de signe constant, pour ensuite étudier la somme des intégrales sur ces tranches. Cela nous ramène à un problème sur les séries. Si la série obtenue diverge, c'est terminé, sinon, il faut encore contrôler ce qui se passe entre les points de la subdivision.

II Méthodes calculatoires

Cette section a pour but d'étendre au cas des intégrales généralisées les méthodes calculatoires établies pour les intégrales sur un segment, à savoir les changements de variable, et les intégrations par parties.
Dans cette section, I désigne un intervalle de bornes $a < b$ dans $\overline{\mathbb{R}}$, les bornes étant indifféremment ouvertes ou fermées. On désigne toujours par $\mathbb{K}$ le corps $\mathbb{R}$ ou $\mathbb{C}$.

II.1 Changement de variables

Théorème 4.2.1 – Formule de changement de variable pour les intégrales généralisées

Soit $f : I \to \mathbb{K}$ et $\varphi : J \to I$, où J est un autre intervalle de $\mathbb{R}$, d'extrémités $\alpha < \beta$. On suppose que :

(i) f est continue sur I ;

(ii) φ est bijective, strictement croissante et de classe $\mathcal{C}^1$ sur J.

Alors les intégrales $\displaystyle\int_I f(x)\,\mathrm{d}x$ et $\displaystyle\int_J f(\varphi(t))\varphi'(t)\,\mathrm{d}t$ sont de même nature, et en cas de convergence,

$$\int_I f(x)\,\mathrm{d}x = \int_J f(\varphi(t))\varphi'(t)\,\mathrm{d}t, \qquad \text{i.e.} \qquad \int_a^b f(x)\,\mathrm{d}x = \int_\alpha^\beta f(\varphi(t))\varphi'(t)\,\mathrm{d}t.$$

◁ **Éléments de preuve.**

Le théorème de la bijection assure que $\lim\limits_{x\to a} \varphi^{-1}(x) = \alpha$ et $\lim\limits_{x\to b} \varphi^{-1}(b) = \beta$.

Ainsi, faire d'abord le changement de variable sur l'intégrale partielle $\int_A^B f(x)\,\mathrm{d}x$, pour $a < A < B < b$, puis faire tendre A vers a et B vers b. ▷

Corollaire 4.2.2 – Adaptation pour un changement de variable strictement décroissant

Soit $f : I \to \mathbb{K}$ et $\varphi : J \to I$, où J est un autre intervalle de $\mathbb{R}$, d'extrémités $\alpha < \beta$. On suppose que :

(i) f est continue sur I ;

(ii) φ est bijective, strictement décroissante et de classe $\mathcal{C}^1$ sur J.

Alors les intégrales $\displaystyle\int_I f(x)\,\mathrm{d}x$ et $\displaystyle\int_J f(\varphi(t))\varphi'(t)\,\mathrm{d}t$ sont de même nature, et en cas de convergence,

$$\int_I f(x)\,\mathrm{d}x = -\int_J f(\varphi(t))\varphi'(t)\,\mathrm{d}t, \qquad \text{i.e.} \qquad \int_a^b f(x)\,\mathrm{d}x = \int_\alpha^\beta f(\varphi(t))\varphi'(t)\,\mathrm{d}t.$$

Exemple 4.2.3

On admet la valeur de l'intégrale de Gauss : $\displaystyle\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t = \sqrt{\pi}$.

1. Que vaut $\displaystyle\int_0^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t$?
2. En déduire la valeur de $\Gamma\left(\dfrac{1}{2}\right) = \displaystyle\int_0^{+\infty} \frac{e^{-t}}{\sqrt{t}}\,\mathrm{d}t$.

Remarques 4.2.4

- Selon les termes du programme, « on applique ce résultat sans justification dans des cas de changements de variables usuels », ce qui est très vague. Cela englobe en tout cas les changements de variable affine. Pour le reste... c'est flou...
- Le théorème de changement de variable permet aussi de comparer la nature des intégales. Donc c'est aussi un moyen d'étude de convergence.

Exemple 4.2.5

Déterminer la nature de $\displaystyle\int_2^{+\infty} \frac{1}{x\ln(x)^\beta}\,\mathrm{d}x$ (intégrale de Bertrand)

Plus généralement, on peut énoncer :

Proposition 4.2.6 – Intégrales de Bertrand, HP

- $x \mapsto \dfrac{1}{x^\alpha \ln(x)^\beta}$ est intégrable en $+\infty$ si et seulement si $\alpha > 1$, ou ($\alpha = 1$ et $\beta > 1$) (*i.e.* $(\alpha, \beta) > (1, 1)$ pour l'ordre lexicographique)
- $x \mapsto \dfrac{1}{x^\alpha \ln(x)^\beta}$ est intégrable en 0 si et seulement si $\alpha < 1$, ou ($\alpha = 1$ et $\beta > 1$).

◁ **Éléments de preuve.**

A savoir faire car classique, bien que HP. Cela doit donc être refait en cas de nécessité.

- Le cas $\alpha \neq 1$ en $+\infty$ s'obtient facilement par comparaison à des séries de Riemann.
- On peut ramener l'étude en 0 à la précédente par changement de variable $y = \frac{1}{x}$.

▷

II.2 Intégration par parties

Notation 4.2.7

Si f est une fonction définie sur I d'extrémités $a < b$ dans $\overline{\mathbb{R}}$ et si f admet des limites finies (à droite) en a et (à gauche) en b, on note

$$\Big[f(x)\Big]_a^b = \lim_{x \to b^-} f(x) - \lim_{x \to a^+} f(x).$$

Théorème 4.2.8 – Intégration par parties

Soit I un intervalle d'extrémités $a < b$ dans $\overline{\mathbb{R}}$. Soit $f, g : I \longrightarrow \mathbb{K}$. On suppose que :

(i) f et g sont de classe $\mathcal{C}^1$ sur I ;

(ii) le produit fg admet une limite finie en a^+ et en b^-.

Alors les intégrales $\displaystyle\int_a^b f(t)g'(t)\,\mathrm{d}t$ et $\displaystyle\int_a^b f'(t)g(t)\,\mathrm{d}t$ sont de même nature, et en cas de convergence,

$$\int_a^b f(t)g'(t)\,\mathrm{d}t = \Big[f(t)g(t)\Big]_a^b - \int_a^b f'(t)g(t)\,\mathrm{d}t.$$

Remarque 4.2.9

Selon les termes du programme, « pour les applications pratiques, on ne demande pas de rappeler les hypothèses de régularité », c'est à dire le caractère $\mathcal{C}^1$ de f et g.

Avertissement 4.2.10

Sous l'hypothèse d'existence des limites de fg, la convergence des intégrales est préservée, mais pas l'intégrabilité, comme le montre l'exemple déjà étudié de $\displaystyle\int_1^{+\infty} \frac{\sin(t)}{t}\,\mathrm{d}t$.

Exemple 4.2.11 – Relation remarquable de Γ

Montrer que pour tout $x > 0$, $\Gamma(x+1) = x\Gamma(x)$.
En déduire que pour tout $n \in \mathbb{N}$, $\Gamma(n) = (n-1)!$

III Études asymptotiques

On s'intéresse pour terminer à l'étude de 2 problèmes asymptotiques liés aux intégrales :

- Le comportement asymptotique d'intégrales dépendant de leur borne, en particulier le comportement de l'intégrale partielle lorsqu'on fait tendre une borne vers la borne de I, lorsque l'intégrale diverge, et de même, le comportement de l'intégrale reste en cas de convergence. On retrouvera ce type d'études pour les séries plus tard.
- L'étude de la convergence d'une suite d'intégrales $\displaystyle\int_I f_n$.

III.1 Intégration des relations de comparaison

Dans ce paragraphe, on suppose que $I = [a, b[$. Les résultats s'adaptent facilement au cas $]a, b]$.

Théorème 4.3.1 – Intégration des relations de comparaison dans le cas convergent

Soit $f, g : I \to \mathbb{R}$. On suppose que :

- f et g sont continues par morceaux
- g est positive (ou au moins de signe constant) et intégrable en b.

Alors :

1. si $f \underset{b}{=} O(g)$, alors $\displaystyle\int_x^b f \underset{x\to b}{=} O\left(\int_x^b g\right)$;
2. si $f \underset{b}{=} o(g)$, alors $\displaystyle\int_x^b f \underset{x\to b}{=} o\left(\int_x^b g\right)$;
3. si $f \underset{b}{\sim} g$, alors $\displaystyle\int_x^b f \underset{x\to b}{\sim} \int_x^b g$

Exemples 4.3.2

1. Soit $x > 0$. Trouver un équivalent simple lorsque y tend vers 0 de $\displaystyle\int_0^y t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t$.
2. Soit $x \in \mathbb{R}$. À l'aide d'une IPP, trouver un équivalent simple lorsque y tend vers $+\infty$ de $\displaystyle\int_y^{+\infty} t^{x-1}\mathrm{e}^{-t}$.

Dans le cas divergent, ce sont les intégrales partielles qu'on pourra comparer.

Théorème 4.3.3 – Intégration des relations de comparaison dans le cas divergent

Soit $f, g : I \to \mathbb{R}$. On suppose que :

- f et g sont continues par morceaux
- g est positive et non intégrable en b.

Alors :

1. si $f \underset{b}{=} O(g)$, alors $\displaystyle\int_a^x f \underset{x\to b}{=} O\left(\int_a^x g\right)$;
2. si $f \underset{b}{=} o(g)$, alors $\displaystyle\int_a^x f \underset{x\to b}{=} o\left(\int_a^x g\right)$;
3. si $f \underset{b}{\sim} g$, alors $\displaystyle\int_a^x f \underset{x\to b}{\sim} \int_a^x g$

Exemples 4.3.4

1. Soit $x < 0$. Déterminer un équivalent en 0 de $y \mapsto \displaystyle\int_y^{+\infty} t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t$.

2. À l'aide d'une IPP, donner un équivalent simple de $\displaystyle\int_2^x \frac{\mathrm{d}t}{\ln(t)}$ quand x tend vers $+\infty$.

III.2 Théorème de convergence dominée

Dans cette section, on donne, sans preuve, un outil extrêmement puissant pour étudier la limite d'une suite d'intégrales. On verra plus tard un autre résultat permettant de « passer à la limite sous le signe $\displaystyle\int$ », lorsque la convergence des f_n est uniforme.

Théorème 4.3.5 – Théorème de convergence dominée, TCD, admis

Soit (f_n) une suite de fonctions de I dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On suppose que :

(i) pour tout $n \in \mathbb{N}$, f_n est c.p.m.

(ii) (f_n) converge simplement vers une fonction f

(iii) (hypothèse de domination) il existe $\varphi : I \to \mathbb{K}$, intégrable sur I et telle que, sur l'intervalle I, on ait :

$$\forall n \in \mathbb{N},\ |f_n| \leqslant \varphi.$$

Alors les (f_n) et f sont intégrables sur I et

$$\int_I f_n \underset{n\to+\infty}{\longrightarrow} \int_I f.$$

◁ **Éléments de preuve.**

L'intégrabilité des f_n et de f provient des théorèmes de comparaison. Le reste de la preuve est (vraiment) hors-programme, le bon cadre n'étant par celui de l'intégrale de Riemann, mais celui de l'intégrale de Lebesgue. Ce théorème sera donc admis. ▷

Exemples 4.3.6

1. Déterminer $\displaystyle\lim_{n\to+\infty} \int_0^{\frac{\pi}{4}} \tan(x)^n \,\mathrm{d}x$
2. Déterminer $\displaystyle\lim_{n\to+\infty} \int_0^{n} \left(1 - \frac{x}{n}\right)^n \,\mathrm{d}x$.
3. Déterminer $\displaystyle\lim_{n\to+\infty} \int_0^{1} (x+n)\mathrm{e}^{-\sqrt{n}x} \,\mathrm{d}x$

Remarques 4.3.7

1. Lorsque la suite (f_n) est décroissante ce qui est le cas du premier exemple, alors on pourra toujours dominer la suite par f_0, à condition qu'elle soit intégrable.
2. Lorsque la suite est croissante, et que la limite f est intégrable, alors on pourra dominer (f_n) par f. C'est le cas du deuxième exemple.
3. Le dernier exemple montre un cas où (f_n) n'est ni croissante ni décroissante.

5

Fonctions continues sur un e.v.n.

Le but de ce chapitre est d'étendre la notion de continuité aux applications d'un e.v.n. dans un autre. Certains résultats d'analyse réelle se généralisent dans ce cadre : le théorème de Heine, le théorème des bornes atteintes (à juste titre aussi appelé théorème de compacité), et le théorème des valeurs intermédiaires, qui nécessite, pour sa généralisation, d'introduire la notion de connexité par arcs. Nous étudierons pour terminer la continuité des applications linéaires.

I Limites dans un e.v.n.

I.1 Limites finies

Définition 5.1.1 – Limite finie en un point adhérent au domaine, définition métrique

Soit $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux e.v.n., et $A \subset E$ une partie de E. Soit $f : A \longrightarrow F$, et $a \in \overline{A}$, et $b \in F$. On dit que $f(x)$ tend vers b lorsque x tend vers a si :

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \ \forall x \in A,\ \|x - a\|_E \leqslant \eta \Longrightarrow \|f(x) - b\|_F \leqslant \varepsilon.$$

Remarque 5.1.2

La signification de cette définition :

- $\forall \varepsilon > 0$: « Quelle que soit la marge d'erreur ε qu'on se donne, aussi petite soit-elle ... »
- $\exists \eta > 0$: « ... il existe une petite boule de rayon η centrée en a, quitte à prendre η très petit ... »
- $\forall x \in A,\ \|x - a\|_E < \eta \Longrightarrow$ ... : « ... tel que si x est à la fois dans A et dans cette boule ... »
- ... $\Longrightarrow \|f(x) - b\|_F < \varepsilon$: « alors $f(x)$ est proche à ε près de b. »

Autrement dit : « À condition de prendre $x \in A$ suffisamment proche de a, on peut rendre $f(x)$ arbitrairement proche de b ».

Remarques 5.1.3

1. L'hypothèse $a \in \overline{A}$ est nécessaire pour pouvoir considérer des points aussi proches qu'on veut de a.
2. L'inégalité à obtenir est d'autant plus contraignante que ε est petit. On peut se contenter d'étudier le cas de valeurs de ε inférieures à une valeur ε_0 donnée.
3. On peut généraliser cette définition à des fonctions définies entre deux espaces métriques. La continuité en a s'exprime alors de la manière suivante :

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in A,\ d_E(x, a) < \eta \Longrightarrow d_F(f(x), b) < \varepsilon,$$

où d_E et d_f sont les distances sur E et F respectivement.

Proposition 5.1.4

On peut remplacer une ou plusieurs des inégalités larges $\|x-a\|_E \leqslant \eta$ et $\|f(x)-b\|_F \leqslant \varepsilon$ par des inégalités strictes, cela donne une définition équivalente

$\lhd$ **Éléments de preuve.**
Le quantificateur universel sur ε nous assure que l'on peut aussi remplacer ε par $\varepsilon/2$. On déduit alors l'équivalence de la chaîne d'inclusions :

$$\overline{B}(b, \frac{\varepsilon}{2}) \subset B(b, \varepsilon) \subset \overline{B}(b, \varepsilon)$$

et des inclusions similaires avec a et $\frac{\eta}{2}$ (il sera alors peut-être nécessaire de considérer comme valeur de sortie $\frac{\eta}{2}$ et non η, ce qui nous donne aussi la validité de notre quantification existentielle). $\rhd$

Proposition 5.1.5 – Limite en un point du domaine

Si $a \in A$, et si $f(x)$ admet une limite en a, alors cette limite est nécessairement égale à $f(a)$.

$\lhd$ **Éléments de preuve.**
Soit ℓ est la limite de f en a. Puisque a vérifie toujours $|a-a| \leqslant \eta$, on doit avoir, pour tout $\varepsilon > 0$, $|f(a)-\ell| \leqslant \varepsilon$. $\rhd$

Avec les notations de la définition, soit B une partie de E. On note

$$f(B) = f(A \cap B) = \{f(x),\ x \in A \cap B\} = \{y \in F \mid \exists x \in B,\ f(x) = y\}$$

Il s'agit donc de l'image directe de B par f (étendu au cas où B n'est pas inclus dans le domaine de f). La définition de la limite peut alors se réécrire de façon plus synthétique.

Proposition 5.1.6 – Réexpression de la définition métrique de la limite

Soit $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux e.v.n., et $A \subset E$ une partie de X. Soit $f : A \longrightarrow F$, et $a \in \overline{A}$, et $b \in F$. Les propositions suivantes sont équivalentes:

(i) $f(x)$ tend vers b lorsque x tend vers a ;

(ii) $\forall \varepsilon > 0,\ \exists \eta > 0,\quad f(B(a,\eta)) \subset B(f(a), \varepsilon)$.

I.2 Limites infinies lorsque $F = \mathbb{R}$

Lorsque $F = \mathbb{R}$, on peut, comme dans le cas des fontions de $\mathbb{R}$ dans $\mathbb{R}$, considérer des limites infinies.

Définition 5.1.7 – Limite infinie en un point adhérent au domaine, définition métrique

Soit $(E, \|\cdot\|_E)$ un e.v.n., et $A \subset E$ une partie de E. Soit $f : A \longrightarrow \mathbb{R}$, et $a \in \overline{A}$. On dit que :

(i) $f(x)$ tend vers $+\infty$ lorsque x tend vers a si :

$$\forall M \in \mathbb{R},\ \exists \eta > 0,\quad \forall x \in A,\ \|x-a\|_E \leqslant \eta \Longrightarrow f(x) \geqslant M;$$

(ii) $f(x)$ tend vers $-\infty$ lorsque x tend vers a si :

$$\forall M \in \mathbb{R},\ \exists \eta > 0,\quad \forall x \in A,\ \|x-a\|_E \leqslant \eta \Longrightarrow f(x) \leqslant M;$$

Exemple 5.1.8

On considère dans cet exemple (et ceux qui suivent) $\mathbb{C}$ comme $\mathbb{R}$-e.v.n., pour la norme définie par le module. Étudier la limite en z_0 de $f : \mathbb{C}\backslash\{z_0\} \to \mathbb{R}$ définie par $f(z) = \frac{1}{|1-z_0|}$.

Remarque 5.1.9

On peut réexprimer la définition de la sorte : $f(x) \underset{x\to a}{\longrightarrow} +\infty$ si et seulement si

$$\forall M \in \mathbb{R}, \quad \exists \eta > 0, \quad f(B(x,\varepsilon)) \subset [M, +\infty[.$$

I.3 Limite en l'infini

Lorsque $E = \mathbb{R}$, et $+\infty$ est adhérent à A (donc A non majoré), on peut définir, comme dans le cas de fonctions réelles, des limites en $+\infty$ ou en $-\infty$

Définition 5.1.10 – Limites en $\pm\infty$ lorsque $E = \mathbb{R}$

Soit $(F, \|\cdot\|_F)$ un e.v.n., et $A \subset \mathbb{R}$ une partie non majorée de $\mathbb{R}$. Soit $f : A \longrightarrow F$, et $a \in \overline{A}$. On dit que $f(x)$ tend vers $b \in F$ lorsque x tend vers $+\infty$ si :

$$\forall \varepsilon > 0,\ \exists \lambda > 0, \quad \forall x \in A,\ x \geqslant \lambda \Longrightarrow \|f(x) - b\|_F < \varepsilon;$$

ou, de façon équivalente, si :

$$\forall \varepsilon > 0,\ \exists \lambda > 0, \quad f([\lambda, +\infty[) \subset B(b, \varepsilon).$$

Exemple 5.1.11

Étudier la limite dans $(\mathcal{C}^0([0,1]), \|\cdot\|_\infty)$ de la fonction $x \mapsto f_x$ lorsque x tend vers $+\infty$, où f_x est la fonction définie par

$$f_x(t) = (1-t)t^x.$$

Remarque 5.1.12

Le cas de la convergence de suites à valeurs dans un e.v.n. $(F, \|\cdot\|_F)$ est un cas particulier de cette définition, lorsque $A = \mathbb{N} \subset \mathbb{R}$.

On peut adapter cette définition dans le cas où $(E, \|\cdot\|_E)$ est un e.v.n. quelconque, mais au lieu de faire tendre x vers un infini dans une direction précise, on fait tendre x vers un infini non directionnel, c'est-à-dire qu'on fait tendre $\|x\|$ vers $+\infty$.

Définition 5.1.13 – Limite lorsque $\|x\| \to +\infty$

Soit $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux e.v.n., et $A \subset E$ une partie de E. Soit $f : A \to F$. On suppose que A n'est pas borné. On dit que $f(x)$ tend vers $b \in F$ lorsque $\|x\|$ tend vers $+\infty$ si

$$\forall \varepsilon > 0,\ \exists \lambda > 0, \quad \forall x \in A,\ \|x\|_E > \lambda \Longrightarrow \|f(x) - b\|_F \leqslant \varepsilon.$$

On note parfois $\lim_{x\to\infty} f(x) = b$.

Exemple 5.1.14

Soit $f : \mathbb{C} \to \mathbb{C}$ définie par $f(z) = \frac{1}{1+z^2}$. Montrer que $f(z)$ tend vers 0 lorsque $|z| \to +\infty$.

Remarque 5.1.15

Toutes les définitions données dans le cadre des fonctions réelles (notamment les limites infinies ou les limites en des points infinis) ont été données, dans le cadre du programme de première année, dans le cas où f est définie sur un intervalle (ou une union finie d'intervalles) et où a est un point de cet intervalle ou une extrémité. Cette contrainte avait pour but de se limiter à des points adhérents au domaine, mais sans le dire. Les définitions s'étendent sans peine au cas d'un point $a \in \overline{X}$, éventuellement infini (on rappelle que $+\infty \in \overline{X}$ si et seulement si X n'est pas majoré).

Exemple 5.1.16

Soit f définie sur $\mathbb{Q}$ par $f(x) = \frac{1}{q}$ où $x = \frac{p}{q}$ est la représentation irréductible de x (le signe éventuel se reportant sur p). Soit $a \in \mathbb{R}\backslash\mathbb{Q}$. Étudier la limite de $f(x)$ lorsque $x \to a$.

I.4 Limites : point de vue topologique

On peut donner une définition globale à l'aide de la notion de voisinage. On rappelle qu'on a étendu, dans le cas réel, la notion de voisinage au cas des infinis :

Définition 5.1.17 – Voisinage de $+\infty$

Soit $E = \mathbb{R}$.

- On convient d'appeler voisinage de $+\infty$ un sous-ensemble $V \subset \mathbb{R}$ tel qu'il existe $\alpha \in \mathbb{R}$ tel que $]\alpha, +\infty[\subset V$.
- On note $\mathcal{V}(+\infty)$ l'ensemble des voisinages de $+\infty$.
- On définit de même $\mathcal{V}(-\infty)$.

Définition 5.1.18 – Voisinage de ∞

Soit $(E, \|\cdot\|_E)$ un e.v.n..

- On convient d'appeler voisinage de ∞ un sous-ensemble $V \subset E$ tel qu'il existe $\alpha \in \mathbb{R}$ tel que $\complement_E \overline{B}(0, a) \subset V$.
- On note $\mathcal{V}(\infty)$ l'ensemble des voisinages de l'infini.

Remarque 5.1.19

Toute comme dans le cas de voisinages de points finis, l'intersection d'un nombre fini de voisinages de l'infini est encore un voisinage de l'infini.

La notion de voisinage permet de condenser tous les cas distincts pour la définition des limites en un seul. C'est surtout intéressant dans le cas réel où il faut distinguer les infinis au départ et à la source.
Dans la caractérisation suivante, les cas a et/ou $b = \pm\infty$ sont pertinents lorsque A et/ou B est une partie de $\mathbb{R}$ non majorée (resp. non minoré). Le cas $a = \infty$ est pertinent lorsque A n'est pas borné.

Théorème 5.1.20 – Caractérisation topologique des limites

Soit $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux e.v.n., $A \subset E$ et $f : A \longrightarrow F$. Soit $a \in \overline{A}$, ou $a \in \{\infty, \pm\infty\}$, et $b \in F$, ou $b = \pm\infty$. Les propositions suivantes sont équivalentes:

(i) $f(x)$ tend vers b lorsque x tend vers a

(ii) $\forall W \in \mathcal{V}(b), \quad \exists U \in \mathcal{V}(a), \quad f(U) \subset V$.

◁ **Éléments de preuve.**

Remarquer que $\|x - a\|_E < \varepsilon$ équivaut à $x \in B(a, \varepsilon)$

Au départ, on passe alors du cas topologique au cas métrique en remarquant que tout voisinage de a contient une boule $B(a,\varepsilon)$, et du cas métrique au cas topologique en remarquant qu'une boule $B(a,\varepsilon)$ est un voisinage de a. Quelle différence à l'arrivée ? ▻

Remarque 5.1.21

On peut aussi avoir des caratérisations mixtes, que j'énonce uniquement dans le cas fini, pour éviter d'avoir à donner trop de cas. Avec les notations de la définition, les propriétés suivantes sont équivalentes :

(i) $f(x)$ tend vers b lorsque x tend vers a ;

(ii) (métrique/topologique) $\forall W \in \mathcal{V}(b), \quad \exists \eta > 0, \quad \forall x \in A, \quad \|x-a\|_E < \eta \Longrightarrow f(x) \in W$ (*i.e.* $f(B(a,\eta) \subset W)$)

(iii) (topologique/métrique) $\forall \varepsilon > 0, \quad \exists V \in \mathcal{V}(a), \quad \forall x \in V \cap A, \quad \|f(x)-b\|_F < \varepsilon$ (*i.e.* $f(V) \subset B(a,\eta)$).

Théorème 5.1.22 – Unicité de la limite

La limite de f en a, si elle existe, est unique. On la note $\lim\limits_{x\to a} f(x)$.

◅ **Éléments de preuve.**

Même principe que pour les limites de suites : si $b_1 \neq b_2$ (y compris infini si on est dans $\overline{\mathbb{R}}$), on peut trouver $V_1 \in \mathcal{V}(b_1)$ et $V_2 \in \mathcal{V}(b_2)$ tels que $V_1 \cap V_2 = \varnothing$. ▻

Proposition 5.1.23

Soit $f : A \to F$ admettant une limite finie en a. Alors $\lim\limits_{x\to a} f(x) \in \overline{f(A)}$.

En notant par abus $x \to \infty$ pour désigner $\|x\| \to +\infty$, et $x_n \to \infty$ lorsque $\|x_n\|_E \to +\infty$, on peut généraliser la caractérisation séquentielle que vous connaissez dans le contexte réel.

Théorème 5.1.24 – Caractérisation séquentielle

Soit $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux e.v.n., $A \subset E$ et $a \in \overline{A}$ (éventuellement $a = \infty$, ou $\pm\infty$), et $b \in F$ ou $b \in \overline{\mathbb{R}}$ (si $F = \mathbb{R}$). Les propositions suivantes sont équivalentes:

(i) $\lim\limits_{x\to a} f(x) = b$

(ii) $\forall (x_n) \in A^{\mathbb{N}}, \quad (x_n \to a) \Longrightarrow (f(x_n) \to b)$.

◅ **Éléments de preuve.**

- Dans le sens direct, utiliser la caractérisation topologique pour ne pas avoir besoin de distinguer les cas.
- Dans le sens réciproque, raisonner par contraposée, et construire des voisinages collant de plus en plus à a, et dont l'image n'est pas incluse dans un certain voisinage bien choisi de b.

▻

Nous dirons qu'une fonction f admet une propriété $\mathcal{P}$ au voisinage d'un point $a \in \overline{A}$, s'il existe un voisinage V de a dans $\mathbb{R}$ (au sens étendu ci-dessus pour a infini) tel que la propriété $\mathcal{P}$ soit vérifiée par f sur l'ensemble $V \cap X$.

La notion de limite permet alors de « contrôler » une fonction au voisinage d'un point. Ainsi, on obtient par exemple :

Proposition 5.1.25

Soit f une fonction admettant une limite finie en un point a de $\overline{A}$. Alors, f est bornée au voisinage de a.

◁ **Éléments de preuve.**
La définition de la limite par ε donne un encadrement local de f. Comment s'arranger pour ne pas avoir besoin de distinguer les cas de limites en un point fini ou en un point infini ? ▷

Définition 5.1.26 – Coïncidence de deux fonctions

Soit f et g deux fonctions définies sur X et Y et a tel que $a \in \overline{X}$ et $a \in \overline{Y}$. On dit que f et g coincident au voisinage de a si et seulement s'il existe un voisinage V de a dans $\mathbb{R}$ tel que $X \cap V = Y \cap V$ et que

$$\forall x \in X \cap V, \quad f(x) = g(x).$$

Proposition 5.1.27 – Comparaison des limites de deux fonctions coïncidant au vois. de a

Soit f et g deux fonctions coïncidant au voisinage d'un point a. Alors, si f admet une limite (finie ou infinie) en a, alors g aussi, et

$$\lim_{x \to a} f(x) = \lim_{x \to a} g(x).$$

◁ **Éléments de preuve.**
Par voisinage pour ne pas avoir de discussion, en remarquant que si V_0 est un voisinage de a sur lequel f et g coïncide, alors pour tout voisinage U de a, $U \cap V_0$ est un voisinage de a, sur lequel f et g coïncident. ▷

Exemple 5.1.28

Trouver un exemple de fonctions réelles telles que f et g coïncident sur $[a, b]$, et que f admette une limite en a mais pas g.

I.5 Opérations sur les limites

La seule opération pertinente dans un e.v.n. quelconque est la combinaison linéaire.

Théorème 5.1.29 – Limite d'une CL

Soit $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux e.v.n., $A \subset E$ et $a \in \overline{A}$. Soit $f, g : A \to F$ et $\lambda \in \mathbb{K}$. Si f et g admettent une limite en a, alors $\lambda f + g$ aussi, et

$$\lim_{x \to a} (\lambda f + g)(x) = \lambda \lim_{x \to a} f(x) + \lim_{x \to a} g(x).$$

◁ **Éléments de preuve.**
Critère séquentiel. ▷

Théorème 5.1.30 – Limite d'une composée

Soit $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ et $(G, \|\cdot\|_G)$ trois e.v.n., $A \subset E$ et $B \subset F$, et $f : A \to F$ et $g : B \to G$ tels que $f(A) \subset B$. Soit $a \in \overline{A}$, tel que f admette une limite b en a et g admette une limite en b. Alors $g \circ f$ admet une limite en c, et

$$\lim_{x \to a} g \circ f(x) = \lim_{y \to b} g(y) = \lim_{y \to \lim_{x \to a} f(x)} g(y).$$

Cet énoncé reste vrai dans les cas où a et b peuvent prendre des valeurs $+\infty$, $-\infty$ ou ∞.

◁ **Éléments de preuve.**
Par voisinages, ce qui permet d'englober aussi les cas évoqués tout à la fin. ▷

Remarque 5.1.31

Le théorème de convergence des suites extraites est un cas particulier de ce résultat.

Proposition 5.1.32 – Limite d'une fonction à valeurs dans un produit cartésien

Soit $(E, \|\cdot\|_E)$ un e.v.n., et $(F_1, \|\cdot\|_{F_1}), \dots (F_p, \|\cdot\|_{F_p})$ une famille finie d'e.v.n.. Soit $f : A \to F_1 \times \cdots \times F_p$ une application définie sur une partie de A. On note, pour $x \in A$,

$$f(x) = (f_1(x), \dots, f_p(x)).$$

Soit $a \in A$ (ou éventuellement $a = \pm\infty$ ou $a = \infty$). Les propositions suivantes sont équivalentes:

(i) $f(x) \underset{x \to a}{\longrightarrow} b = (b_1, \dots, b_p)$ (pour la norme produit)

(ii) $\forall k \in [\![1, p]\!]$, $f_k(x) \underset{x \to a}{\longrightarrow} b_k$

En termes plus concis, l'existence d'une limite d'une fonction à valeurs dans un produit cartésien équivaut à l'existence de limites de chacune de ses composantes.

Remarque 5.1.33

La norme produit est la norme $\|\cdot\|_\infty$ définie sur le produit cartésien. Comme on l'a vu, elle est équivalente aux normes $\|\cdot\|_1$ et $\|\cdot\|_2$, donc cette propriété reste vraie avec ces normes-là.

Exemple 5.1.34

Le cas des limites des fonctions $f : \mathbb{R} \to \mathbb{C}$ est un cas particulier, la norme usuelle sur $\mathbb{C}$ correspondant à la norme $\|\cdot\|_2$ sur le produit cartésien $\mathbb{R} \times \mathbb{R}$. Ainsi, f admet une limite en a si et seulement si sa partie réelle et sa partie imaginaire admettent une limite en a.

I.6 Limites partielles

On considère souvent des limites partielles (sur une partie du domaine), notamment dans le cas où une fonction f est définie de façon différente sur diverses parties de son domaine de définition.

Définition 5.1.35 – Limite de f selon X

Soit E et F deux e.v.n. et $A, X \subset E$ deux parties de E. Soit $f : A \to F$. Soit $a \in \overline{A \cap X}$. La limite de f selon X est, si elle existe,

$$\lim_{\substack{x \to a \\ x \in X}} f(x) = \lim_{x \to a} f_{|X}.$$

L'hypothèse $a \in \overline{A \cap X}$ est nécessaire du fait que $A \cap X$ est le domaine de définition de la restriction $f_{|X}$.

Exemples 5.1.36

1. Pour f une fonction d'un sous-ensemble de $\mathbb{R}$ dans $\mathbb{R}$, la limite à gauche de f en a est la limite selon $]-\infty, a[$.
2. De même, la limite à droite de f en a est la limite selon $]a, +\infty[$.
3. Si $X = \{x_0\}$, a est adhérent à $X \cap A$ si et seulement si $x_0 = a$, et $a \in A$. La limite selon $\{a\}$ est alors $f(a)$.

4. Limite épointée $\lim\limits_{\substack{x \to a \\ x \neq a}} f(x)$.

Vous savez caractériser l'existence d'une limite en un point $a \in \mathbb{R}$ par l'existence et l'égalité des limites à gauche, à droite et de la valeur au point ; c'est-à-dire l'existence et l'égalité des limites selon $]-\infty, a[$, $]a, +\infty$ et $\{a\}$. Voici une généralisation de cette caractérisation

Proposition 5.1.37 – Caractérisation de la limite par les limites partielles

Soit E, F deux e.v.n., $A \subset E$, et $a \in \overline{A}$. Soit $f : A \to F$. Soit $b \in F$. Soit $(X_i)_{i \in I}$ une famille *finie* de parties de E tels que :

(i) il existe un voisinage $V \in \mathcal{V}(a)$ tel que $V \subset \bigcup\limits_{i \in I} X_i$;

(ii) pour tout $i \in I$, $a \in \overline{A \cap X_i}$;

(iii) pour tout $i \in I$, $\lim\limits_{\substack{x \to a \\ x \in X_i}} f(x) = b$.

Alors $\lim\limits_{x \to a} f(x) = b$.

Exemple 5.1.38

Étudier la limite en $(x_0, 0)$ de la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x,y) = \begin{cases} \frac{\sin(xy)}{\sqrt{y}} & \text{si } y > 0 \\ 0 & \text{sinon.} \end{cases}$$

Remarque 5.1.39

Les propriétés sur les limites (et celles sur la continuité qui suivent) sont notamment valides dans le contexte de fonctions $f : A \to \mathbb{R}$, où A est une partie de $\mathbb{R}^n$ (fonction réelle de n variables réelles). Il faut a priori se donner une norme sur $\mathbb{R}^n$, mais comme on le verra à la fin du chapitre, toutes les normes sur $\mathbb{R}^n$ sont équivalentes, et le choix de la norme importe peu. On peut ainsi choisir la norme qui convient le mieux à l'étude en cours.
Vous pouvez admettre ce point, et utiliser dès à présent dans les exercices l'équivalence des normes sur $\mathbb{R}^n$.

II Continuité

II.1 Continuité en un point

Dans tout le paragraphe, $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ sont deux e.v.n., et $A \subset E$ est une partie de E.

Définition 5.2.1 – Continuité

Soit $f : A \to F$ et $a \in A$. On dit que f est continue en a si f admet une limite (dans F) en a.

Remarque 5.2.2

Le point a est dans le domaine de f. Ainsi, si f est continue en a, la limite en a est nécessairement $f(a)$.

De façon équivalente :

Proposition 5.2.3 – Réexpression métrique et topologique

Soit $f : A \to F$, et $a \in A$. Les propriétés suivantes sont équivalentes :

(i) f est continue en a ;

(ii) $\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in A,\quad \|x-a\|_E \leqslant \eta \Longrightarrow \|f(x)-f(a)\|_F \leqslant \varepsilon$;

(iii) pour tout $W \in \mathcal{V}(f(a))$, il existe un voisinage $V \in \mathcal{V}(a)$ tel que $f(V) \subset W$.

◁ **Éléments de preuve.**
Équivalences déjà vues dans le cadre des limites. Le seul ingrédient supplémentaire est l'égalité de la limite avec $f(a)$, ce qui, comme on l'a vu, est toujours le cas lorsque a est dans le domaine de f. ▷

Remarque 5.2.4

Comme on l'a vu dans le cadre des limites, les inégalités $\|x-a\|_E \leqslant \eta$ et $\|f(x)-f(a)\|_F \leqslant \varepsilon$ peuvent être indifféremment larges ou strictes.

Définition 5.2.5 – Continuité sur une partie

Soit $f : A \to F$, et $B \subset A$. On dit que f est continue sur B si f est continue en tout point de B.

Avertissement 5.2.6

Cela n'équivaut pas à la continuité de $f_{|B}$. Pouvez-vous trouver un contre-exemple ?

Proposition 5.2.7 – Restriction de l'ensemble source

Si f et g coïncident sur un *voisinage* de a, f est continue en a si et seulement si g est continue en a.

◁ **Éléments de preuve.**
On a déjà vu dans cette situation l'équivalence de l'existence des limites, et leur égalité. ▷

Corollaire 5.2.8

Soit f et g définies sur $A \subset E$, et U un ouvert tel que $U \cap A \neq \varnothing$. Alors, si $f_{|U\cap A} = g_{|U\cap A}$ et si g est continue sur $U \cap A$, alors f aussi.

Exemple 5.2.9

Étude de la continuité de la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x,y) = \begin{cases} \frac{\sin(xy)}{\sqrt{y}} & \text{si } y > 0 \\ 0 & \text{sinon.} \end{cases}$$

Enfin, on traduit dans le cadre de la continuité la caractérisation séquentielle de la limite.

Proposition 5.2.10 – Caractérisation séquentielle de la continuité

Soit $f : A \to F$ une application définie sur une partie A d'un e.v.n. E, et $a \in A$. Les propositions suivantes sont équivalentes:

(i) f est continue en a ;

(ii) pour tout $(u_n) \in A^{\mathbb{N}}$, $(u_n \to a) \Longrightarrow (f(u_n) \Longrightarrow f(a))$.

◁ **Éléments de preuve.**
Conséquence directe de la caractérisation séquentielle pour les limites, et de la valeur de la limite lorsque a est dans le domaine. ▷

II.2 Opérations sur les fonctions continues

Les opérations sur les limites ont pour corollaire immédiat les opérations similaires sur les fonctions continues.
On considère toujours $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux e.v.n., et A une partie de E, et $a \in A$.

Proposition 5.2.11 – Continuité d'une combinaison linéaire

Soit $f, g : A \to F$, et $\lambda : A \to \mathbb{K}$. Si f, g et λ sont continues en a, alors $\lambda f + g$ aussi.

Proposition 5.2.12 – Continuité d'une composition

Soit $A \subset F$, et $B \subset F$. Soit $f : A \to F$ et $g : B \to G$ tels que $f(A) \subset B$

- Si f est continue en $a \in A$ et g continue en $f(a)$, alors $g \circ f$ est continue en a.
- En particulier, si f est continue sur A et g continue sur B, alors $g \circ f$ est continue sur A.

Proposition 5.2.13 – Continuité d'une fonction à valeurs dans un produit cartésien

On suppose que $F = F_1 \times \cdots \times F_p$ où $(F_k, \|\cdot\|_{F_k})$ sont des e.v.n., et que $\|\cdot\|_F$ est la norme produit $\|\cdot\|_\infty$. Soit $f = (f_1, \ldots, f_p) : A \to F$. Les propositions suivantes sont équivalentes:

(i) f est continue en a ;

(ii) pour tout $k \in [\![1, p]\!]$, f_k est continue en a.

Remarque 5.2.14

1. Cela reste valide en remplaçant la norme produit par $\|\cdot\|_1$ ou $\|\cdot\|_2$ qui lui sont équivalentes.
2. En particulier, cela permet d'étudier la continuité d'une application à valeurs dans $\mathbb{R}^n$ en étudiant chacun de ses coordonnées, si $\mathbb{R}^n$ est muni d'une des 3 normes $\|\cdot\|_1$, $\|\cdot\|_2$ ou $\|\cdot\|_\infty$, ou plus généralement de n'importe quelle norme, en admettant provisoirement l'équivalence des normes en dimension finie.

II.3 Continuité uniforme, fonctions lipschitziennes

La notion de continuité traduit le fait que localement, $f(x)$ approche $f(a)$ à ε près fixé arbitrairement à l'avance. L'aspect local de cette approximation se traduit par ε intervenant dans la définition : ε nous donne le domaine de validité de l'approximation. Plus ε est petit, plus il faut rester proche de x pour que l'approximation soit correcte. La taille de ce domaine de validité peut d'ailleurs dépendre de x. En général, il n'y a pas de raison de pouvoir trouver un réel η convenant pour toutes les valeurs de x du domaine de définition d'une fonction continue. Cela signifie que, une marge ε étant donnée, il peut exister des points pour lesquels il faudra rester très très proche (infiniment proche si on fait tendre x vers un des bords du domaine) de x pour que l'approximation à ε près reste vraie.

Exemples 5.2.15

1. $x \mapsto e^x$ sur $\mathbb{R}$.
2. $x \mapsto \frac{1}{x}$ sur $]0, 1]$.

Si pour tout ε, on peut trouver un η convenable pour toute valeur de x, on parlera de continuité uniforme. Remarquez qu'il ne s'agit de rien de moins que d'une interversion de quantificateurs par rapport à la définition de la continuité sur un domaine.

Définition 5.2.16 – Continuité uniforme

Soit $f : A \to F$ une fonction définie sur une partie A d'un e.v.n. E. Soit $B \subset A$. On dit que f est uniformément continue sur B si

$$\forall \varepsilon > 0,\quad \exists \eta > 0,\quad \forall (x,y) \in B^2,\quad \|x-y\|_E < \eta \Longrightarrow \|f(x)-f(y)\| < \varepsilon.$$

Le réel η est appelé *module de continuité uniforme de f pour l'approximation ε*

Proposition 5.2.17 – Continuité uniforme implique continuité

Soit $f : A \to F$ et $B \subset A$. Si f est uniformément continue sur B, elle est continue sur B.

◁ **Éléments de preuve.**

Évident : c'est juste enlever une contrainte de dépendance. ▷

Proposition 5.2.18 – Critère séquentiel de la continuité uniforme

Soit $f : A \to F$ une fonction définie sur une partie A d'un e.v.n. E. Soit $B \subset A$. Les deux propriétés suivantes sont équivalentes :

(i) f est uniformément continue sur B

(ii) Pour toutes suites (x_n) et (y_n) d'éléments de B telles que $x_n - y_n \to 0$, on a aussi $f(x_n) - f(y_n) \to 0$.

◁ **Éléments de preuve.**

Le sens direct est une conséquence immédiate de la définition de la continuité uniforme et de la convergence de suites. Réciproquement, raisonner par la contraposée. ▷

Un exemple important de classe de fonctions uniformément continues est la classe des fonctions lipschitziennes.

Définition 5.2.19 – Applications lipschitziennes

Soit $f : A \to F$ une application définie sur une partie A d'un e.v.n. E. Soit $B \subset A$. Soit $k \in \mathbb{R}_+$.

- On dit que f est k-lipschitzienne sur B si

$$\forall (x,y) \in B^2,\quad \|f(x)-f(y)\|_E \leqslant k\|x-y\|_F.$$

- On dit dans ce cas que k est un facteur de Lipschitz pour la fonction f sur B.
- On dit que f est lipschitzienne sur B s'il existe $k \in \mathbb{R}_+$ tel que f soit k-lipschitzienne.
- On dit que f est contractante s'il existe $k \in [0,1[$ tel que f soit k-lipschitzienne.

Proposition 5.2.20 – Lipschitz implique u.c.

Une application lipschitzienne sur B est uniformément continue sur B.

II.4 Exemples importants de fonctions continues

Proposition 5.2.21 – Continuité de la norme

Soit $(E, \|\cdot\|_E)$ un e.v.n.. L'application $x \mapsto \|x\|_E$ de E dans $\mathbb{K}$ est 1-lipschitzienne, donc uniformément continue sur E.

◁ **Éléments de preuve.**

Par inégalité triangulaire. ▷

Proposition 5.2.22 – Continuité de la distance à une partie

Soit $(E, \|\cdot\|_E)$, et $A \subset E$ une partie non vide. L'application $d(\cdot, A)$ est 1-lipschitzienne donc uniformément continue sur E.

◁ **Éléments de preuve.**

Majorer $d(x, z) - d(y, z)$ par IT, puis $d(x, A) - d(y, A)$, puis intervertir les rôles. ▷

Proposition 5.2.23 – Continuité de la projection

Soit $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux e.v.n.. On munit $E \times F$ de la norme produit. Alors les projections $p_E : E \times F \to E$ et $p_F : E \times F \to F$ définies par

$$p_E(x, y) = x \qquad \text{et} \qquad p_F(x, y) = y$$

sont uniformément continues sur $E \times F$.

◁ **Éléments de preuve.**

Probablement encore un coup de M. Rudolf Lipschitz. ▷

Voici un corollaire qui permet souvent de se ramener à la continuité de fonctions usuelles d'une variable.

Corollaire 5.2.24

Soit $(E_1, \|\cdot\|_{E_1})$, $(E_2, \|\cdot\|_{E_2})$ et $(F, \|\cdot\|_F)$ trois e.v.n., et $f : E_1 \to F$ une application continue. Alors

$$\tilde{f} = \left\{ \begin{array}{ccc} E_1 \times E_2 & \longrightarrow & F \\ (x, y) & \longmapsto & f(x) \end{array} \right.$$

est continue.

◁ **Éléments de preuve.**

Composer par la projection. ▷

Exemple 5.2.25

Étudier la continuité pour $\|\cdot\|_\infty$ de la fonction $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par

$$f(x, y) = \left(\sin(x)\cos(y), \mathrm{e}^{x+y}\right).$$

Un autre résultat, permettant de travailler « à isomorphisme près » (par exemple en assimilant $\|\cdot\|_{\mathcal{B},\infty}$ sur E à $\|\cdot\|_\infty$ sur l'espace $\mathbb{K}^n$ des coordonnées dans la base $\mathcal{B}$) :

Théorème 5.2.26 – Continuité de l'isomorphisme de transfert

Soit E et F deux $\mathbb{K}$-ev, et $\varphi : E \to F$ un isomorphisme. Soit $\|\cdot\|_F$ une norme sur F, et $\|\cdot\|_E$ la norme sur E induite par φ. Alors φ et φ^{-1} sont 1-lipschitziennes donc continues (pour ces normes).

◁ **Éléments de preuve.**

Ce sont même des isométries ! ▷

II.5 Prolongements

Proposition/Définition 5.2.27 – Prolongement par continuité

Soit $f : A \to F$ définie sur une partie A d'un e.v.n., et continue sur A. Soit $a \in \overline{A} \backslash A$ tel que $f(x)$ admet une limite $b \in F$ lorsque x tend vers a.
Alors il existe une unique fonction $\tilde{f} : A \cup \{a\} \to F$, coïncidant avec f sur A et continue sur $A \cup \{a\}$.
Elle est définie par :

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in A \\ \lim\limits_{y \to a} f(y) & \text{si } x = a. \end{cases}$$

La fonction $\tilde{f}$ est appelée prolongement par continuité de f en a.

Les situations de densité permettent aussi d'obtenir des propriétés d'unicité de prolongements.

Proposition 5.2.28 – Coïncidence sur une partie dense

Soit $f, g : A \to F$ deux applications continues sur A, et B une partie de E dense dans A. Si f et g coïncident sur B, alors $f = g$.

Corollaire 5.2.29

Soit A une partie de E et $B \subset A$ une partie dense dans A. Soit $f : B \to F$ une application continue. Alors f admet au plus un prolongement g défini et continu sur A.

Exemple 5.2.30

Déterminer toutes les applications continues $f : \mathbb{R} \to \mathbb{R}$ telles que pour tout $(x, y) \in \mathbb{R}^2$, $f(x + y) = f(x) + f(y)$.

III Continuité et topologie

III.1 Images réciproques d'ouverts et fermés

Comme souvent, les images réciproques donnent des propriétés plus simples que les images directes (surtout du fait d'une quantification en moins dans la définition, ce qui permet d'échapper aux problèmes d'interversion de quantificateurs).

Théorème 5.3.1 – Image réciproque d'un ouvert, d'un fermé

Soit $f : A \to F$ une application continue sur A, partie d'un e.v.n. E.

1. Pour tout U ouvert de F, $f^{-1}(U)$ est un ouvert relatif de A.
2. Pour tout C fermé de F, $f^{-1}(C)$ est un fermé relatif de A.

Ce théorème est notamment très utile pour montrer de façon efficace que certaines parties de $\mathbb{R}^n$ sont des ouverts ou des fermés.

Exemples 5.3.2

1. Le demi-plan de $\mathbb{R}^2$ d'équation $x \geqslant 0$ est un fermé.
2. $\{(x, y) \in \mathbb{R}^2 \mid x < 3y + 2 \text{ et } x^2 > y\}$ est un ouvert de $\mathbb{R}^2$.
3. Redémontrer le fait que $\overline{B}(a, r)$ et $S(a, r)$ sont des fermés de E, et que $\mathring{B}(a, r)$ est un ouvert.

Comme on l'avait annoncé plus haut :

Théorème 5.3.3 – $\mathrm{GL}_n(\mathbb{K})$ est un ouvert de $\mathcal{M}_n(\mathbb{K})$

Tout est dit dans le titre (et ceci indépendamment de la norme, si on admet l'équivalence des normes en dimension finie).

◁ **Éléments de preuve.**
Quelle fonction continue permet de caractériser l'inversibilité ? ▷

Avertissement 5.3.4

En général, l'image directe d'un ouvert (resp. fermé) par une fonction continue n'est pas ouvert (resp. fermé)

Exemple 5.3.5

Trouvez (graphiquement) des exemples dans le cas de fonctions $\mathbb{R} \to \mathbb{R}$.

Pour les fermés, c'est cependant presque vrai, avec une petite condition en plus (la compacité).

III.2 Fonction continue sur compact

Théorème 5.3.6 – Théorème de compacité

L'image par une fonction continue $f : A \to F$ d'une partie compacte $K \subset A$ est une partie compacte de F

◁ **Éléments de preuve.**
La continuité permet de préserver la convergence d'une suite extraite. ▷

Corollaire 5.3.7 – Théorème des bornes atteintes

Soit $f : K \to \mathbb{R}$ continue sur une partie K compacte d'un e.v.n. E. Alors f est bornée et atteint ses bornes.

◁ **Éléments de preuve.**
Par le théorème de compacité, $f(K)$ est fermé et borné. Donc $\sup f(K) \in f(K)$. ▷

Remarque 5.3.8

1. On retrouve l'énoncé de première année dans le cadre de fonctions définies sur un intervalle fermé borné de $\mathbb{R}$ en remarquant qu'un intervalle $[a, b]$ est compact.
2. L'importance de ce théorème se fait particulièrement sentir dans le cadre de la recherche d'extremas globaux : le théorème des bornes atteintes se réexprime en effet en disant qu'une fonction $f : K \to \mathbb{R}$ définie sur un compact admet un minimum et un maximum sur $\mathbb{K}$.

Exemple 5.3.9

Soit P une fonction polynomiale sur $\mathbb{R}^2$ prenant au moins une valeur strictement positive. Existence Montrer que l'application $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x, y) = \mathrm{e}^{-(x^2+y^2)} P(x, y)$$

admet un maximum global.

Une autre propriété des fonctions continues sur un compact généralise une propriété de première année.

Théorème 5.3.10 – Heine

Une fonction continue sur un compact K est uniformément continue sur K.

◁ **Éléments de preuve.**
De même que le cas réel, par contraposée, en niant le critère séquentiel. ▷

IV Connexité

On cherche dans cette partie à généraliser le théorème des valeurs intermédiaires.

IV.1 Parties connexes par arcs

Définition 5.4.1 – Arc, ou chemin continu

Soit $(E, \|\cdot\|_E)$ un e.v.n., et $(a,b) \in E^2$. Un arc continu (ou chemin continu) de a à b est une application continue $\gamma : [0,1] \to E$ telle que $\gamma(0) = a$ et $\gamma(1) = b$.

Remarque 5.4.2

Faites un dessin pour illustrer la notion !

Soit A une partie de E. On peut définir une relation sur A par

$$a\mathcal{R}b \iff \text{il existe un arc continu reliant } a \text{ à } b.$$

Proposition/Définition 5.4.3 – Composantes connexes de A

La relation $\mathcal{R}$ est une relation d'équivalence. Les classes d'équivalence de A pour cette relation sont appelées composantes connexes de A.

Définition 5.4.4 – Partie connexe par arcs

Soit A une partie d'un e.v.n. E. On dit que A est connexe par arcs si A ne possède qu'une seule composante connexe, donc si pour tout $(a,b) \in A^2$, il existe un arc continu γ reliant a à b.

Exemple 5.4.5

1. Déterminer les composantes connexes par arcs de $\mathbb{R}^*$
2. Montrer que $\mathbb{C}^*$ est connexe par arcs.
3. Montrer que le « plan fendu » $\mathbb{C}\backslash\mathbb{R}_-$ est connexe par arcs
4. $\mathbb{C}\backslash\mathbb{R}$ est-il connexe par arc ?
5. Étudier la connexité par arcs de $S(a,r)$ dans un e.v.n..

Définition 5.4.6 – Partie étoilée

Soit A une partie d'un e.v.n. E.

(i) On dit que A est étoilée par rapport à $a \in A$ si pour tout $b \in A$, le segment $[a,b]$ est inclus dans A, c'est-à-dire :

$$\forall t \in [0,1], \quad (1-t)a + tb \in A.$$

(ii) On dit que A est étoilée s'il existe a tel que A soit étoilée par rapport à A.

Proposition 5.4.7

Une partie convexe est étoilée.

Proposition 5.4.8 – Étoilé implique connexe

Soit A une partie d'un e.v.n. E.

- Si A est étoilée, alors A est connexe par arcs.
- En particulier, toute partie convexe est connexe par arcs.

◁ **Éléments de preuve.**
On peut joindre b à c en 2 temps, en faisant un petit détour. ▷

Exemple 5.4.9

1. Les boules sont connexes par arcs.
2. Montrer que tout sous-espace vectoriel de E est connexe par arcs.

Proposition 5.4.10 – Les connexes par arcs de $\mathbb{R}$

Les parties connexes par arcs de $\mathbb{R}$ sont exactement les intervalles.

IV.2 Image continue d'une partie connexe par arcs

Le théorème suivant est une généralisation du TVI.

Théorème 5.4.11 – Image continue d'un connexe par arcs

Soit $f : A \to F$ une fonction continue sur une partie connexe par arcs A de E. Alors $f(A)$ est aussi connexe par arcs.

Dans le cas où $E = F = \mathbb{R}$, les connexes par arcs étant les intervalles, on retrouve le fait que l'image d'un intervalle par une fonction continue est un intervalle, qui est l'un des énoncés possibles du TVI.
On peut se servir de cette propriété pour montrer que certains ensembles ne sont pas connexes par arcs.

Exemple 5.4.12

1. En considérant det, montrer que $\mathrm{GL}_n(\mathbb{R})$ n'est pas connexe par arcs.
2. De même pour $\mathrm{O}_n(\mathbb{R})$.

V Continuité des applications linéaires

On termine ce chapitre par l'étude de la continuité des applications linéaires entre deux e.v.n..

Théorème 5.5.1 – Caractérisation de la continuité d'une AL

Soit $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux e.v.n., et $u \in \mathcal{L}(E, F)$. Les propriétés suivantes sont équivalentes.

(i) u est continue sur E ;

(ii) u est continue en 0 ;

(iii) il existe $k \in \mathbb{R}_+$ tel que pour tout $x \in E$, $\|u(x)\|_E \leqslant k\|x\|_F$

(iv) $u(\overline{B}(0,1))$ est borné.

(v) $u(S(0,1))$ est borné.

Définition 5.5.2 – Norme subordonnée

Soit $u \in \mathcal{L}(E, F)$ une application linéaire continue. On définit

$$|||u||| = \sup_{x \in S(0,1))} \|u(x)\|_F = \sup_{x \in E \setminus \{0\}} \frac{\|u(x)\|_F}{\|x\|_E}$$

On dit que $|||u|||$ est la norme subordonnée (aux normes $\|\cdot\|_E$ et $\|\cdot\|_F$) de l'application linéaire continue u.
On parle aussi de norme triple, ou de norme d'opérateur, et on rencontre parfois la notation $\|\cdot\|_{\text{op}}$.

Exemples 5.5.3

1. Soit $(E, \|\cdot\|)$ un e.v.n.. Déterminer $|||\text{Id}_E|||$.
2. Soit $E = \mathbb{R}[X]$ muni de $\|\cdot\|_\infty$ sur les coefficients, et $a \in]-1, 1[$. Montrer que $u : P \mapsto P(a)$ est continue, et déterminer $|||u|||$. Que dire lorsque $|a| > 1$?

Proposition 5.5.4 – L'e.v.n. $\mathcal{L}_c(E, F)$

Soit $\mathcal{L}_c(E, F)$ l'ensemble des applications linéaires continues de E dans F. Alors $\mathcal{L}_c(E, F)$ est un $\mathbb{K}$-ev, et $|||\cdot|||$ est une norme sur $\mathcal{L}_c(E, F)$.

Proposition 5.5.5 – Sous-multiplicativité de $|||\cdot|||$

Soit E, F et G trois e.v.n., et $u \in \mathcal{L}_c(E, F)$, $v \in \mathcal{L}_c(F, G)$. Alors $v \circ u \in \mathcal{L}_c(E, G)$, et

$$|||v \circ u||| \leqslant |||v||| \cdot |||u|||.$$

Corollaire 5.5.6

La norme $|||\cdot|||$ est une norme d'algèbre sur $\mathcal{L}_c(\mathbb{R})$.

Exemple 5.5.7

Donner un exemple montrant qu'on peut avoir $|||u \circ v||| < |||u||| \cdot |||v|||$.

On généralise l'étude de la continuité des applications linéaires au cas des applications multilinéaires.

Théorème 5.5.8 – Caractérisation de la continuité des applications multilinéaires

Soit $E_1, \dots, E_n, F$ des e.v.n., de normes $\|\cdot\|_{E_1}, \dots, \|\cdot\|_{E_n}$, $\|\cdot\|_F$. Soit $\varphi : E_1 \times \cdots \times E_n \to F$ une application n-linéaire. Les propositions suivantes sont équivalentes:

(i) φ est continue ;

(ii) φ est continue en $(0, \dots, 0)$;

(iii) il existe $k \in \mathbb{R}_+$ tel que pour tout $(x_1, \dots, x_n)$,

$$\|\varphi(x_1, \dots, x_n)\|_F \leqslant k \prod_{i=1}^{n} \|x_i\|_{E_i}.$$

◁ **Éléments de preuve.**

Le démonstration n'est pas exigible.

- $(ii) \Longrightarrow (iii)$: Dans le sens direct, prendre $k = \frac{1}{\eta^n}$, où η est tel que

$$\|(x_1, \dots, x_n)\| \leqslant \eta \Longrightarrow \|\varphi(x_1, \dots, x_n))\| \leqslant 1.$$

- $(iii) \Longrightarrow (i)$ Faire un télescopage (comme pour la continuité d'un produit), de sorte à former une différence $y_i - x_i$ sur l'une des coordonnées dans chaque terme.

▷

Exemples 5.5.9

1. L'application $\mathbb{K} \times E \to E$, telle que $(\lambda, x) \mapsto \lambda x$.
2. L'application d'évalution $\mathcal{L}_c(E, F) \times E \to F$ telle que $(u, x) \mapsto u(x)$.
3. L'application $\mathcal{L}_c(E, F) \times \mathcal{L}_c(F, G) \to \mathcal{L}_c(E, G)$ telle que $(u, v) \mapsto v \circ u$.
4. Si E est un espace euclidienne, et $\| \cdot \|$ la norme euclidienne associée, $\langle \cdot, \cdot \rangle$ est continue.
5. Si E est une algèbre tel que $\| \cdot \|_E$ est une norme d'algèbre, le produit $(x, y) \mapsto xy$ de $E \times E$ dans E.
6. Retrouver avec l'exemple précédent la propriété de continuité d'un produit de deux fonctions à valeurs dans une algèbre muni d'une norme d'algèbre (donc même si cette propriété n'est pas au programme, on la retrouver facilement avec les résultats au programme).

VI E.v.n. de dimension finie, et conséquences sur la continuité

VI.1 Équivalence des normes

Nous avons déjà évoqué ce théorème d'une importance capitale :

Théorème 5.6.1 – Équivalence des normes en dimension finie

Soit E un espace vectoriel de dimension finie. Toutes les normes définies sur E sont équivalentes

◁ **Éléments de preuve.**

Fixer une base $\mathcal{B}$ et comparer une norme N quelconque à la norme $\| \cdot \|_\infty = \| \cdot \|_{\mathcal{B},\infty}$.

- Dans un sens, c'est évident par IT et majoration des coordonnées.
- Dans l'autre, utiliser la deuxième IT et la majoration précédente pour montrer que N est lipschitzienne donc continue. Terminer en utilisant la compacité de $B(0, 1)$ pour $\| \cdot \|_{\mathcal{B},\infty}$.

▷

Un espace de dimension finie pouvant toujours être muni d'une norme (par exemple la norme $\| \cdot \|_{\mathcal{B},\infty}$ associée à une base), on a une topologie naturelle d'e.v.n. unique, indépendante de la norme choisie. En particulier, toutes les propriétés de continuité et de convergence sont les mêmes quelle que soit la norme choisie. Le caractère borné est aussi préservé d'une norme à l'autre.

Il en résulte que le choix de la norme n'a pas beaucoup d'importance, et en général, elle ne sera pas préciser. Parler de convergence ou de continuité dans un espace de dimension finie E, sans préciser de norme, se réfère de façon implicite à la topologie définie par l'une quelconque des normes qui peuvent être définies sur E.

Ce théorème nous permet notamment de choisir une norme à notre convenance, dans le but de simplifier les calculs et les majorations, certaines normes pouvant être plus pratiques que d'autres suivant le contexte. Ainsi, comme on l'a déjà vu, il peut être intéressant, dans un contexte matriciel (donc en dimension finie), d'utiliser une norme matricielle (donc sous-multiplicative).

En particulier, la convergence peut toujours s'étudier avec la norme $\| \cdot \|_{\mathcal{B},\infty}$, ce qui permet de justifier le résultat suivant.

Corollaire 5.6.2 – Caractérisation de la convergence par les coordonnées en dim finie

Soit (E, N) un e.v.n. de dimension finie, et $\mathcal{B} = (b_1, \dots, b_d)$ une base de E, (E', N') un e.v.n. quelconque, $A \subset E'$ et $a \in \overline{A}$. Soit $(u_n) \in E^{\mathbb{N}}$ et $f \in E^A$ décomposées sur la base $\mathcal{B}$:

$$\forall n \in \mathbb{N}, \quad u_n = \sum_{k=1} u_{n,k} b_k \qquad \text{et} \qquad \forall x \in A, \quad f(x) = \sum_{k=1}^{d} f_k(x) b_k.$$

1. La suite $(u_n)_{n\in\mathbb{N}}$ converge vers L tel que $[L]_{\mathcal{B}} = (\ell_1, \dots, \ell_d)$ si et seulement si pour tout $k \in [\![1, d]\!]$, $(u_{n,k})_{n\in\mathbb{N}}$ converge vers ℓ_k.
2. L'application f admet une limite $L = (\ell_1, \dots, \ell_d)$ lorsque x tend vers a si et seulement si pour tout $k \in [\![1, d]\!]$, $f_k(x) \underset{x\to a}{\longrightarrow} \ell_k$.
3. L'application f est continue en $b \in A$ si et seulement si pour tout $k \in [\![1, d]\!]$, f_k est continue en b.

Autrement dit, la convergence (de suites, de fonctions) et la continuité peuvent s'étudier coordonnée par coordonnée.

Exemple 5.6.3

Trouver dans $\mathbb{R}[X]$ une suite non convergente pour la norme $\|\cdot\|_\infty$ sur les coefficients, mais convergeant coordonnée par coordonnée dans la base canonique.

Avertissement 5.6.4

On peut définir des métriques sur E associées à d'autres topologies ! Ainsi, l'unicité de la topologie sur E de dimension finie est valide dans le contexte où on ne regarde que les topologies issues de normes.

VI.2 Conséquences topologiques

Comme on l'a vu, en dimension finie, la topologie définie par une norme ne dépend pas du choix de la norme. Cela implique en particulier que certaines propriétés vues dans le cadre d'une norme donnée restent valides pour toutes les normes. En particulier :

Théorème 5.6.5 – Bolzano-Weierstrass en dimension finie

Soit $(E, \|\cdot\|)$ un e.v.n. de dimension finie. Alors toute suite bornée de vecteurs de E admet une valeur d'adhérence.

◁ **Éléments de preuve.**

Cela avait été démontré pour $\|\cdot\|_{\mathcal{B},\infty}$, après choix d'une base. Par équivalence des normes, cela reste vrai pour toute norme. ▷

Corollaire 5.6.6 – Caractérisation des compacts en dimension finie

Soit E un e.v.n. de dimension finie, et $K \subset E$. Les propositions suivantes sont équivalentes:

(i) K est compact ;

(ii) K est fermé et borné.

◁ **Éléments de preuve.**

Un sens a déjà été vu. L'autre résulte de BW et du caractère fermé. ▷

Corollaire 5.6.7 – Compacité des boules en dimension finie

Soit E un e.v.n. de dimension finie, $a \in E$ et $r \in \mathbb{R}_+^*$. La boule fermée $\overline{B}(a, r)$ est un compact.

Corollaire 5.6.8 – Caractérisation de la convergence par les v.a. dans un e.v.n. de dim finie

Soit E un e.v.n. de dimension finie, et (u_n) un suite bornée. Alors (u_n) converge si et seulement si elle a une unique valeur d'adhérence.

◁ **Éléments de preuve.**
(u_n) est à valeurs dans une boule fermée, donc compacte. ▷

Théorème 5.6.9 – Propriété topologique des sous-espaces

Soit E un e.v.n. quelconque et F un sous-espace vectoriel de dimension finie de E. Alors F est fermé.

◁ **Éléments de preuve.**
Utiliser une base de $\text{Vect}(F \cup \{\ell\})$, où ℓ est la limite d'une suite convergente d'éléments de E, et introduire une base convenable de cet espace pour étudier la convergence coordonnée par coordonnée. ▷

Exemple 5.6.10

C'est faux en dimension infinie. En effet soit E l'espace des suites réelles telles que $\sum \frac{|u_n|}{n^2}$ converge. Pour $u = (u_n)$, on définit

$$N(u) = \sum_{n=0}^{+\infty} \frac{|u_n|}{n^2}.$$

Soit F le sous-espace de E des suites à support fini (donc nulles àpcr). Montrer que N est une norme sur E, et que F n'est pas fermé (on pourra constuire une suite d'éléments de F convergeant vers la suite constante de valeur 1).

VI.3 Conséquences sur la continuité des applications linéaires

Théorème 5.6.11 – Continuité des AL en dimension finie

Soit E, F deux e.v.n.. Si E est de dimension finie, toute application linéaire $u \in \mathcal{L}(E, F)$ est continue, donc

$$\mathcal{L}_c(E, F) = \mathcal{L}(E, F).$$

◁ **Éléments de preuve.**
Munir E d'une norme $\|\cdot\|_\infty = \|\cdot\|_{\mathcal{B},\infty}$, et majorer $u(\overline{B}(0, 1))$. ▷

La norme triple $|||\cdot|||$ peut donc être définie sur $\mathcal{L}(E, F)$ tout entier.
Si F est lui-même de dimension finie, $\mathcal{L}(E, F)$ est isomorphe à un espace de matrices $\mathcal{M}_{n,p}(\mathbb{K})$. Plus précisément, $M \in \mathcal{M}_{n,p}(\mathbb{K})$ représente une application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$, définie par

$$X \mapsto MX.$$

Ainsi, quelle que soit la norme $\|\cdot\|_n$ choisie sur $\mathbb{K}^n$ et la norme $\|\cdot\|_p$ choisie sur $\mathbb{K}^p$, l'application linéaire définie par M est continue, ce qui permet de définir la norme triple associée à M

Définition 5.6.12 – Norme triple associée à une matrice

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$. La norme triple de M (coordonnée aux normes $\|\cdot\|_p$ et $\|\cdot\|_n$ de $\mathbb{K}^p$ et $\mathbb{K}^n$) est définie par

$$|||M||| = \sup_{X \in \mathbb{K}^p \setminus \{0\}} \frac{\|MX\|_n}{\|X\|_p} = \sup_{X \in \overline{B}(0,1) \setminus \{0\}} \frac{\|MX\|_n}{\|X\|_p} = \sup_{X \in S(0,1)} \|MX\|_n.$$

Proposition 5.6.13 – Encore une norme matricielle

La norme triple est matricielle (*i.e.* sous-multiplicative) : si $M \in \mathcal{M}_{n,p}(\mathbb{K})$ et $N \in \mathcal{M}_{p,q}(\mathbb{K})$, et si la norme utilisée sur $\mathbb{K}^p$ est la même pour définir $|||N|||$ et $|||M|||$, alors

$$|||MN||| \leqslant |||M||| \cdot |||N|||.$$

◁ **Éléments de preuve.**
On l'a déjà vu dans le contexte des AL continues. ▷

Plus généralement,

Théorème 5.6.14 – Continuité des applications multilinéaires en dimension finie

Soit $E_1, \dots, E_p, F$ des e.v.n.. Si les E_k, $k \in [\![1,p]\!]$ sont de dimension finie, alors toute application multilinéaire $u : E_1 \times \cdots \times E_p \to F$ est continue.

◁ **Éléments de preuve.**
Récurrence sur p. ▷

Exemples 5.6.15

1. Soit E de dimension finie, muni d'une base $\mathcal{B}$. Le déterminant $\det_{\mathcal{B}} : E \to \mathbb{K}$ est multilinéaire sur E, donc continue, quelle que soit la norme de E.
2. det : $\mathcal{M}_n(\mathbb{K})$ est continue par rapport à ses colonnes, pour n'importe quelle norme sur l'espace $\mathbb{K}^n$ des colonnes.
3. $(M,N) \mapsto MN$ de $\mathcal{M}_{n,p}(\mathbb{K}) \times \mathcal{M}_{p,q}(\mathbb{K})$ dans $\mathcal{M}_{n,q}(\mathbb{K})$ est continue.
4. De même, la composition des AL entre espaces de dimension finie.

VI.4 Applications polynomiales

Définition 5.6.16 – Application polynomiales de n variables

Soit $A \subset \mathbb{K}^n$ et $f : A \to \mathbb{K}$. On dit que f est une application polynomiale si f est une combinaison linéaire d'applications

$$(x_1, \dots, x_n) \mapsto x_1^{a_1} \cdots x_n^{a_n},$$

où $(a_1, \dots, a_n) \in \mathbb{N}^n$.

Exemple 5.6.17

L'application f définie sur $\mathbb{R}^3$ par $f(x,y,z) = 3x^2yz^4 - xy + 2$ est une application polynomiale.

Proposition 5.6.18

Les applications polynomiales sont continues sur leur domaine de définition, et ceci indépendamment des normes choisies.

◁ **Éléments de preuve.**

Travailler avec la norme ∞ pour obtenir la continuité des projections, puis les règles sur le produit permettent de conclure. ▷

Exemple 5.6.19

Étudier la continuité de $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par

$$f : (x, y) \mapsto \left(\ln(x^2 + xy + y^2 + 1), \frac{2\cos(xy)}{x^2 + y^4 + 1} \right)$$

Théorème 5.6.20 – Continuité de tr et $\det$

$\mathrm{tr} : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ et $\det : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ sont continues en tant qu'applications de n^2 variables scalaires.

◁ **Éléments de preuve.**

Avec ce point de vue, tr et $\det$ sont polynomiales. La continuité de tr provient également de sa linéarité (et de la dimension finie). ▷

Plus généralement, on peut définir des applications polynomiales sur un espace quelconque :

Définition 5.6.21 – Application polynomiale

Soit E et F deux espaces de dimension finie, et $f : E \to F$. Les propositions suivantes sont équivalentes:

(i) Il existe une base $\mathcal{B}$ de E et une base $\mathcal{C}$ de F tels que les coordonnées $f_k(x)$ de $f(x)$ sur $\mathcal{C}$ soient polynomiales en les coordonnées de x sur $\mathcal{B}$;

(ii) Pour toute base $\mathcal{B}$ de E et toute base $\mathcal{C}$ de F, les coordonnées $f_k(x)$ de $f(x)$ sur $\mathcal{C}$ sont polynomiales en les coordonnées de x sur $\mathcal{B}$;

On dit dans ce cas que f est une application polynomiale.

En travaillant coordonnée par coordonnée (sur F) avec la norme $\|\cdot\|_{\mathcal{B},\infty}$ sur E, les raisonnements ci-dessus s'adaptent bien, et on obtient :

Proposition 5.6.22 – Continuité des applications polynomiales

Soit E et F des espaces de dimension finie, et $f : E \to F$ une application polynomiale. Alors f est continue.

6

Séries numériques et vectorielles – rappels et compléments

I Rappels et compléments sur les séries numériques

Toutes les séries de cette section sont à terme général dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.
On rappelle rapidement les définitions et propriétés principales.

I.1 Convergences

Définition 6.1.1 – Convergence d'une série

Soit (u_n) une suite de réels ou complexes.

(i) Les sommes partielles S_n de la série de terme général u_n sont :

$$\forall n \in \mathbb{N}, \quad S_n = \sum_{k=0}^{n} u_k.$$

(ii) On dit que la série $\sum u_n$ de terme général u_n *converge* si la suite $(S_n)_{n\in\mathbb{N}}$ de ses sommes partielles admet une limite finie.

(iii) On note alors

$$\sum_{n=0}^{+\infty} u_n = \lim_{n\to+\infty} S_n = \lim_{n\to+\infty} \sum_{k=0}^{n} u_k.$$

Cette quantité est appelée *somme* de la série de terme général u_n.

(iv) Une série non convergente est dite *divergente*.

(v) Si $\sum\limits_{n\in\mathbb{N}} u_n$ converge, on définit les restes de la série par

$$\forall n \in \mathbb{N}, \quad R_n = \sum_{k=n+1}^{+\infty} u_k = \sum_{k=0}^{+\infty} u_k - S_n.$$

Proposition/Définition 6.1.2 – Divergence grossière

Si (u_n) ne tend pas vers 0, alors $\sum u_n$ diverge. On dit que la divergence est grossière.

◁ **Éléments de preuve.**
Si la série converge, $u_n = S_n - S_{n-1}$ tend vers 0. ▷

Avertissement 6.1.3 – Divergence non grossière

La réciproque est fausse. Autrement dit, il existe des divergences non grossières, comme la série harmonique $\sum \frac{1}{n}$.

Théorème 6.1.4 – Théorème de comparaison des séries à termes positifs, TCSTP

Si pour tout $n \in \mathbb{N}$, $0 \leqslant u_n \leqslant v_n$ et $\sum v_n$ est convergente, alors $\sum u_n$ est convergente.

◁ **Éléments de preuve.**

Sous ces hypothèses, la somme partielle de la série $\sum u_n$ est croissante, et majorée par $\sum_{k=0}^{+\infty} v_k$. ▷

Définition 6.1.5 – Convergence absolue

On dit que $\sum u_n$ converge absolument si $\sum |u_n|$ converge

Théorème 6.1.6 – CVA implique CV

Toute série (réelle ou complexe) absolument convergente est convergente.

◁ **Éléments de preuve.**

- Pour le cas réel, utiliser le TCSTP avec $0 \leqslant u_n^+ \leqslant |u_n|$ et $0 \leqslant u_n^- \leqslant |u_n|$.
- Pour le cas complexe, utiliser le cas réel, et le TCSTP avec $0 \leqslant |\mathrm{Re}(u_n)| \leqslant |u_n|$ et $0 \leqslant |\mathrm{Im}(u_n)| \leqslant |u_n|$

▷

Proposition 6.1.7 – Linéarité

Soit λ et μ deux complexes.

1. Si $\sum u_n$ et $\sum v_n$ convergent, alors $\sum(\lambda u_n + \mu v_n)$ converge, et :
$$\sum_{n=0}^{+\infty}(\lambda u_n + \mu v_n) = \lambda \sum_{n=0}^{+\infty} u_n + \mu \sum_{n=0}^{+\infty} v_n.$$
2. Si $\sum u_n$ converge et $\sum v_n$ diverge, alors $\sum(u_n + v_n)$ diverge.
3. Si $\sum u_n$ et $\sum v_n$ divergent, on ne peut rien conclure sur $\sum u_n + v_n$.

◁ **Éléments de preuve.**

Combinaison de la propriété de linéarité des sommes finies et de la propriété de linéarité de la limite. ▷

I.2 Comparaisons séries-intégrales

Lemme 6.1.8 – Caractérisation discrète de la convergence de $\int_a^{+\infty} f(t)\,\mathrm{d}t$, $f \geqslant 0$

Soit f une application positive, continue par morceaux sur $[a, +\infty[$, et $n_0 \geqslant a$. L'intégrale $\int_a^{+\infty} f(t)\,\mathrm{d}t$ converge si et seulement si la série de terme général $\left(\int_n^{n+1} f(t)\,\mathrm{d}t\right)_{n \geqslant n_0}$ converge.

◅ **Éléments de preuve.**
C'est bien sûr Chasles, d'abord. Puis la croissance de $x \mapsto \int_{n_0}^{x} f(t)\ \mathrm{d}t$, qui assure l'existence d'une limite dans $\overline{\mathbb{R}}$ quand x tend vers $+\infty$. ▻

Théorème 6.1.9 – Théorème de comparaison entre série et intégrale

Soit $a \in \mathbb{R}_+$ et soit $f : [a, +\infty[\to \mathbb{R}$ une fonction décroissante continue par morceaux et positive. Les propositions suivantes sont équivalentes:

(i) $\sum_{n \geqslant a} f(n)$ converge ;

(ii) $\int_a^{+\infty} f(t)\ \mathrm{d}t$ converge ;

(iii) f est intégrable sur $[a, +\infty[$

◅ **Éléments de preuve.**
Les points (ii) et (iii) sont équivalents par positivité de f.
Pour l'équivalence entre (ii) et (iii), encadrer sur un intervalle $[n, n+1]$ par décroissance de f :

$$f(n+1) \leqslant \int_n^{n+1} f(t)\ \mathrm{d}t \leqslant f(n),$$

puis utiliser le TCSTP et le lemme. ▻

On en déduit notamment :

Théorème 6.1.10 – Nature des séries de Riemann

Soit $\alpha \in \mathbb{R}$. La série de Riemann $\sum_{n \geqslant 1} \frac{1}{n^\alpha}$ converge si et seulement si $\alpha > 1$.

Théorème 6.1.11 – Nature des séries de Bertrand, HP

Soit $(\alpha, \beta) \in \mathbb{R}^2$. La série de Bertrand $\sum_{n \geqslant 1} \frac{1}{n^\alpha \ln(n)^\beta}$ de paramètre (α, β) converge si et seulement si $(\alpha, \beta) > (1, 1)$ (pour l'ordre lexicographique)

◅ **Éléments de preuve.**
On avait montré le théorème similaire pour les intégrales. Il faut juste s'assurer de la monotonie àpcr. Pour $\alpha \neq 1$, on peut aussi le retrouver facilement par comparaison à une série de Riemann. ▻

Les techniques de comparaison entre séries et intégrales permettent en général d'obtenir mieux que juste la convergence ou la divergence d'une série. En effet, il s'agit d'un encadrement en général suffisamment fin de la série par deux intégrales pour obtenir une estimation (majoration, ou équivalent) de sommes partielles ou de restes. Il n'y a pas de résultat particulier à connaître, simplement une bonne maîtrise de ces techniques de comparaison. Pour citer le programme officiel : « Les étudiants doivent savoir utiliser la comparaison série-intégrale pour établir des convergences et des divergences de séries, estimer des sommes partielles de séries divergentes, ou des restes de séries convergentes, notamment dans le cas d'une fonction monotone. »
Nous voyons donc un certain nombre d'exemples, à bien étudier pour être en mesure de les adapter à d'autres situations.

Exemples 6.1.12

1. Trouver un équivalent de $\sum_{k=1}^{n} \frac{1}{k^\alpha}$ pour $\alpha \leqslant 1$, de $\sum_{k=n+1}^{+\infty} \frac{1}{k^\alpha}$ pour $\alpha > 1$.

2. Déterminer un équivalent de $\displaystyle\sum_{k=n+1}^{+\infty} \frac{\ln(k)}{k^2}$.

3. Déterminer un équivalent de $\displaystyle\sum_{k=0}^{n} \frac{\ln(k)}{\sqrt{k}}$

I.3 Théorèmes de comparaison pour la convergence absolue

On dispose d'un certain nombre d'outils de comparaison, permettant d'obtenir la convergence d'une série en la comparant à une série de référence. Les séries de référence sont :

- Les séries géométriques $\sum a^n$, qui convergent ssi $|a| < 1$;
- Les séries de Riemann $\sum \frac{1}{n^\alpha}$, qui convergent ssi $\alpha > 1$;
- (HP) les séries de Bertrand $\sum \frac{1}{n^\alpha \ln(n)^\beta}$, qui convergent ssi $(\alpha, \beta) > (1, 1)$ pour l'ordre lexicographique.

Les comparaisons s'établissent en général dans le cadre des séries à termes positifs mais s'utilisent de fait souvent pour l'étude de la convergence absolue. Les énoncer dans ce contexte ne diminue pas la généralité, puisque pour une série à termes positifs, la convergence équivaut à la convergence absolue.

Théorème 6.1.13 – Théorème de comparaison des séries par o, O, $\sim$

Soit (u_n) et (v_n) des suites réelles ou complexes.

1. Si $u_n = O(v_n)$ (et a fortiori si $u_n = o(v_n)$), et si $\sum v_n$ converge absolument, alors $\sum u_n$ converge absolument.
2. Si $u_n \underset{+\infty}{\sim} v_n$ alors $\sum u_n$ CVA si et seulement si $\sum v_n$ CVA.

◁ **Éléments de preuve.**

Un O se traduit par une inégalité àpcr sur les valeurs absolues, ce qui nous ramène au TCSTP. Un équivalent se réécrit avec o. ▷

Avertissement 6.1.14

La comparaison par équivalents permet de comparer la nature de séries de signe constant, ou la CVA, mais ne permet pas d'établir une semi-convergence, par exemple.

Une façon commode de comparer à une série géométrique consiste à comparer asymptotiquement le quotient $\frac{u_{n+1}}{u_n}$ au quotient similaire (constant égal à la raison) qu'on obtiendrait avec une série géométrique. Cela est à la base de la règle suivante :

Proposition 6.1.15 – Règle de d'Alembert

Soit $\sum u_n$ une série réelle ou complexe telle que pour tout $n \in \mathbb{N}$, $u_n \neq 0$ (ou au moins à partir d'un certain rang). On suppose que $\left|\frac{u_{n+1}}{u_n}\right| \longrightarrow \ell \in \overline{\mathbb{R}}_+$.

(i) Si $0 \leqslant \ell < 1$, alors $\sum u_n$ converge absolument.

(ii) Si $\ell > 1$, alors $\sum u_n$ diverge grossièrement.

(iii) Le cas $\ell = 1$ est indéterminé.

◁ **Éléments de preuve.**

(i) si $\ell < a < 1$, àpcr, $|u_n| \leqslant a^{n-n_0}|u_{n_0}|$.

(ii) àpcr, $|u_{n+1}| > |u_n|$, ce qui empêche la convergence vers 0 du terme général

(iii) Considérer par exemple les séries de Riemann qui se placent toutes dans ce cas.

▷

Remarque 6.1.16

La règle de Riemann sera un outil important dans l'étude du domaine convergence des séries entières $\sum a_n x^n$. En effet si $\left|\frac{a_{n+1}}{a_n}\right| \to \ell$, alors

$$\left|\frac{a_{n+1}x^{n+1}}{a_n x^n}\right| \longrightarrow \ell|x|,$$

et la série va converger absolument si $|x| < \frac{1}{\ell}$ et diverger grossièrement si $|x| > \frac{1}{\ell}$ (dans $\mathbb{C}$). Ainsi, le domaine est un disque pouvant contenir (ou non) certains points de sa frontière. Le réel $\frac{1}{\ell}$ est appelé rayon de convergence de la série entière.
On montrera que le domaine de convergence des séries entières est toujours de cette forme, même lorsque la règle de d'Alembert n'est pas applicable.

Exemples 6.1.17

1. Rayon de convergence de la série exponentielle $\sum \frac{z^n}{n!}$
2. Rayon de convergence de la série géométrique dérivée $\sum (n+1)x^n$.
3. Rayon de convergence de la série du logarithme $\sum\limits_{n\geqslant 1} \frac{x^n}{n}$.

I.4 Semi-convergence

Définition 6.1.18 – Semi-convergence

Une serie semi-convergente est une série convergente, mais pas absolument convergente.

Le seul outil au programme est :

Théorème 6.1.19 – Critère spécial de convergence des séries alternées

Soit (a_n) une suite décroissante de limite nulle. Alors $\sum(-1)^n a_n$ est appelée série alternée, et est convergente.

◁ **Éléments de preuve.**
En notant S_n la somme partielle, montrer que (S_{2n}) et (S_{2n+1}) sont adjacentes. ▷

Méthode 6.1.20

L'étude des séries $\sum \frac{(-1)^n}{n^\alpha}$ ou $\sum \frac{(-1)^n}{n^\alpha \ln(n)^\beta}$ est alors immédiate. On peut souvent s'y ramener par DL.

Exemple 6.1.21

Nature de $\sum \frac{(-1)^n}{(-1)^n + n^\alpha}$.

Méthode 6.1.22

- Dans des situations d'agencements de signes plus complexes, une bonne idée consiste souvent à regrouper les termes consécutifs de même signe (faire des paquets), puis :
  - ∗ vérifier que la somme sur chaque paquet ne tend pas vers 0, ce qui suffit à établir la divergence ;
  - ∗ ou vérifier que la sommation par paquet se ramène à l'étude d'une série alternée. Il faut ensuite contrôler ce qui se passe entre les paquets, mais c'est assez « automatique » vu

les propriétés de signes.
- On peut aussi faire des paquets qui permettent des compensations de signe pour se ramener à des paquets de signe constant.

I.5 Techniques asymptotiques

Pour commencer, on adapte au cas discret les théorèmes de sommation des relations de comparaison vus pour les intégrales.

Théorème 6.1.23 – Sommation des relations de comparaison

Soit $\sum u_n$ et $\sum v_n$ deux séries numériques. On suppose que $\sum v_n$ est à termes réels de signe constant (au moins àpcr).
On note $S_n(u)$ (resp. $S_n(v)$) les sommes partielles de $\sum u_n$ (resp. de $\sum v_n$), et en cas de convergence, on note

$$R_n(u) = \sum_{k=n+1}^{+\infty} u_k \qquad \left(\text{resp. } R_n(v) = \sum_{k=n+1}^{+\infty} v_k\ \right)$$

les restes de la série $\sum u_n$ (resp. $\sum v_n$).

1. Si $\sum v_n$ converge,
   - si $u_n \underset{+\infty}{=} O(v_n)$, alors $\sum u_n$ converge, et $R_n(u) \underset{+\infty}{=} O(R_n(v))$;
   - si $u_n \underset{+\infty}{=} o(v_n)$, alors $\sum u_n$ converge, et $R_n(u) \underset{+\infty}{=} o(R_n(v))$;
   - si $u_n \underset{+\infty}{\sim} v_n$, alors $\sum u_n$ converge, et $R_n(u) \underset{+\infty}{\sim} R_n(v)$;
2. Si $\sum v_n$ diverge,
   - si $u_n \underset{+\infty}{=} O(v_n)$, alors $S_n(u) \underset{+\infty}{=} O(S_n(v))$;
   - si $u_n \underset{+\infty}{=} o(v_n)$, alors $S_n(u) \underset{+\infty}{=} o(S_n(v))$;
   - si $u_n \underset{+\infty}{\sim} v_n$, alors $S_n(u) \underset{+\infty}{\sim} S_n(u)$.

◁ **Éléments de preuve.**
La démonstration peut se faire par ε exactement comme dans le cas des intégrales. On peut aussi être flemmard et se ramener au cas des comparaisons intégrales, en introduisant les fonctions $f_u : x \mapsto u_{\lfloor x \rfloor}$ et $f_v : x \mapsto v_{\lfloor x \rfloor}$. ▷

Corollaire 6.1.24 – Théorème de Cesàro

Soit $u_n \longrightarrow \ell \in \mathbb{C}$ (ou $\overline{\mathbb{R}}$ si $u_n \in \mathbb{R}^{\mathbb{N}}$). Alors la suite des moyennes a la même limite :

$$\frac{1}{n+1} \sum_{k=0}^{n} u_k \longrightarrow \ell.$$

◁ **Éléments de preuve.**
- Si ℓ fini, écrire $u_n - \ell = o(1)$ et sommer.
- Si $\ell = +\infty$, $1 = o(u_n)$ et sommer. Adapter pour $-\infty$.

▷

On retrouve, avec une technique un peu allégée, les équivalents trouvés précédemment pour les séries de Riemann. Le résultat lui-même est HP, mais il est attendu qu'on sache le retrouver rapidement.

Corollaire 6.1.25 – Équivalent des séries de Riemann revisité (HP)

1. Si $\alpha < 1$, $\displaystyle\sum_{k=1}^{n} \frac{1}{k^a} \underset{+\infty}{\sim} \frac{1}{1-\alpha} \frac{1}{k^{\alpha-1}}$;

2. Si $\alpha > 1$, $\displaystyle\sum_{k=n+1}^{+\infty} \frac{1}{k^a} \underset{+\infty}{\sim} \frac{1}{\alpha-1}\frac{1}{k^{\alpha-1}}$;
3. Si $\alpha = 1$, $\displaystyle\sum_{k=1}^{n} \frac{1}{k} \underset{+\infty}{\sim} \ln(n)$

◁ **Éléments de preuve.**

Procéder par télescopage, en trouvant un équivalent simple de $\dfrac{1}{(n+1)^{\alpha-1}} - \dfrac{1}{n^{\alpha-1}}$, ou de $\ln(n+1) - \ln(n)$ dans le dernier cas. ▷

Méthode 6.1.26 – DA de sommes partielles ou de restes

- Trouver un équivalent simple du tg, puis, suivant le cas, un équivalent du reste ou de la somme partielle par sommation.
- Retrancher l'équivalent à la somme partielle ou au reste. Cela définit une nouvelle suite (u_n) qu'on peut étudier via la série téléscopique $\sum u_n - u_{n-1}$.
- On peut itérer. Dans les situations « standard », à chaque étape, on est ramené à une somme de Riemann, via un développement limité.
- Attention à la bascule à faire des sommes partielles aux restes, quand on affine suffisamment le DA pour passer du cas divergent au cas convergent. Cela introduit en général une constante dans le DA, qu'on n'est pas en mesure de calculer.

On peut appliquer cela (sauf la première étape déjà faite) aux séries de Riemann elles-mêmes. Par exemple :

Exemple 6.1.27 – DA de la série harmonique

Montrer qu'il existe une constante γ telle que

$$\sum_{k=1}^{n} = \ln(n) + \gamma + \frac{1}{2n} + o\left(\frac{1}{n}\right).$$

La constante γ est appelée constante d'Euler-Mascheroni. À l'heure actuelle, malgré de nombreuses recherches à ce propos, on ne sait toujours pas si γ est rationnelle ou non.

Exemples 6.1.28

- Trouver un DA à la précision $o\left(\frac{1}{n^2}\right)$ de $\displaystyle\sum_{k=1}^{n} \sin\left(\frac{1}{k}\right)$. On n'explicitera pas la constante.
- Trouver un DA à la précision $o\left(\frac{1}{n}\right)$ de $\displaystyle\sum_{k=1}^{n} k^2 \mathrm{e}^{\frac{1}{k}}$. À nouveau, on n'explicitera pas la constante.

II Séries dans un e.v.n.

Le but de ce paragraphe est d'adapter rapidement les définitions et propriétés générales vues sur les séries numériques dans le cadre de séries à valeurs dans un e.v.n.

II.1 Définitions et propriétés générales

Définition 6.2.1 – Convergence d'une série dans un e.v.n.

- Soit $(E, \|\cdot\|)$ un e.v.n., et $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. On dit que la série $\sum u_n$ converge dans E si la suite des sommes partielles $\displaystyle\sum_{k=0}^{n} u_k$ admet une limite $S \in E$ (au sens de la norme $\|\cdot\|$)

- Le vecteur S est alors appelé somme de la série $\sum u_n$, et on note

$$S = \sum_{n=0}^{+\infty} u_n = \lim_{N\to+\infty} \sum_{n=0}^{N} u_n.$$

On adapte évidemment au cas d'une série ne commençant pas au rang 0.

Exemple 6.2.2

1. Soit $E = \mathbb{R}^{\mathbb{N}}$ et $u_n = (\delta_{n,m})_{m\in\mathbb{N}} = \mathbb{1}_{\{n\}}(m)$. On munit E de la norme
$$\|(x_m)_{m\in\mathbb{N}}\| = \sup_{k\in\mathbb{N}} \frac{|u_k|}{k}.$$
Montrer que $\sum u_n$ converge, et en donner sa somme.
2. On reprend le même exemple, mais on munit cette fois E de la norme
$$\|(x_m)_{m\in\mathbb{N}}\| = \sup_{k\in\mathbb{N}} |u_k|.$$
La série est-elle encore convergente ?

La technique mise en oeuvre dans le deuxième exemple nous montre que certains principes d'étude des suites numériques s'adaptent, notamment :

Proposition/Définition 6.2.3 – Divergence grossière

1. Soit $\sum u_n$ une série convergente dans un e.v.n. E. Alors $u_n \to 0_E$.
2. Par contraposée, une série dont le terme général ne tend pas vers 0 est divergente. On dit dans ce cas que la série est grossièrement divergente.

Un certain nombre de propriétés connues dans le cadre numérique restent vraies. Par exemple :

Proposition 6.2.4 – Linéarité de la somme

Soit E un e.v.n. et $\sum u_n$ et $\sum v_n$ deux séries d'éléments de E, et $\lambda \in \mathbb{K}$.

1. Si $\sum u_n$ et $\sum v_n$ convergent, alors $\sum(u_n + \lambda v_n)$ aussi, et
$$\sum_{n=0}^{+\infty}(u_n + \lambda v_n) = \sum_{n=0}^{+\infty} u_n + \lambda \sum_{n=0}^{+\infty} v_n.$$
2. Si $\sum u_n$ converge et $\sum v_n$ diverge, $\sum u_n + v_n$ diverge.

Remarque 6.2.5

1. Que dire si $\sum u_n$ et $\sum v_n$ divergent ?
2. Si on considère une somme de p séries, à quelles conditions de convergence peut-on déduire à coup sûr la nature de la série de la somme ?

Dans la plupart des cas, on essayera de se ramener au cadre réel positif, de la même manière que lorsqu'on étudie la convergence absolue dans le cadre des séries numériques, grâce au théorème qui suit. Attention à l'hypothèse de finitude indispensable dans ce théorème.
Comme dans le cas réel, l'étude de la convergence de suites peut se ramener à l'étude de la convergence de séries télescopiques :

Théorème 6.2.6 – Lien suite-série

Soit E un e.v.n. et $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. La suite $(u_n)_{n\in\mathbb{N}}$ est de même nature que la série télescopique $\sum u_{n+1} - u_n$.

II.2 Séries à valeurs dans un e.v.n. de dimension finie

Dans le cadre de la dimension finie, on peut souvent se ramener au cas de séries réelles positives, comme ce qu'on fait en étudiant la convergence absolue de séries numériques. On remarquera que dans ce cas, la propriété de convergence est indépendante du choix de la norme.
On commence par donner une définition qui a du sens en dimension quelconque.

Définition 6.2.7 – Convergence absolue, CVA

Soit $(E, \|\cdot\|)$ un e.v.n. et $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. On dit que $\sum u_n$ converge absolument (ou normalement) si la série numérique $\sum \|u_n\|$ est convergente.

Attention en revanche à l'hypothèse de finitude de la dimension dans le théorème suivant.

Théorème 6.2.8 – CVA implique CV

Soit E un e.v.n. de dimension finie, et $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. Si $\sum u_n$ converge absolument, alors $\sum u_n$ converge.

$\triangleleft$ **Éléments de preuve.**
Travailler avec la norme $\|\cdot\|_\infty$ pour se ramener à la convergence absolue coordonnée par coordonnée.
$\triangleright$

Exemple 6.2.9

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors la série $\sum \frac{A^n}{n!}$ converge.

Définition 6.2.10 – Exponentielle matricielle

Soit $n \in \mathbb{N}^*$, et $A \in \mathcal{M}_n(\mathbb{K})$. On définit

$$\exp(A) = \sum_{n=0}^{+\infty} \frac{A^n}{n!}.$$

De même, si E est un e.v.n. de dimension finie, on définit, pour $u \in \mathcal{L}(E)$:

$$\exp(u) = \sum_{n=0}^{+\infty} \frac{u^n}{n!}.$$

L'exponentielle matricielle est notamment l'outil adéquat pour la résolution des systèmes différentiels, généralisant ainsi la résolution connue en dimension 1.
Nous étudierons plus en détail les propriétés de l'exponentielle matricielle dans un chapitre ultérieur.

7

Structures algébriques

Ce chapitre a pour but de compléter l'étude des structures algébriques vues en MPSI, en approfondissant notamment les propriétés combintoires, algébriques et arithmétiques.

I Groupes

I.1 Rappels

Définition 7.1.1 – Groupe

On rappelle qu'un groupe G est un ensemble muni d'une loi de composition interne $\star$ telle que :
- $\star$ est associative ;
- il existe un élément neutre e pour la loi $\star$;
- tout élément x admet un symétrique x^s.

Le neutre est alors unique, ainsi que le symétrique d'un élément x donné.

Définition 7.1.2 – Groupe abélien ou commutatif

On dit qu'un groupe $(G, \star)$ est abélien (ou commutatif) si la loi de G est commutative.

On rencontre dans la pratique les deux conventions d'écriture suivantes.

Notation 7.1.3 – Notation additive, notation multiplicative

- Dans le cas général, on utilise souvent une notation multiplicative. Ainsi, la loi est $\times$, le neutre est souvent noté 1, le signe opératoire peut être omis (ab désigne $a \times b$) et on désignera sous forme de puissance l'itération de la loi, ainsi que la prise de symétrique ($x^s = x^{-1}$, on parlera alors d'inverse).
- Dans le cas abélien, on peut trouver la notation additive $+$, avec les conventions usuelles : le neutre est souvent noté 0, l'itération est notée nx, le symétrique est noté $-x$ et est appelé opposé.

Une structure algébrique vient toujours accompagnée des applications qui ont un sens dans le contexte, c'est-à-dire qui respectent la structure.

Définition 7.1.4 – Homomorphisme de groupes

Soit $(G, \star)$ et $(H, \diamond)$ deux groupes.
- On dit qu'une application $f : G \longrightarrow H$ est un homomorphisme de groupes (ou plus simplement morphisme de groupes) si pour tout $(x, y) \in G$, $f(x \star y) = f(x) \diamond f(y)$.
  On note $\mathrm{Hom}(G, H)$ l'ensemble des homomorphismes de G dans H.

- Si $(G, \star) = (H, \diamond)$, on dit que f est un endomorphisme de $(G, \star)$.
- Un homomorphisme bijectif est appelé isomorphisme.
- Un endomorphisme bijectif est appelé automorphisme.
  On note $\mathrm{Aut}(G)$ l'ensemble des automophismes de G.

Proposition 7.1.5 – Image du neutre par un morphisme

Soit $f : G \to H$ un morphisme de groupes. Alors $f(e_G) = e_H$.

Proposition 7.1.6 – Image par un morphisme d'un inverse

Soit G, H deux groupes (notés multiplicativement), et f un morphisme de G dans H. Alors $f(x^{-1}) = f(x)^{-1}$. On adaptera aisément cette propriété au cas ou l'un (ou les deux) groupe(s) est (sont) en notation additive.

◁ **Éléments de preuve.**
Considérer $f(x)f(x^{-1})$. ▷

Proposition 7.1.7 – composée et réciproque de morphismes

Soit G, H et K trois groupes.

1. Si $f : G \to H$ et $g : H \to K$ sont deux morphismes de groupes, alors $g \circ f$ est aussi un morphisme de groupes.
2. Si $f : G \to H$ est un isomorphisme, alors $f^{-1} : H \to F$ est aussi un isomorphisme.

◁ **Éléments de preuve.**
Vérifications faciles pour la composée
Pour la réciproque, considérer $f^{-1}(f(f^{-1}(x)f^{-1}(y)))$. ▷

I.2 Sous-groupes

Définition 7.1.8 – Sous-groupe

Soit $(G, \star)$ un groupe. Un sous-ensemble H de G est appelé *sous-groupe de* G si H est stable pour la loi de G et si la loi induite définit sur H une structure de groupe.

Le neutre de H est alors nécessairement le même que celui de G.
Dans la pratique, pour vérifier que H est un sous-groupe de G on utilise le résultat suivant, ou sa version condensée :

Théorème 7.1.9 – Caractérisation des sous-groupes

Un sous-ensemble H d'un groupe $(G, \star)$ (de neutre e_G) est un sous-groupe de G si et seulement si :

(i) H est non vide,

(ii) H est stable pour $\star$: $\forall (x, y) \in H, \quad x \star y \in H$,

(iii) H est stable par prise de symétrique : $\forall x \in H, \quad x^s \in H$.

Dans la pratique, pour montrer que H est non vide, on vérifie souvent que $e_G \in H$.

Théorème 7.1.10 – Caractérisation des sous-groupes, version condensée

Un sous-ensemble H d'un groupe $(G, \star)$ (de neutre e_G) est un sous-groupe de G si et seulement si :

(i) H est non vide,

(ii) $\forall (x,y) \in H^2, \quad x \star y^s \in H$.

On traduit cette dernière propriété dans les deux cas les plus fréquents :

- pour un sous-groupe d'un groupe additif, la vérification de stabilité à faire est donc :

$$\forall (x,y) \in H^2,\ x - y \in H;$$

- pour un sous-groupe d'un groupe multiplicatif, la vérification de stabilité à faire est donc :

$$\forall (x,y) \in H^2, \quad xy^{-1} \in H.$$

On trouve notamment des sous-groupes en prenant des images directes ou réciproques de sous-groupes connus.

Proposition 7.1.11 – Image directe et réciproque de sous-groupes par un homomorphisme

Soit G et H deux groupes, et soit $f \in \mathrm{Hom}(G,H)$ un morphisme de groupes.

1. Soit G' un sous-groupe de G. Alors $f(G')$ est un sous-groupe de H.
2. Soit H' un sous-groupe de H. Alors $f^{-1}(H')$ est un sous-groupe de G.

Un cas particulier important en est le noyau

Définition 7.1.12 – Noyau

Soit G et H deux groupes et $f \in \mathrm{Hom}(G,H)$ un morphisme de groupes. Le noyau de f est le sous-groupe de G défini par :

$$\mathrm{Ker}(f) = f^{-1}(\{e_H\}) = \{y \in G \mid f(y) = e_H\}.$$

Une propriété importante du noyau est qu'il mesure le défaut d'injectivité :

Théorème 7.1.13 – Caractérisation de l'injectivité

Soit $f \in \mathrm{Hom}(G,H)$. Le morphisme f est injectif si et seulement si $\mathrm{Ker}(f) = \{e_G\}$.

Une autre façon de construire un sous-groupe à partir d'autres est d'en prendre l'intersection :

Proposition 7.1.14 – Intersection de sous-groupes

Soit G un groupe, et $(H_i)_{i\in I}$ une famille de sous-groupes de G. Alors $\bigcap\limits_{i\in I} H_i$ est un sous-groupe de G.

◁ **Éléments de preuve.**

Sans difficulté à l'aide de l'une ou l'autre des caractérisations des sous-groupes. ▷

Cela nous permet de définir le plus petit sous-groupe contenant une partie donnée :

Proposition/Définition 7.1.15 – Sous-groupe engendré par une partie

Soit $(G,\times)$ un groupe, et X une partie de G. Le sous-groupe $\langle X\rangle$ de G engendré par X est le plus petit sous-groupe de G contenant X. Ce sous-groupe est défini explicitement par

$$\langle X\rangle = \bigcap_{X\subset H} H,$$

l'intersection étant prise sur tous les sous-groupes H de G.

Proposition 7.1.16 – Description par le bas du sous-groupe engendré par une partie

Soit X une partie d'une groupe G. Alors $\langle X\rangle$ est l'ensemble des éléments pouvant s'écrire sous la forme $x_1 \cdots x_n$, $n \in \mathbb{N}$, où les x_i vérifient soit $x_i \in X$, soit $x_i^{-1} \in X$. Le produit vide est par convention égal au neutre e de G.

◁ **Éléments de preuve.**
Justifier que l'ensemble de ces éléments est forcément inclus dans $\langle X\rangle$, et que c'est une sous-groupe contenant X. ▷

Définition 7.1.17 – Partie génératrice d'un groupe

Soit $X \subset G$ une partie de G. On dit que X est une partie génératrice de G si $\langle X\rangle = G$.

Lorsque X est un singleton on écrira plutôt $\langle x\rangle$ au lieu de $\langle\{x\}\rangle$, pour alléger les notations.
Par exemple $\mathbb{Z} = \langle 1\rangle$. Plus généralement, tous les sous-groupes de $\mathbb{Z}$ vont pouvoir se décrire avec un unique générateur :

Proposition 7.1.18

Les sous-groupes de $\mathbb{Z}$ sont les groupes $a\mathbb{Z} = \langle a\rangle$, $a \in \mathbb{N}$.

◁ **Éléments de preuve.**
Considérer a l'élément minimal de $G \cap \mathbb{N}^*$. Justifier que $a\mathbb{Z} \subset G$, et si l'inclusion est stricte, contredire la minimalité de a en effectuant une division euclidienne. ▷

I.3 Le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$

Soit $n \in \mathbb{N}^*$.
On rappelle les définitions suivantes, cas particulier des définitions générales liées aux relations d'équivalence.

Définition 7.1.19 – Classe de congruence, représentant

- Pour $a \in \mathbb{Z}$, on note $\overline{a}$ la classe de congruence de a modulo n, c'est-à-dire la classe d'équivalence de a pour la relation de congruence modulo n
- Explicitement, $\overline{a} = a + n\mathbb{Z}$.
- Étant donné c une classe de congruence, tout élément $a \in c$ est appelé représentant de la classe c. On a alors $c = \overline{a}$

Ainsi, a est un représentant de $\overline{a}$, mais aussi $a + n$, ou $a - 42n$.
On rappelle que la congruence modulo n respecte la somme et le produit, c'est-à-dire :

$$\forall (a, a', b, b') \in \mathbb{Z}^4, \quad \begin{cases} a \equiv a' \ [n] \\ b \equiv b' \ [n] \end{cases} \implies \begin{cases} a + b \equiv a' + b' \ [n] \\ ab \equiv a'b' \ [n] \end{cases}$$

Cela peut se résumer ainsi : la classe de congruence d'une somme ou d'un produit de deux représentants de deux classes de congruences ne dépend pas du choix de ces représentants. Cela permet de définir la somme et le produit directement sur les classes de congruence, ce qui donne la très importante définition suivante.

Définition 7.1.20 – $\mathbb{Z}/n\mathbb{Z}$ et ses lois

L'ensemble des classes de congruence modulo n est noté $\mathbb{Z}/n\mathbb{Z}$. Il peut être muni d'un somme et d'un produit, définis, pour $\overline{a}$ et $\overline{b}$ deux classes de $\mathbb{Z}/n\mathbb{Z}$, par

- $\overline{a} + \overline{b} = \overline{a + b}$;
- $\overline{a} \cdot \overline{b} = \overline{ab}$.

En vertu de ce qui précède, ces définitions sont indépendantes du choix des représentants a et b des deux classes.

Exemples 7.1.21

Décrire la table des lois de $\mathbb{Z}/n\mathbb{Z}$, pour $n \in \{2, 3, 4\}$.

Proposition 7.1.22 – Le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$

- $(\mathbb{Z}/n\mathbb{Z}, +)$ est un groupe abélien d'ordre n.
- L'application $\pi : \mathbb{Z} \mapsto \mathbb{Z}/n\mathbb{Z}$ définie par $\varphi(a) = \overline{a}$ est un morphisme de groupe, et $\mathrm{Ker}(\varphi) = a\mathbb{Z}$.

Proposition 7.1.23 – Générateurs de $\mathbb{Z}/n\mathbb{Z}$

Soit $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$. Les propositions suivantes sont équivalentes:

(i) $\mathbb{Z}/n\mathbb{Z} = \langle \overline{a} \rangle$

(ii) il existe $k \in \mathbb{N}$ tel que $\overline{ka} = 1$

(iii) $a \wedge n = 1$

◁ **Éléments de preuve.**

Montrer $(i) \iff (ii)$, qui est assez clair, et $(ii) \iff (iii)$ qui provient de Bezout. ▷

Proposition 7.1.24 – Définir un morphisme $\mathbb{Z}/n\mathbb{Z} \to G$

Soit $(G, \times)$ un groupe, et $f : \mathbb{Z} \to G$ un morphisme de groupe. Les propositions suivantes sont équivalentes:

(i) il existe un morphisme $g : \mathbb{Z}/n\mathbb{Z} \to G$ tel que $f = g \circ \pi$;

(ii) $n\mathbb{Z} \subset \mathrm{Ker}(f)$.

◁ **Éléments de preuve.**

Le sens direct est évident. Réciproquement (ii) donne la condition pour que la valeur de f soit la même pour tout représentant d'une classe de congruence, ce qui donne une façon naturelle de définit g. Le reste n'est que de la vérification facile. ▷

Ainsi, pour définir un morphisme de groupe issu de $\mathbb{Z}/n\mathbb{Z}$, on le définira d'abord sur $\mathbb{Z}$, puis on vérifiera l'hypothèse (ii) pour le « passer au quotient ».

Proposition 7.1.25 – Comparaison de $\mathbb{Z}/n\mathbb{Z}$ et $\mathbb{U}_n$

Les groupes $(\mathbb{Z}/n\mathbb{Z}, +)$ et $(\mathbb{U}_n, \times)$ sont isomorphes.

◁ **Éléments de preuve.**

Considérer $\mathbb{Z} \longrightarrow \mathbb{U}_n$ défini par $f(k) = \mathrm{e}^{\frac{2\,\mathrm{i}\,\pi k}{n}}$, puis passer au quotient. ▷

I.4 Groupes monogènes

Définition 7.1.26 – Groupe monogène

1. Un groupe G est monogène s'il existe $x \in G$ tel que $G = \langle x \rangle$.
2. Si G est fini et monogène, on dit que G est un groupe cyclique.
3. Un générateur d'un groupe monogène est un élément x tel que $G = \langle x \rangle$.

Exemples 7.1.27

1. $(\mathbb{Z}, +)$ est monogène, engendré par 1 ou par -1. Ce sont les seuls générateurs possibles de $\mathbb{Z}$.
2. Tous les sous groupes de $\mathbb{Z}$ sont monogènes. En effet, ils sont de la forme $a\mathbb{Z} = \langle a \rangle = \langle -a \rangle$. Ce sont les seuls éléments générateurs de $a\mathbb{Z}$.
3. $(\mathbb{Z}/n\mathbb{Z}, +)$ est monogène, ses éléments générateurs sont les $\overline{k}$, avec $k \wedge n = 1$.
4. $(\mathbb{U}_n, \times)$ est monogène. Ses générateurs sont les images par l'isomorphisme de la proposition 7.1.25 des générateurs de $\mathbb{Z}/n\mathbb{Z}$, à savoir $\mathrm{e}^{\frac{2\,\mathrm{i}\,k\pi}{n}}$, avec $k \wedge n = 1$. Ces éléments générateurs sont appelés racines $n$ièmes primitives de l'unité.

Proposition 7.1.28 – Description explicite d'un sous-groupe monogène

Soit $(G, \times)$ un groupe, et $x \in G$. Alors

$$\langle x \rangle = \{x^k, \quad k \in \mathbb{N}\} = \mathrm{Im}(\varphi_x),$$

où $\varphi_x : \mathbb{Z} \mapsto G$ est le morphisme défini par $\varphi_x(k) = x^k$.

Le moyau de φ_x étant un sous-groupe de $\mathbb{Z}$, il est égal à un $n\mathbb{Z}$. En distinguant suivant que $n = 0$ ou $n > 0$, on obtient alors la description suivante des groupes monogènes.

Théorème 7.1.29 – Structure des groupes monogènes

Tout groupe monogène $\langle x \rangle$ est isomorphe :
- soit à $(\mathbb{Z}, +)$, si $\mathrm{Ker}(\varphi_x) = \{0\}$;
- soit à $(\mathbb{Z}/n\mathbb{Z}, +)$, si $\mathrm{Ker}(\varphi_x) = n\mathbb{Z}$; dans ce cas, $n = \min\{k \in \mathbb{N}^*,\ x^k = e\}$.

◁ **Éléments de preuve.**
Dans le cas 2, passer φ_x au quotient par la proposition 7.1.24 pour définir $\widetilde{\varphi_x} : \mathbb{Z}/n\mathbb{Z} \to G$. Vérifier que c'est un isomorphisme. ▷

I.5 Ordre d'un élement

Le dernier théorème nous incite à définir, s'il existe, le premier retour à l'élément neutre dans la suite des puissances d'un élément x de G

Définition 7.1.30 – Ordre d'un élément

Soit $(G, \times)$ un groupe, de neutre e, et $x \in G$. L'ordre de x est
- $\mathrm{ord}(x) = +\infty$ si pour tout $n \in \mathbb{N}^*$, $x^n \neq e$;
- $\mathrm{ord}(x) = \min\{n \in \mathbb{N}^* \mid x^n = e\}$ sinon.

On dit que x admet un ordre fini si $\mathrm{ord}(x) \neq +\infty$.

Exemples 7.1.31

1. Dans $(\mathbb{Z}, +)$, $\mathrm{ord}(1) = +\infty$
2. Dans $(\mathbb{Z}/n\mathbb{Z}, +)$, $\mathrm{ord}(\overline{1}) = n$.
3. Dans $(\mathbb{Z}/42\mathbb{Z}, +)$, $\mathrm{ord}(\overline{12}) = 7$.

Avec les notations du paragraphe précédent
- $\mathrm{ord}(x) = +\infty$ si et seulement si $\mathrm{Ker}(\varphi_x) = \{0\}$
- $\mathrm{ord}(x) = n \in \mathbb{N}^*$ si et seulement si $\mathrm{Ker}(\varphi_x) = n\mathbb{Z}$,

On en déduit en particulier :

Proposition 7.1.32 – Résolution de $x^k = e$, Spé

Soit $(G, \times)$ un groupe de neutre e. Soit x un élément d'ordre fini n. Alors

$$x^k = e \Longleftrightarrow n \mid k.$$

Par ailleurs, en vertu du théorème 7.1.29, on obtient donc :

Proposition 7.1.33 – Description de $\mathrm{ord}(x)$ par le groupe engendré

Soit $(G, \times)$ un groupe, et $x \in G$. Alors

$$\mathrm{ord}(x) = |\langle x \rangle|.$$

◁ **Éléments de preuve.**

Dans le cas où $\mathrm{ord}(x)$ est fini, égal à n, le théorème 7.1.29 donne $\langle x \rangle \simeq \mathbb{Z}/n\mathbb{Z}$. Passer au cardinal. ▷

Définition 7.1.34 – Ordre d'un groupe

L'ordre d'un groupe est son cardinal. C'est un élément de $\mathbb{N}^* \cup \{+\infty\}$

Ainsi, le dernier résultat se réexprime en disant que l'ordre d'un élément est égal à l'ordre du sous-groupe qu'il engendre.

Corollaire 7.1.35

Si G est un groupe fini, tout $x \in G$ admet un ordre fini.

◁ **Éléments de preuve.**

Sinon, $\mathbb{Z} \simeq \langle x \rangle \subset G$. ▷

On peut préciser un peu le corollaire précédent :

Théorème 7.1.36 – Lagrange, admis sauf dans le cas abélien

Soit G un groupe fini, et $x \in G$. Alors $\mathrm{ord}(x) \mid \mathrm{Card}(G)$.

Ainsi l'ordre d'un élément d'un groupe fini divise l'ordre de ce groupe.

◁ **Éléments de preuve.**

La démonstration n'est exigible que dans le cas où G est abélien. Montrer que l'application $\varphi : g \mapsto xg$ est une bijection et en déduire que $\prod_{g \in G} g = \prod_{g \in G} (xg)$, puis simplifier. ▷

I.6 Un deuxième théorème de Lagrange, HP

On esquisse dans cette section, pour les lecteurs intéressés, la preuve générale du théorème de Lagrange qui clôt le paragraphe précédent. Tout ce paragraphe est hors-programme. Les groupes sont tous notés multiplicativement.

Lemme 7.1.37 – Partition par les classes à gauche modulo H

Soit G un groupe fini et H un sous-groupe de G. Alors $\{aH,\ a \in G\}$ est une partition de G

◁ **Éléments de preuve.**

C'est la partition associée à la relation d'équivalence définie sur G par

$$x \equiv_g y \quad [H] \Longleftrightarrow x^{-1}y \in H$$

Cette relation est appelée relation de congruence à gauche modulo H. ▷

Remarque 7.1.38

Lorsque $G = \mathbb{Z}$ et $H = n\mathbb{Z}$ (et en revenant aux notations additives), la congruence modulo H correspond à la congruence modulo n bien connue.

Lemme 7.1.39 – Cardinal des classes à gauche modulo H

Pour tout $a \in G$, $|aH| = |H|$.

◁ **Éléments de preuve.**
L'application $h \mapsto ah$ est une bijection de H dans aH. ▷

Théorème 7.1.40 – Lagrange pour les sous-groupes

Soit G un groupe fini et H un sous-groupe de G. Alors l'ordre de H divise l'ordre de G.

◁ **Éléments de preuve.**
On a partitionné G avec des parts qui ont toutes le même cardinal que H. ▷

Le théorème de Lagrange pour l'ordre des éléments s'en déduit en appliquant le théorème de Lagrange pour l'ordre des sous-groupes avec $H = \langle x \rangle$.

Remarque 7.1.41 – Groupes quotients

1. On aurait pu faire la même chose avec les classes à droite Ha.
2. Quand G n'est pas abélien, la partition des classes à gauche n'est en général par la même que la partition des classes à droite.
3. Cela peut néanmoins arriver. Cela se produit lorsque pour tout $a \in G$, $aH = Ha$, c'est-à-dire $aHa^{-1} = H$. On dit dans ce cas que H est un sous-groupe normal de G (terminologie anglaise) ou un sous-groupe distingué de G (terminologie française).
4. Dans ce cas, on peut montrer que la loi de G passe sur les classes, dans le sens où la classe (à droite ou gauche) de xy ne dépend pas du choix de x et y dans leur classe respective. Cela permet de définir une loi sur l'ensemble des classes modulo H, qui en fait un groupe. C'est ce qu'on appelle le groupe quotient G/H
5. Cette situation se produit systématiquement lorsque G est abélien, puisque dans ce cas, les classes à gauche et à droite sont clairement confondues. Ainsi, on peut toujours définir le groupe quotient G/H dans cette situation.
6. Un cas particulier est le cas du sous-groupe $n\mathbb{Z}$ de $\mathbb{Z}$: en formant le groupe quotient, on retrouve le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$ déjà étudié.

II Anneaux et idéaux

II.1 Rappels

Un anneau est un groupe abélien (additif) auquel on rajoute une loi multiplicative :

Définition 7.2.1 – Anneau

$(A, +, \times)$ (ou plus simplement A) est un anneau si :

(i) $(A, +)$ est un groupe abélien ; le neutre est noté 0_A ou 0 ;

(ii) $\times$ est associative, et admet un neutre 1_A ;

(iii) $\times$ est distributive sur $+$.

On dit que A est un anneau commutatif si de plus $\times$ est commutative.

Définition 7.2.2 – Corps

$(K, +, \times)$ est un corps si c'est un anneau commutatif, tel que tout $x \neq 0$ de K admet un inverse pour $\times$.

Les exemples importants sont :

Exemples 7.2.3

1. $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{Q}$ et $\mathbb{C}$ munis des opérations usuelles sont des anneaux. $\mathbb{R}$, $\mathbb{Q}$ et $\mathbb{C}$ sont même des corps.
2. Pour tout $n \in \mathbb{N}^*$, $\mathbb{Z}/n\mathbb{Z}$ est un anneau commutatif.
3. L'ensemble $\mathbb{K}[X]$ des polynômes à coefficients dans un corps $\mathbb{K}$ est un anneau commutatif
4. L'ensemble $\mathcal{M}_n(\mathbb{K})$ des matrices carrées est un anneau non commutatif.

Les éléments non nul d'un anneau ne sont donc en général pas inversibles, contrairement à un corps. Leur ensemble possède tout de même une structure intéressante.

Théorème 7.2.4 – Groupe des inversibles d'un anneau

Soit A un anneau. Alors l'ensemble des éléments inversibles de A, généralement noté $A^\times$ ou $U(A)$, est stable pour la loi $\times$, et la loi induite munit $A^\times$ d'une structure de groupe multiplicatif.

Définition 7.2.5 – Régularité et diviseurs de 0

Soit A un anneau commutatif. Un élément $a \in A$ est appelé :
- élément régulier s'il est simplifiable : $ax = ay$ implique $x = y$;
- diviseur de 0 s'il est non nul, et qu'il existe $b \neq 0$ tel que $ab = 0$.

Proposition 7.2.6 – Description des éléments réguliers

Soit A un anneau commutatif et $a \in A\backslash\{0\}$. Les propositions suivantes sont équivalentes:

(i) a est régulier.

(ii) a n'est pas diviseur de 0.

La régularité est à la base de la résolution de grand nombre d'équations, par factorisation. Voici un cadre qui nous permet d'exploiter ce type de technique :

Définition 7.2.7 – Anneau intègre

Un anneau A est intègre si :

(i) $A \neq \{0\}$;

(ii) A est commutatif ;

(iii) A ne possède pas de diviseur de 0.

La troisième propriété se traduit par : $ab = 0 \Longrightarrow (a = 0)$ ou $b = 0$.
Beaucoup de manipulations calculatoires se passent dans un anneau A comme dans $\mathbb{C}$ ou $\mathbb{Z}$, mais il faut se méfier :
- des inversions intempestives (si A n'est pas un corps)
- des simplifications abusives (si A n'est pas intègre)
- des problèmes de commutativité, qui nécessitent parfois l'introduction d'hypothèses supplémentaires, à vérifier scrupuleusement.

Exemple 7.2.8 – Calculs dans un anneau

Sous certaines conditions de commutativité, certains outils calculatoires connus sur $\mathbb{C}$ restent valides dans un anneau :

1. la formule du binôme pour le développement de $(A+B)^n$, lorsque $AB = BA$;
2. la factorisation de $A^n - B^n$ par $A - B$ (factorisation de Bernoulli), lorsque $AB = BA$.

Définition 7.2.9 – Homomorphisme d'anneaux

Soit A et B deux anneaux.

1. Un homomorphisme d'anneaux de A à B est une application $f : A \longrightarrow B$, soit égale à la fonction nulle, soit vérifiant :
$$\forall (x,y) \in A^2, \quad f(x+y) = f(x) + f(y), \qquad f(xy) = f(x)f(y) \qquad \text{et} \qquad f(1_A) = 1_B.$$
2. Un isomorphisme est un homomorphisme bijectif. Sa réciproque est alors aussi un morphisme d'anneau.

Définition 7.2.10 – Sous-anneau

Soit $(A, +, \times)$ un anneau. Un sous-anneau de A est un sous-ensemble $B \subset A$ tel que :

(i) B soit stable pour les lois $+$ et $\times$,

(ii) $1_A \in B$

(iii) les lois induites sur B définissent sur B une structure d'anneau.

Comme pour les groupes, on a une caractérisation assez efficace des sous-anneaux :

Proposition 7.2.11 – Caractérisation des sous-anneaux

Un sous-ensemble B d'un anneau A est un sous-anneau de A si et seulement si :

(i) $1_A \in B$.

(ii) pour tout $(x,y) \in B$, $x - y \in B$

(iii) pour tout $(x,y) \in B$, $xy \in B$

Enfin, comme pour les groupes, on peut munir un produit cartésien d'anneaux d'une structure d'anneau produit. On le définit pour un produit de deux anneaux, mais cela se généralise évidemment pour un produit cartésien d'un nombre fini d'anneaux.

Proposition/Définition 7.2.12 – Anneau produit

- Soit $(A, +, \times)$ et $(B, +, \times)$ deux anneaux. On munit le produit cartésien $A \times B$ des deux opérations :
$$\forall ((a,b),(a',b')) \in (A\times B)^2, \quad (a,b)+(a',b') = (a+a', b+b') \text{ et } (a,b)\times(a',b') = (a\times a', b\times b').$$
- Muni de ces deux lois, $A \times B$ est un anneau, donc les neutres sont $(0_A, 0_B)$ et $(1_A, 1_B)$.
- Si de plus, A et B sont commutatifs, il en est de même de $A \times B$.

Proposition 7.2.13 – Inversibles d'un anneau produit

Soit $(A, +, \times)$ et $(B, +, \times)$. Alors
$$(A \times B)^\times = A^\times \times B^\times$$

II.2 Idéaux

Les anneaux sont des environnements propices à l'arithmétique. En effet, les propriétés liés aux produits permettent, sous certaines conditions, d'envisager des notions de divisibilité.
Dans le cas de l'anneau $\mathbb{Z}$ (et donc de l'arithmétique usuelle sur les entiers), la relation de congruence modulo n est omniprésente. C'est elle qui permet notamment de synthétiser de façon concise toutes les propriétés se rapportant à l'étude des restes de la division euclidienne.
Dans ce cadre, le sous-objet qui nous intéresse dans $\mathbb{Z}$ est l'ensemble $n\mathbb{Z}$ des multiples de n, et les différentes classes de congruence, qui correspondent à des translatés $k+n\mathbb{Z}$ de $n\mathbb{Z}$. Cependant, $n\mathbb{Z}$ n'est pas un sous-anneau de $\mathbb{Z}$, mais est un peu plus riche qu'un sous-groupe additif (car aussi stable par multiplication). Ainsi, même si la notion de sous-anneau n'est pas ininterressante en soi, elle n'est pas la notion appropriée pour les études arithmétiques. Cela motive la définition suivante, introduisant un type de sous-ensembles particuliers d'un anneau. On se limite, conformément au programme, au cas des anneaux commutatifs. Une notion similaire peut être définie également pour des anneaux non commutatifs, mais il convient alors de distinguer les idéaux à droites des idéaux à gauche. Si un sous-ensemble est alors à la fois idéal à gauche et idéal à droite, on dira dans ce cas que c'est un idéal bilatère. Ces distinctions n'ont pas lieu d'être dans la situation qui nous intéresse.

Définition 7.2.14 – Idéal d'un anneau commutatif

Soit A un anneau commutatif. Un idéal de A est une partie $I \subset A$ telle que :

(i) I est un sous-groupe de $(A,+)$

(ii) I est fortement stable par produit : $A \cdot I \subset I$, c'est-à-dire :

$$\forall a \in A, \quad \forall x \in I, \quad ax \in I.$$

On rappelle que le noyau d'un morphisme d'anneau est le noyau du morphisme de groupe additif sous-jacent, c'est-à-dire $\mathrm{Ker}(f) = f^{-1}(\{0_A\})$.

Proposition 7.2.15 – Noyau d'un homomorphisme

Soit A et B deux anneaux, et $f : A \to B$ un morphisme d'anneaux. On suppose A commutatif. Alors $\mathrm{Ker}(f)$ est un idéal de A.

Un autre exemple important d'idéal qui sera à la base de l'arithmétique est l'idéal engendré par un élément.

Proposition 7.2.16 – Idéal engendré par x

Soit A un anneau commutatif, et $x \in A$.

1. $xA = \{xa,\ a \in A\}$ est un idéal.
2. C'est le plus petit idéal de A contenant x.

xA est appelé idéal engendré par x.

Définition 7.2.17 – Idéal principal

Soit A un anneau commutatif et I un idéal de A. On dit que I est un idéal principal, s'il existe x tel que $I = xA$.

Remarque 7.2.18

La plupart des idéaux que nous rencontreront seront principaux, mais il existe des idéaux non principaux.

Définition 7.2.19 – Anneau principal

Un anneau principal est un anneau intègre dont tous les idéaux sont principaux.

En particulier, un anneau principal est commutatif (par définition de l'intégrité).
Comme on le constatera plus loin, c'est le cadre idéal pour faire de l'arithmétique. La raison essentielle en est que la notion de déivisibilité se traduit très facilement sur les idéaux principaux.

Définition 7.2.20 – Divisibilité dans un anneau intègre

Soit A un anneau (commutatif) intègre, et a et b deux éléments de A. On dit que a divise b s'il existe $c \in A$ tel que $ac = b$.

Pr intégrité de A, l'élément c est alors unique.

Proposition 7.2.21 – Caractérisation de la divisibilité par les idéaux

Soit A un anneau commutatif intègre, et a et b deux éléments de A. Les propositions suivantes sont équivalentes: :

(i) a divise b

(ii) $bA \subset aA$.

Enfin, voici deux constructions sur les idéaux, qui seront à la base de caractérisations du pgcd et du ppcm dans $\mathbb{Z}$ ou $\mathbb{K}[X]$ (et pourront même permettre de définir des pgcd et ppcm dans des anneaux principaux quelconques).

Proposition 7.2.22 – Somme d'idéaux

Soit I et J deux idéaux d'un anneau commutatif A. Alors :

1. $I + J$ est un idéal de A
2. C'est le plus petit idéal K tel que $I \subset K$ et $J \subset K$.

Cela s'étend par récurrence immédiate à un nombre fini quelconque d'idéaux.

Dans la suite du chapitre, on s'intéresse à 3 familles d'anneaux en particulier, tout d'abord $\mathbb{Z}$, puis les anneaux $\mathbb{Z}/n\mathbb{Z}$, qui nous fournissent des outils algébriques importants pour l'arithmétique des entiers, et enfin les anneaux de polynômes $\mathbb{K}[X]$, dont l'étude des idéaux permettra également d'établir un certain nombre de propruétés arithmétiques qui seront d'une grande utilité dans l'étude des endomorphismes d'un $\mathbb{K}$-espace vectoriel, et notamment en vue de réduire (diagonaliser par exempl) ces endomorphismes.

II.3 Idéaux de $\mathbb{Z}$ et arithmétique

Théorème 7.2.23 – Les idéaux de $\mathbb{Z}$

$\mathbb{Z}$ est un anneau principal, dont les idéaux sont exactement les $a\mathbb{Z}$, $a \in \mathbb{N}$.

◁ **Éléments de preuve.**
Les idéaux de $\mathbb{Z}$ sont en particulier des sous-groupes, dont on connaît la description ! ▷

Les idéaux permettent de redéfinir, ou de caractériser le PGCD et le PPCM tel qu'ils ont été définis en MPSI.

Définition 7.2.24 – (Re)définition du pgcd d'un nombre fini d'entiers

Soit $n \geqslant 2$ et $a_1, \dots, a_n$ des entiers naturels, non tous nuls, et m un entier naturel. Le pgcd $m = a_1 \wedge a_2 \wedge \cdots \wedge a_n$ de $a_1, \dots, a_n$ est l'unique entier positif engendrant l'idéal $a_1\mathbb{Z} + a_2\mathbb{Z} + \cdots + a_n\mathbb{Z}$.

Ainsi, il est caractérisé par l'égalité

$$m\mathbb{Z} = a_1\mathbb{Z} + a_2\mathbb{Z} + \cdots + a_n\mathbb{Z}.$$

On fait le lien avec la définition de l'année dernière grâce au résultat suivant

Proposition 7.2.25 – Interprétation du pgcd par divisibilité

Le pgcd m de $a_1, \dots, a_n$ est le plus grand entier (au sens de la divisibilité) tel que m divise chaque a_k.
Ainsi, si d est un autre entier divisant chaque a_k, alors $d \mid m$.

Remarque 7.2.26 – Lien avec Bézout

La définition 7.2.24 correspond en fait au théorème de Bézout. Il affirme en effet que si m désigne le pgcd de $a_1, \dots, a_n$:

- il existe $u_1, \dots, u_n$ des entiers relatifs tels que

$$m = a_1u_1 + \cdots + a_nu_n;$$

- réciproquement, si c vérifie

$$c = a_1u_1 + \cdots + a_nu_n,$$

  alors m divise c.

Remarque 7.2.27 – Définition du PGCD et du PPCM dans un anneau principal

On peut définir de la même façon un pgcd (mais non nécessairement unique) dans un anneau principal, ainsi qu'un ppcm, en remplaçant la somme par l'intersection. Par exemple, pour 2 éléments a et b :

- $aA + bA$ est un idéal principal de A, donc de la forme mA. Un élément m vérifiant cela est appelé pgcd de a et b.
- $aA \cap bA$ est un idéal principal de A, donc de la forme MA. Un élément M vérifiant cela est appelé ppcm de a et b.

On retrouvera cette situation dans $\mathbb{K}[X]$.

III Anneau $\mathbb{Z}/n\mathbb{Z}$

III.1 Propriétés d'inversibilité

Nous avons déjà justifié plus haut, au cours d'un exemple que $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un anneau commutatif. Il n'est en général pas intègre. Par exemple, dans $\mathbb{Z}/6\mathbb{Z}$, $\overline{2} \times \overline{3} = \overline{0}$. Ainsi, il n'est pas principal (même si ses idéaux sont principaux, comme on peut le montrer facilement), donc pas un cadre favorable pour l'arithmétique via les idéaux.
En revanche, cet anneau étant étroitement lié à la notion de congruence modulo n, son étude permet tout de même d'apporter des outils algébriques à certains problèmes arithmétiques.
Pour commencer, on décrit les propriétés d'inversibilité.

Proposition 7.3.1 – Inversibles de $\mathbb{Z}/n\mathbb{Z}$

Soit $n \in \mathbb{N}^*$ et $\overline{k} \in \mathbb{Z}/n\mathbb{Z}$. Les propositions suivantes sont équivalentes:

(i) $\overline{k}$ est inversible ;

(ii) il existe u et v dans $\mathbb{Z}$ tels que $uk + vn = 1$;

(iii) $k \wedge n = 1$

◁ **Éléments de preuve.**
Immédiat par traduction des congruences, et par le théorème de Bézout. ▷

Remarque 7.3.2

Le point (ii) nous dit comment trouver un inverse de k modulo n : il suffit de déterminer une relation de Bézout, et de la réduire modulo n.

Corollaire 7.3.3 – Le corps $\mathbb{F}_p$

En particulier, pour $p \in \mathbb{N}^*$, $(\mathbb{Z}/p\mathbb{Z}, +, \times)$ est un corps si et seulement si p est premier.
Ce corps est noté $\mathbb{F}_p$.

◁ **Éléments de preuve.**
Si $p \in \mathbb{P}$, tout élément de $[\![1, p-1]\!]$ est premier avec p. Réciproquement, si $p \notin \mathbb{P}$, un diviseur strict de p n'est pas premier avec p et sa classe n'est pas inversible dans $\mathbb{Z}/p\mathbb{Z}$. ▷

Exemple 7.3.4

Le corps $\mathbb{F}_2$ a seulement deux éléments. Le corps $\mathbb{F}_3$ en a 3. Décrire leurs lois.

III.2 Théorème chinois, et systèmes de congruences

Théorème 7.3.5 – Théorème chinois

Soit m et n dans $\mathbb{N}^*$, tels que $m \wedge n = 1$. Pour $k \in \mathbb{Z}$, on note :
- $\overline{k}$ la classe de k dans $\mathbb{Z}/mn\mathbb{Z}$;
- $\tilde{k}$ la classe de k dans $\mathbb{Z}/m\mathbb{Z}$;
- $\hat{k}$ la classe de k dans $\mathbb{Z}/n\mathbb{Z}$;

L'application $\varphi : \mathbb{Z}/mn\mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ définie par

$$\varphi(\overline{k}) = (\tilde{k}, \hat{k})$$

est bien définie, et est un isomorphisme d'anneaux.
Ceci s'étend sans difficulté au cas d'un plus grand nombre d'entiers 2 à 2 premiers entre eux.

◁ **Éléments de preuve.**
Pour montrer que c'est un isomorphisme, étudier le noyau en se ramenant à des propriétés arithmétiques. ▷

Corollaire 7.3.6 – Unicité modulo $a_1 \cdots a_n$ de la solution d'un système

Soit $a_1, \dots, a_n$ des entiers deux à deux premiers entre eux, et $b_1, \dots, b_n$ des entiers. Alors le système

$$\begin{cases} x \equiv b_1 \ [a_1] \\ \vdots \\ x \equiv b_n \ [a_n] \end{cases}$$

admet une solution, unique modulo $a_1 \cdots a_n$.

En d'autres termes, l'ensemble des solutions est de la forme $x_0 + (a_1 \cdots a_n)\mathbb{Z}$. C'est donc une classe de congruence modulo $a_1 \dots a_n$.

Corollaire 7.3.7 – cas d'un second membre commun

En particulier, si $b_1 = \cdots = b_n = b$, la seule solution modulo $a_1 \cdots a_n$ est b.

Méthode 7.3.8 – Résolution d'un système de congruence

Pour résoudre

$$\begin{cases} x \equiv b_1 \ [a_1] \\ \vdots \\ x \equiv b_n \ [a_n] \end{cases},$$

avec $a_1, \dots, a_n$ deux à deux premiers entre eux, on combine les solutions de

$$\begin{cases} x \equiv \delta_{1,k} \ [a_1] \\ \vdots \\ x \equiv \delta_{n,k} \ [a_n] \end{cases},$$

Pour ces dernières, par exemple pour $k = 1$, le système équivaut à

$$\begin{cases} x \equiv 1 \ [a_1] & x \equiv 0 \ [a_2 \dots a_n] \end{cases}$$

Ce système se résout en écrivant une relation de Bézout entre a_1 et $a_2 \dots a_n$.

Exemple 7.3.9

$$\begin{cases} x \equiv 2 \ [5] \\ x \equiv 6 \ [9] \\ x \equiv -2 \ [11] \end{cases}$$

Lorsque les modules de congruence ne sont pas premiers entre eux, on peut se ramener à cette situation en séparant la contribution de chaque facteur premier, et en comparant les congruences modulo les puissances d'un même facteur premier. Plutôt que de longs discours, voici un exemple :

Exemple 7.3.10

Résoudre les deux systèmes suivants :

1. $\begin{cases} x \equiv 3 \ [42] \\ x \equiv 10 \ [49] \end{cases}$
2. $\begin{cases} x \equiv 3 \ [42] \\ x \equiv 9 \ [49] \end{cases}$

III.3 L'indicatrice d'Euler

Définition 7.3.11 – Indicatrice d'Euler φ

Soit $n \in \mathbb{N}^*$. L'indicatrice d'Euler est définie par

$$\varphi(n) = \mathrm{Card}(\{k \in [\![1, n]\!], \quad k \wedge n = 1\}).$$

D'après la proposition 7.3.1 :

Lemme 7.3.12 – Ordre du groupe des inversibles de $\mathbb{Z}/n\mathbb{Z}$

: Soit $n \in \mathbb{N}^*$. Alors $\mathrm{Card}\left((\mathbb{Z}/n\mathbb{Z})^\times\right) = \varphi(n)$.

Le théorème de Lagrange amène alors :

Théorème 7.3.13 – Théorème d'Euler

Soit $(x, n) \in \mathbb{Z} \times \mathbb{N}^*$ tels que $n \wedge x = 1$. Alors $x^{\varphi(n)} \equiv 1\ [n]$

◁ **Éléments de preuve.**
C'est le théorème de Lagrange dans le groupe $(\mathbb{Z}/n\mathbb{Z})^\times$ ▷

Remarque 7.3.14 – Lien avec le petit théorème de Fermat

Pour $n = p \in \mathbb{P}$, $\varphi(p) = p - 1$, et on retrouve, pour $k \notin p\mathbb{Z}$, $k^{p-1} \equiv 1\ [p]$.
Ainsi, le théorème d'Euler est une généralisation du petit théorème de Fermat.

On cherche maintenant à expliciter l'indicatrice d'Euler $\varphi(n)$, connaissant la décomposition en facteurs premiers de n.

Proposition 7.3.15 – Multiplicativité arithmétique de φ

Soit $(m, n) \in (\mathbb{N}^*)^2$ tel que $m \wedge n = 1$. Alors $\varphi(mn) = \varphi(m)\varphi(n)$.

◁ **Éléments de preuve.**
En passant aux inversibles dans le théorème choinois, sachant qu'un isomorphisme préserve l'inversibilité. ▷

Lemme 7.3.16 – $\varphi(p^k)$

Soit $p \in \mathbb{P}$ et $k \in \mathbb{N}^*$. Alors $\varphi(p^k) = p^{k-1}(p - 1)$.

◁ **Éléments de preuve.**
Dénombrement direct, sachant que k n'est pas premier avec $p \in \mathbb{P}$ si et seulement s'il est multiple de p. ▷

Théorème 7.3.17 – Expression de $\varphi(n)$

Soit $n =\in \mathbb{N}^*$, et $n = p_1^{a_1} \dots p_m^{a_m}$ sa décomposition en facteurs premiers, avec $a_i > 0$. Alors

$$\varphi(n) = \prod_{\ell=1}^{m} p_\ell^{a_\ell - 1}(p_\ell - 1) = n \prod_{\ell=1}^{m} \left(1 - \frac{1}{p_\ell}\right).$$

IV Arithmétique dans l'anneau $\mathbb{K}[X]$

Conformément au programme, soit $\mathbb{K}$ un sous-corps de $\mathbb{C}$ (mais tous les résultats de cette section restent valide si $\mathbb{K}$ est un corps quelconque).

IV.1 Idéaux de $\mathbb{K}[X]$

Tout se passe un peu pareil dans $\mathbb{K}[X]$ que dans $\mathbb{Z}$, du fait de la description des idéaux de $\mathbb{K}[X]$.

Théorème 7.4.1 – Description des idéaux de $\mathbb{K}[X]$

Soit $\mathbb{K}$ un corps. Alors :

1. $\mathbb{K}[X]$ est un anneau principal.
2. De plus, deux polynômes P et Q engendrent le même idéal si et seulement s'il existe un $\lambda \in \mathbb{K}^*$ tel que $Q = \lambda P$.
3. En particulier, tout idéal non nul I de $\mathbb{K}[X]$ s'écrit de manière unique sous la forme $I = P \cdot \mathbb{K}[X]$, où P est un polynôme unitaire.

◁ **Éléments de preuve.**
S'aider de la division euclidienne, comme pour montrer que $\mathbb{Z}$ est principal. C'est une propriété général de tout anneau muni d'une division euclidienne (anneau euclidien). La deuxième partie provient de la description des inversibles de $\mathbb{K}[X]$. ▷

Définition 7.4.2 – Divisibilité dans $\mathbb{K}[X]$

Soit A et B deux polynômes de $\mathbb{K}[X]$. On dit que B divise A s'il existe un polynôme Q tel que $A = BQ$. Inversement, on dit que A est un multiple de B.

Ainsi, B divise A si et seulement si le reste de la division euclidienne de A par B est nul.
Comme dans $\mathbb{Z}$, on a la caractérisation suivante :

Proposition 7.4.3 – Caractérisation en termes d'idéaux

Soit A et B deux polynômes de $\mathbb{K}[X]$. Alors A divise B si et seulement si $B \in (A)$, ou encore si et seulement si $(B) \subset (A)$.

On peut alors caracrétiser (ou définir) le pgcd de la même manière que dans $\mathbb{Z}$:

Définition 7.4.4 – (Re)définition du PGCD de n polynômes

Soit $n \geqslant 2$ et $P_1, \dots, P_n$ des polynômes non tous nuls de $\mathbb{K}[X]$.
Le PGCD Q de $P_1, \dots, P_n$, noté $Q = P_1 \wedge P_2 \wedge \cdots \wedge P_n$, est l'unique polynôme unitaire tel que

$$Q \cdot \mathbb{K}[X] = P_1 \cdot \mathbb{K}[X] + \cdots + P_n \cdot \mathbb{K}[X].$$

Comme pour les entiers, on retrouve la notion définie en MSPI via le résultat suivant :

Proposition 7.4.5 – Caractérisation du PGCD par divisibilité

Soit $n \geqslant 2$ et $P_1, \dots, P_n$ des polynômes non tous nuls de $\mathbb{K}[X]$. Alors

- $Q = P_1 \wedge \cdots \wedge P_n$ est l'unique polynôme unitaire de plus grand degré divisant chacun des P_i ;
- Tout autre polynôme D divisant chaque P_i vérifie $D \mid Q$.

IV.2 Décomposition en facteurs irréductibles

Étendons rapidement au cas où $\mathbb{K}$ est un sous-corps quelconque de $\mathbb{C}$ les propriétés arithémtiques vues pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ en MPSI.

Définition 7.4.6 – Polynôme irréductible

Un polynôme non constant P de $\mathbb{K}[X]$ est irréductible si et seulement s'il n'est, à une constante multiplicative non nulle près, divisible que par lui-même et par 1.

Exemples 7.4.7

1. Les polynômes $X - \lambda$ sont irréductibles ($\lambda \in \mathbb{K}$)
2. Dans $\mathbb{R}[X]$, tout polynôme $aX^2 + bX + c$ tel que $\Delta < 0$ est irréductible.
3. Ces polynômes ne sont pas irréductibles dans $\mathbb{C}[X]$.

On rappelle que :

Théorème 7.4.8 – Polynômes irréductibles de $\mathbb{C}[X]$ et $\mathbb{R}[X]$

1. Les polynômes irréductibles de $\mathbb{C}[X]$ sont exactement les polynômes de degré 1 (théorème de d'Alembert-Gauss)
2. Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 à discriminant strictement négatif.

Lemme 7.4.9 – Existence d'un diviseur irréductible

Tout polynôme non constant $P \in \mathbb{K}[X]$ admet un diviseur irréductible.

◁ **Éléments de preuve.**
Considérer un diviseur non constant de degré minimal. ▷

Lemme 7.4.10 – Euclide

Soit A et B deux polynômes de $\mathbb{K}[X]$ et P un polynôme irréductible. Alors si P divise AB, P divise A ou P divise B.

Le lemme d'Euclide se généralise facilement à davantage de facteur, par récurrence immédiate.

◁ **Éléments de preuve.**
Si P ne divise pas A, $P\mathbb{K}[X] + A\mathbb{K}[X] = \mathbb{K}[X]$. Multiplier par B pour obtenir $B\mathbb{K}[X] \subset P\mathbb{K}[X]$. ▷

Enfin, voici l'analogue du théorème de la décomposition primaire :

Théorème 7.4.11 – Décomposition en facteurs irréductibles

Soit P un polynôme non nul de $\mathbb{K}[X]$. Il existe un unique élément $\lambda \in \mathbb{K}^*$ et des polynômes irréductibles unitaires $P_1, \dots, P_k$, uniques à permutation près, tels que

$$P = \lambda P_1 \cdots P_k.$$

◁ **Éléments de preuve.**
Par récurrence, en divisant par un facteur irréductible unitaire. Pour l'unicité, aussi par récurrence, en remarquant qu'un facteur irréductible unitaire de P se retrouve dans toute décomposition, d'après le lemme d'Euclide. ▷

8

Compléments d'algèbre linéaire

On admet que toute l'algèbre linéaire vue en MPSI reste valide lorsque $\mathbb{K}$, le corps de base, est un corps quelconque. En particulier, les coefficients peuvent, par exemple, être pris sur $\mathbb{K} = \mathbb{F}_p$, $p \in \mathbb{P}$. Cependant, lorsque des polynômes interviendront, (notamment dans la troisième partie de ce chapitre), il convient de faire un peu plus attention, car des choses particulières peuvent se produire. Par exemple, les règles sur les degrés peuvent ne pas être les mêmes. Par exemple, dans $\mathbb{F}_p[X]$, la dérivée de X^p n'est pas de degré $p-1$, et à peu près pour cette raison, les polynômes de degré $p-1$ n'admettent pas de primitive. On observe par la même occasion que pour ce polynôme X^p, la caractérisation de la multiplicité de la racine 0 par les dérivées successives n'est pas valide.

Pour éviter ce type de problème, dès lors que des polynômes interviendront dans notre étude, nous supposerons que $\mathbb{K}$ est un sous-corps de $\mathbb{C}$, même si un grand nombre de résultats associés restent valides dans un contexte plus général.

Nous ne reprenons pas le cours de MPSI. Nous nous contentons d'en rappeler ci-dessous les grandes lignes, à bien maîtriser :

- Notion d'espace vectoriel, de sous-espace vectoriel ;
- Notion de famille libre, de famille génératrice, de base ; une famille libre maximale est une base ;
- Notion de somme de 2 espaces, de somme directe de 2 espaces. Lien avec les bases.
- Applications linéaires, détermination d'une application linéaire sur les vecteurs d'une base, endomorphisme, isomorphisme, automorphisme
- Projecteurs et symétries, caractérisation algébrique ($p^2 = p$, $s^2 = \mathrm{id}$), caractérisation géométrique.
- Sous-espaces affines d'un espace vectoriel
- Théorie de la dimension : espace de dimension finie (par existence d'une famille génératrice finie), théorème de la dimension, dimension d'une somme directe, formule de Grassmann
- Applications linéaires en dimension finie, dont surtout, le théorème du rang et la caractérisation de la bijectivité par l'injectivité ou la surjectivité en cas d'égalité des dimensions.
- Représentation matricielle d'une AL par choix de 2 bases. Cas d'un endomorphisme (avec la même base au départ et à l'arrivée). Propriétés de composition, d'évaluation.
- Rang d'une matrice, version matricielle du théorème du rang, calcul du rang d'une matrice par méthode du pivot.
- Changements de base. Matrices équivalentes. Toute matrice est équivalente à une $J_{n,p,r}$; deux matrices sont équivalentes ssi elles ont même rang.
- Déterminant d'une application linéaire dans une base, déterminant d'une matrice. Lien entre les deux.
- Méthodes calculatoires pour les déterminants : pivot, développement suivant une ligne ou une colonne.
- Exemple important : déterminant de Vandermonde.

I Sommes directes et blocs

I.1 Somme directe de n sous-espaces vectoriels

On définit la somme de n espaces vectoriels de façon globale, plutôt qu'itérative :

Définition 8.1.1 – Somme d'un nombre fini de sous-espaces

Soit $n \geqslant 1$, et $F_1, \dots, F_n$ des sous-espaces vectoriel de E. On définit

$$\sum_{i=1}^{n} F_i = F_1 + \cdots + F_n = \left\{ \sum_{i=1}^{n} x_i \mid x_1 \in F_1, \dots, x_n \in F_n \right\}.$$

Théorème 8.1.2 – Minimalité de la somme

$\sum_{i=1}^{n} F_i$ est le plus petit sous-espace vectoriel de E incluant tous les E_i. En d'autres termes,

$$\sum_{i=1}^{n} F_i = \operatorname{Vect}\left(\bigcup_{i=1}^{n} F_i \right).$$

◁ **Éléments de preuve.**

- C'est un sev, par vérifications simples
- Il inclut tous les F_i
- Par stabilité par somme, si un sev G inclut tous les F_i, il inclut aussi leur somme.

▷

Proposition/Définition 8.1.3 – Somme directe

Soit $F_1, \dots, F_n$ des sev de E. Les propositions suivantes sont équivalentes:

(i) Pour tout $x \in F_1 + \cdots + F_n$, il existe un unique n-uplet $(x_1, \dots, x_n) \in F_1 \times \cdots \times F_n$ tel que $x = x_1 + \cdots + x_n$

(ii) Pour tout $(x_1, \dots, x_n) \in F_1 \times \cdots \times F_n$, tel que $(x_1, \dots, x_n) \neq (0, \dots, 0)$, $x_1 + \cdots + x_n \neq 0$.

(iii) Pour tout $i \in [\![1, n]\!]$, $F_i \cap \sum_{j \in [\![1,n]\!] \setminus \{i\}} F_j = \{0\}$

(iv) Pour tout $i \in [\![2, n]\!]$, $F_i \cap \sum_{j=1}^{i-1} F_j = \{0\}$

(v) Pour tout $i \in [\![2, n]\!]$, la somme de F_i et de $\sum_{j=1}^{i-1} F_j$ est directe.

Dans ce cas, on dit que la somme $F_1 + \cdots + F_n$ est directe, et pour indiquer cette propriété, on note :

$$F_1 + \cdots + F_n = F_1 \oplus \cdots \oplus F_n = \bigoplus_{i=1}^{n} F_i$$

Remarque 8.1.4

La propriété (v) montre que la somme globale est directe, si et seulement si à chaque étape, le terme qu'on ajoute est en somme directe avec les précédents. Combiné aux propriétés d'associativité, cela permet d'écrire :

$$\bigoplus_{i=1}^{n} F_i = (((F_1 \oplus F_2) \oplus F_3) \oplus \cdots \oplus F_{n-1}) \oplus F_n.$$

Avertissement 8.1.5

Attention à l'erreur courante consistant à caractériser la somme directe par le fait que les intersections deux à deux sont réduites à $\{0\}$ (autrement dit les E_i sont 2 à 2 en somme directe.
C'est clairement insuffisant, comme le montre l'exemple de 3 droites vectorielles non confondues dans le plan.

Proposition 8.1.6 – Dimension d'une somme

Soit $F_1, \dots, F_n$ des sous-espaces de dimension finie de E. Alors :

1. $F_1 + \cdots + F_n$ est aussi de dimension finie ;
2. $\dim(F_1 + \cdots + F_n) \leqslant \sum_{i=1}^{n} \dim(F_i)$;
3. si la somme est directe, $\dim\left(\bigoplus_{i=1}^{n} F_i\right) = \sum_{i=1}^{n} \dim(F_i)$.

Notation 8.1.7 – Concaténation

Pour $X = (x_1, \dots, x_k) \in E^k$ et $Y = (y_1, \dots, y_\ell) \in E^\ell$, on note

$$X \smile Y = (x_1, \dots, x_k, y_1, \dots, y_\ell)$$

la concaténation du k-uplet X et du ℓ-uplet Y. Il s'agit donc d'un $k + \ell$-uplet.
C'est une notation à redéfinir si vous l'utilisez.

Proposition/Définition 8.1.8 – Base adaptée à une décomposition en somme directe

1. Soit E un espace vectoriel de dimension finie, tel que $E = \bigoplus_{i=1}^{n} E_i$, et tout $i \in [\![1, n]\!]$, $\mathcal{B}_i$ une base de E_i Alors $\mathcal{B} = \mathcal{B}_1 \smile \cdots \smile \mathcal{B}_n$ est une base de E.
2. Une base $\mathcal{B}$ de E est dite adaptée à la décomposition en somme directe $E = \bigoplus_{i=1}^{n} E_i$ si elle est obtenue de la sorte, par concaténation de bases des E_i.

Le point 1 montre qu'il existe toujours une base adaptée à une décomposition en somme directe.

Proposition 8.1.9 – Projecteurs associés à une décomposition en somme directe

Soit E un espace vectoriel tel que $E = \bigoplus_{i=1}^{n} E_i$. On définit pour tout $i \in [\![1, n]\!]$, $\pi_i : E \to E$ par

$$\pi_i : x_1 + \cdots + x_n \mapsto x_i.$$

1. Pour tout $i \in [\![1, n]\!]$, π_i est un projecteur. Plus précisément, c'est le projecteur sur E_i parallèlement à la somme des autres E_k.
2. Les π_i vérifient :

$$\sum_{i=1}^{n} \pi_i = \mathrm{id}_E \qquad \text{et} \qquad \forall (i, j) \in [\![1, n]\!]^2 \quad \pi_i \circ \pi_j = \delta_{i,j}.$$

On peut montrer que réciproquement, si $(\pi_1, \dots, \pi_n)$ sont des projecteurs vérifiant le point 2, leurs images forment une déomposition de E en somme directe dont ils sont les projecteurs associés.

I.2 Détermination d'une application linéaire

Vous avez déjà vu l'année dernière qu'une application linéaire se détermine entièrement et sans ambiguité par l'image des vecteurs d'une base. C'est cette représentation qui donne lieu ensuite aux interprétations matricielles : la matrice d'une application linéaire u détermine entièrement u, justement car elle décrit l'image d'une base.

Nous allons donner une version un peu plus générale de ce résultat, en montrant qu'une application linéaire est déterminée par ses restrictions à des sous-espaces formant une décomposition de E en somme directe. En considérant une décomposition de E en somme directe de droites engendrées par les vecteurs d'une base, on retrouve plus ou moins le théorème de détermination d'une AL sur une base, pouvant ainsi considérer que nous avons généralisé ce résultat.

Proposition 8.1.10 – Détermination d'une AL sur une décomposition en somme directe

Soit E et F deux $\mathbb{K}$-ev. Soit $E = \bigoplus_{j=1}^{n} E_j$ une décomposition de E en somme directe, et pour tout $j \in [\![1, n]\!]$, $u_j \in \mathcal{L}(E_j, F)$. Alors il existe une et une seule application linéaire $u \in \mathcal{L}(E, F)$ telle que pour tout $j \in [\![1, n]\!]$, $u_{|E_j} = u_j$.

◁ **Éléments de preuve.**

Une analyse synthèse sans difficulté. ▷

Proposition 8.1.11 – Interprétation matricielle de la détermination sur une $\oplus$

Soit $E = \bigoplus_{i=j}^{n} E_j$ et F deux espaces vectoriels de dimension finie. On note, pour $i \in [\![1, n]\!]$, $p_j = \dim E_j$, $p = \dim E$ et $q = \dim F$. On se donne $\mathcal{B}_1, \dots, \mathcal{B}_n$ des bases de $E_1, \dots, E_n$, $\mathcal{C}$ une base de F, et on définit $\mathcal{B} = \mathcal{B}_1 \smile \cdots \smile \mathcal{B}_n$, base adaptée à la décomposition de E en somme directe.

1. Soit pour tout $j \in [\![1, n]\!]$, $u_j \in \mathcal{L}(E_j, F)$, et $A_j = \mathrm{Mat}_{\mathcal{B}_j, \mathcal{C}}(u_j) \in \mathcal{M}_{q, p_j}(\mathbb{K})$. Alors l'unique application linéaire $u \in \mathcal{L}(E, F)$ coïncidant avec chacun des u_j est l'application déterminée par la matrice
$$\mathrm{Mat}_{\mathcal{B}, \mathcal{C}}(u) = \begin{pmatrix} \mathrm{Mat}_{\mathcal{B}_1, \mathcal{C}}(u_1) & \mathrm{Mat}_{\mathcal{B}_2, \mathcal{C}}(u_2) & \cdots & \mathrm{Mat}_{\mathcal{B}_n, \mathcal{C}}(u_n) \end{pmatrix} \in \mathcal{M}_{q,p}(\mathbb{K}),$$
c'est-à-dire la matrice obtenue en juxtaposant dans une même grande matrice les matrices des u_i dans l'ordre.
2. Réciproquement, pour $u \in \mathcal{L}(E, F)$, et une décomposition en blocs
$$\mathrm{Mat}_{\mathcal{B}, \mathcal{C}}(u) = \begin{pmatrix} A_1 & A_2 & \cdots & A_n \end{pmatrix},$$
où pour tout $j \in [\![1, n]\!]$, $A_j \in \mathcal{M}_{q, p_j}$, alors
$$A_j = \mathrm{Mat}_{\mathcal{B}_j, \mathcal{F}}(u_{|E_j}).$$
3. En particulier, pour $u \in \mathcal{L}(E, F)$, avec les notations précédentes,
$$\mathrm{Mat}_{\mathcal{B}, \mathcal{C}}(u \circ \pi_j) = \begin{pmatrix} 0_{q,p_1} & \cdots & 0_{q,p_{j-1}} & A_j & 0_{q,p_{j+1}} & \cdots & 0_{q,p_n} \end{pmatrix}.$$

On s'intéresse maintenant au problème dual, lorsque $F = \bigoplus_{i=1}^{m} F_i$ admet une décomposition en somme directe. On note $\rho_1, \dots, \rho_m$ les projections associées, et ${}_j u$ la corestriction à F_j de $\rho_i \circ u$. Ainsi, pour tout $x \in E$, ${}_j u(x)$ est la composante sur F_j de la décomposition de $u(x)$ dans la somme directe $\bigoplus_{i=1}^{m} F_i$. On a alors ${}_j u \in \mathcal{L}(E, F_j)$.

Proposition 8.1.12 – Interprétation matricielle d'une décomposition en $\oplus$ à l'arrivée

Soit E et F deux espaces vectoriels de dimension finie, et $F = \bigoplus_{i=1}^{m} F_i$ une décomposition de F en somme directe. Soit pour tout $i \in [\![1, m]\!]$ une base $\mathcal{C}_i$ de F_i, et $\mathcal{C} = \mathcal{C}_1 \smile \cdots \smile \mathcal{C}_m$ la base de F correspondante, adaptée à la décomposition. Soit $\mathcal{B}$ une base de E. Alors

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u) = \begin{pmatrix} \mathcal{M}_{\mathcal{B},\mathcal{C}_1}({}_1u) \\ \vdots \\ \mathcal{M}_{\mathcal{B},\mathcal{C}_m}({}_mu) \end{pmatrix}.$$

De plus, on a alors, pour tout $i \in [\![1, m]\!]$

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\rho_i \circ u) = \begin{pmatrix} 0_{q_1,p} \\ \vdots \\ 0_{q_{i-1},p} \\ \mathcal{M}_{\mathcal{B},\mathcal{C}_i}({}_iu) \\ 0_{q_{i+1},p} \\ \vdots \\ 0_{q_m,p} \end{pmatrix}$$

En combinant ces deux représentations on peut avoir une représentation d'un endomorphisme u relativement à des bases adaptées à deux décompositions en somme directe, l'une de l'espace de départ $E = \bigoplus_{j=1}^{n} E_j$, l'autre de l'espace d'arrivée $E = \bigoplus_{i=1}^{m} F_i$, les deux espaces E et F étant supposés de dimension finie. Pour $u \in \mathcal{L}(E, F)$, pour tout $j \in [\![1, n]\!]$, u_j désigne comme avant la restriction de u à E_j, et pour tout $i \in [\![1, m]\!]$, on peut alors considérer ${}_iu_j$ la corestriction de $\rho_i \circ u_j$ à F_i. Ainsi, ${}_iu_j \in \mathcal{L}(E_j, F_i)$.

Théorème 8.1.13 – Interprétation d'une écriture matricielle par blocs

Soit E et F deux espaces vectoriels de dimension finie

1. On suppose que
$$E = \bigoplus_{j=1}^{n} E_j \qquad \text{et} \qquad F = \bigoplus_{i=1}^{m} F_i.$$
Soit, pour tout $j \in [\![1, n]\!]$, $\mathcal{B}_j$ une base de E_j et pour tout $i \in [\![1, m]\!]$ une base de F_i. On note $\mathcal{B} = \mathcal{B}_1 \smile \cdots \smile \mathcal{B}_n$ et $\mathcal{C} = \mathcal{C}_1 \smile \cdots \smile \mathcal{C}_m$ les bases adaptées correspondantes de E et F. Alors on obtient la représentation par blocs suivante :
$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u) = \begin{pmatrix} \mathrm{Mat}_{\mathcal{B}_1,\mathcal{C}_1}({}_1u_1) & \cdots & \mathrm{Mat}_{\mathcal{B}_n,\mathcal{C}_1}({}_1u_j) \\ \vdots & & \vdots \\ \mathrm{Mat}_{\mathcal{B}_1,\mathcal{C}_m}({}_mu_1) & \ldots & \mathrm{Mat}_{\mathcal{B}_n,\mathcal{C}_m}({}_mu_n) \end{pmatrix},$$

2. On a alors en particulier, en notant $p_j = \dim E_j$ et $q_i = \dim F_i$,
$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\rho_j \circ u \circ \pi_i) = \begin{pmatrix} 0_{q_1,p_1} & \cdots & 0_{q_1,p_n} \\ \vdots & \mathrm{Mat}_{\mathcal{B}_j,\mathcal{C}_i}({}_iu_j) & \vdots \\ 0_{q_m,p_1} & \cdots & 0_{q_m,p_n} \end{pmatrix},$$
l'unique bloc non nul étant en position (i, j).

3. Réciproquement, si $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$ se décompose en blocs tels que

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix},$$

tels que :
- à i fixé, les $A_{i,j}$, $j \in [\![1,n]\!]$ ont toutes même nombre de ligne ;
- à j fixé, les $A_{i,j}$, $i \in [\![1,m]\!]$ ont toutes même nombre de colonnes ;

alors, en regroupant convenablement les vecteurs de la base $\mathcal{B}$, il existe des décompositions en somme directe de E et F dont $\mathcal{B}$ et $\mathcal{C}$ sont des bases adaptées, telles que, avec les notations préédentes,

$$\forall (i,j) \in [\![1,n]\!] \times [\![1,m]\!], \quad A_{i,j} = \mathrm{Mat}_{\mathcal{B}_j,\mathcal{C}_i}({}_iu_j).$$

Définition 8.1.14 – Décomposition en blocs d'une matrice

On appellera décomposition (ou découpage) par blocs d'une matrice A une représentation de A dans laquelle on a regroupé certaines lignes consécutives ou colonnes consécutives de sorte à former des sous-matrices, vérifiant les conditions du 3 du théorème 8.1.13

Corollaire 8.1.15 – Décomposition de E stable par u

Soit E un espace vectoriel de dimension finie, et une décomposition en somme directe :

$$E = \bigoplus_{j=1}^{n} E_j.$$

Soit pour tout $i \in [\![1,n]\!]$, $\mathcal{B}_i$ une base de E_i, et $\mathcal{B} = \mathcal{B}_1 \smile \cdots \smile \mathcal{B}_n$ une base adaptée.
Soit $u \in \mathcal{L}(E)$ tel que pour tout $i \in [\![1,n]\!]$, E_i est stable par u. Alors

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \mathrm{Mat}_{\mathcal{B}_1}(u_{E_1}) & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathrm{Mat}_{\mathcal{B}_n}(u_{E_n}) \end{pmatrix},$$

où $u_{E_i} = {}_iu_i$ est l'endomorphisme de $\mathcal{L}(E_i)$ induit par u.
Une telle matrice est appelée matrice diagonale par blocs.

I.3 Calcul matriciel par blocs

Les descriptions précédentes vont permettre d'établir assez facilement (disons que le plus dur est fait...) les règles de calcul matriciel par blocs.

Terminologie 8.1.16 – Découpages par blocs de tailles compatibles

Soit $A \in \mathcal{M}_{q,r}(\mathbb{K})$ et $B \in \mathcal{M}_{r,s}(\mathbb{K})$ deux matrices, munis de découpages par blocs

$$A = (A_{i,j})_{(i,j)\in[\![1,m]\!]\times[\![1,n]\!]} \qquad \text{et} \qquad B = (B_{j,k})_{(j,k)\in[\![1,n]\!]\times[\![1,p]\!]}.$$

On dira que les deux découpages par blocs ont des tailles compatibles si pour tout $j \in [\![1,n]\!]$, le nombre de colonnes des $A_{i,j}$ est égal au nombre de lignes des $B_{j,k}$.

Remarques 8.1.17

1. Attention, on ne peut pas permuter A et B dans cette définition.
2. Cette condition permet de considérer, pour tout $j \in [\![1, n]\!]$, le produit $A_{i,j}B_{j,k}$. De plus, le format de cette matrice ne dépend pas de j. Ainsi, la matrice suivante est bien définie :
$$\sum_{j=1}^{n} A_{i,j}B_{j,k}.$$
3. D'un point de vue vectoriel, si A et B sont interprétés comme des matrices d'applications linéaires $u \in \mathcal{L}(F, G)$ et $v \in \mathcal{L}(E, F)$ respectivement, relativement à des bases $\mathcal{B}$, $\mathcal{C}$ et $\mathcal{D}$, les découpages par blocs définissent des décompositions en somme directe de E, F et G. La compatibilité des tailles des blocs dit simplement que les blocs de A et B définissent la même décomposition en somme directe de F.

Théorème 8.1.18 – Produit matriciel par blocs

Soit $A \in \mathcal{M}_{q,r}(\mathbb{K})$ et $B \in \mathcal{M}_{r,s}(\mathbb{K})$ deux matrices, munis de découpages par blocs
$$A = (A_{i,j})_{(i,j)\in[\![1,m]\!]\times[\![1,n]\!]} \qquad \text{et} \qquad B = (B_{j,k})_{(j,k)\in[\![1,n]\!]\times[\![1,p]\!]}.$$
de tailles compatibles. Alors le produit AB peut se calculer par blocs :
$$AB = \left(\sum_{j=1}^{n} A_{i,j}B_{j,k}\right)_{(i,k)\in[\![1,m]\!]\times[\![1,p]\!]}.$$

◁ **Éléments de preuve.**

La démonstration n'est pas exigible.

Considérer $u \in \mathcal{L}(F, G)$ et $v \in \mathcal{L}(E, F)$ des endomorphismes associés à A et B, avec des décompositions en somme directe associées de E, F et G, de projections respectives π_k, ρ_j et σ_i. Décomposer
$$\pi_k \circ u \circ v \circ \sigma_i = \sum_{j} \pi_k \circ u \circ \rho_j \circ \rho_j \circ v \circ \sigma_i$$
et utiliser les résultats du paragraphe précédent. ▷

Ainsi, le produit matriciel par blocs suit les mêmes règles que le produit matriciel sur les coefficients, en remplaçant les coefficients par des blocs matriciels. Il faut juste pour cela que les tailles soient compatibles pour former les produits.

Corollaire 8.1.19 – Puissances de matrices triangulaires par blocs

Soit $A \in \mathcal{M}_p(\mathbb{K})$ découpée en blocs $A = (A_{i,j})_{1\leqslant i,j\leqslant n}$, de tailles compatibles pour former le produit $A \times A$ (autrement dit, les regroupements de lignes ont la même taille que les regroupements de colonnes, ce qui équivaut à dire que les blocs diagonaux sont carrés).

1. On dit que A est triangulaire (supérieure) par blocs si pour tout $i > j$, $A_{i,j} = 0$, donc
$$A = \begin{pmatrix} A_{1,1} & \cdots & \cdots & A_{1,n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & A_{q,q} \end{pmatrix}.$$
Alors pour tout $n \in \mathbb{N}$, A^n est aussi triangulaire par blocs, avec des blocs de même taille que

A, et

$$A^n = \begin{pmatrix} A_{1,1}^n & \bullet & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bullet \\ 0 & \cdots & 0 & A_{q,q}^n \end{pmatrix}.$$

2. On dit que A est diagonale par blocs si pour tout $i \neq j$, $A_{i,j} = 0$, donc

$$A = \begin{pmatrix} A_{1,1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_{q,q} \end{pmatrix}.$$

Alors pour tout $n \in \mathbb{N}$, A^n est aussi diagonale par blocs, avec des blocs de même taille que A, et

$$A^n = \begin{pmatrix} A_{1,1}^n & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_{q,q}^n \end{pmatrix}.$$

◁ **Éléments de preuve.**

Comme pour les matrices exprimées avec des coefficients scalaires. Inutile de refaire. ▷

Exemple 8.1.20

Calcul des puissances successives de $\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$

I.4 Déterminants par blocs

Le but de ce paragraphe est de donner quelques outils supplémentaires sur le calcul des déterminants, notamment lorsqu'on dispose d'un découpage par blocs de la matrice. En règle générale, on ne peut pas exploiter un découpage quelconque. Mais certaines manipulations sont permises.

Définition 8.1.21 – Transvection par blocs

Soit $A = (A_{i,j})_{(i,j)\in[\![1,m]\!]\times[\![1,n]\!]}$ une représentation de A par blocs matriciels $A_{i,j} \in \mathcal{M}_{m_i,n_j}(\mathbb{K})$. Soit $T \in \mathcal{M}_{m_i,m_k}(\mathbb{K})$. Pour $i \neq k$, la transvection $L_i \leftarrow L_i + TL_k$ consiste à remplacer les $A_{i,j}$ de la ligne L_i par $A_{i,j} + TA_{k,j}$.

La ligne L_i correspond évidemment à la i-ème ligne (matricielle) du découpage par blocs.

Proposition 8.1.22 – Interprétation matricielle d'une transvection par blocs

la transvection par blocs $L_i \leftarrow L_i + TL_k$ correspond à la multiplication à gauche par la matrice

$$E_{i,k,T} = \begin{array}{cc} \begin{pmatrix} I & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & T & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & I \end{pmatrix} & \begin{array}{c} \\ i \\ \\ \\ \\ \\ \end{array} \\ \begin{array}{cccccc} & & & k & & \end{array} & \end{array},$$

les blocs diagonaux étant triangulaires, et les blocs étant de taille compatible avec ceux de A pour effectuer le produit.

◁ **Éléments de preuve.**
C'est la même chose que dans le cas scalaire, à adapter grâce aux règles de produit matriciel par blocs. ▷

Corollaire 8.1.23 – Invariance du déterminant par transvection par blocs

Soit B la matrice obtenue de A par transvection $L_i \leqslant L_i + TL_k$. Alors $\det(A) = \det(B)$.

◁ **Éléments de preuve.**
En effet, la matrice $E_{i,k,T}$ est triangulaire (vraiment, pas seulement par blocs), de coefficients diagonaux tous égaux à 1. ▷

Lemme 8.1.24

1. $\det \begin{pmatrix} I & 0 \\ C & B \end{pmatrix} = \det(B)$
2. $\det \begin{pmatrix} I & C \\ 0 & B \end{pmatrix} = \det(B)$

◁ **Éléments de preuve.**

1. Une transvection fait sauter C, puis faire des développements suivant des colonnes pour se débarrasser du I.
2. Transposer.

▷

Théorème 8.1.25 – Déterminant d'une matrice triangulaire par blocs

Soit T une matrice triangulaire par blocs, c'est-à-dire de la forme

$$T = \begin{pmatrix} A_1 & \bullet & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bullet \\ 0 & \cdots & 0 & A_k \end{pmatrix}$$

où les A_k sont des matrices carrées. Alors

$$\det(T) = \det(A_1) \cdots \det(A_k).$$

◁ **Éléments de preuve.**
Par récurrence, se ramener au cas $n = 2$, pour lequel on peut décomposer

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \begin{pmatrix} I & C \\ 0 & B \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix}$$

puis utiliser le lemme 8.1.24. ▷

Exemple 8.1.26

Exprimer $\det \begin{pmatrix} A & B \\ B & A \end{pmatrix}$ en fonction de $\det(A - B)$ et $\det(A + B)$.

II Algèbres

II.1 Définitions générales

Soit $\mathbb{K}$ un corps quelconque.

Définition 8.2.1 – Algèbre

Soit $(A, +, \cdot, \times)$ un ensemble muni de 2 l.c.i. $+$ et $\times$ et d'une l.c.e. $\cdot$ sur le corps $\mathbb{K}$. On dit que A est une algèbre sur $\mathbb{K}$ si :

(i) $(A, +, \cdot)$ est un espace vectoriel sur $\mathbb{K}$

(ii) $(A, +, \times)$ est un anneau

(iii) Pour tout $(\lambda, x, y) \in \mathbb{K} \times A \times A$, $(\lambda \cdot x) \times y = \lambda \cdot (x \times y) = x \times (\lambda \cdot x)$.

Remarque 8.2.2

1. En particulier, une algèbre est unitaire, dans le sens où elle possède un neutre multiplicatif. Dans certains ouvrages plus généraux, ce n'est pas toujours le cas. Se méfier.
2. La définition est redondante, puisque certaines propriétés de $+$ se retrouvent à la fois dans la structure d'anneau et dans la structure d'espace vectoriel
3. Pour éviter ces redondances, on peut définir une algèbre comme un $\mathbb{K}$-ev enrichi d'une multiplication $\times$ associative, unitaire, et bilinéaire. La bilinéarité englobe à la fois les propriétés de distributivité et la propriété (iii).
4. Les signes opératoires $\times$ et $\cdot$ sont souvent omis. Il n'y a pas de confusion possible entre les deux, puisque l'une des deux lois implique des scalaires et pas l'autre. L'examen de la nature des objets en jeu permet donc de lever toute ambiguité.

Exemples 8.2.3

1. $(\mathbb{K}, +, \cdot, \cdot)$ est une $\mathbb{K}$-algèbre (le produit externe et le produit interne sont ici égaux).
2. $\mathbb{K}[X]$ est une $\mathbb{K}$-algèbre.
3. $\mathbb{K}[X, Y]$ (combinaison linéaire de monômes $X^i Y^j$) aussi.
4. $\mathcal{M}_n(\mathbb{K})$ est une $\mathbb{K}$-algèbre
5. Si E est un $\mathbb{K}$-ev, $\mathcal{L}(E)$ est une $\mathbb{K}$-algèbre.
6. $\mathbb{K}^X = \mathcal{F}(X, \mathbb{K})$ est une $\mathbb{K}$-algèbre.
7. $\mathbb{C}$ est une $\mathbb{R}$-algèbre.

8. Soit $a \in \mathbb{R}$, et $\mathbb{Q}[a] = \{P(a),\ P \in \mathbb{Q}[X]\}$. Alors $\mathbb{Q}[a]$ est une $\mathbb{Q}$-algèbre.

Définition 8.2.4 – Sous-algèbre

Une sous-algèbre B d'une algèbre A est un sous-ensemble $B \subset A$ tel que :

(i) B soit stable par les 3 lois de A

(ii) $1_A \in B$

(iii) La restriction à B des lois de A définit sur B une structure d'algèbre.

Proposition 8.2.5 – Reformulation de la définition des sous-algèbres

Ainsi, les propositions suivantes sont équivalentes: :

(i) B est une sous-algèbre de A ;

(ii) B est un sous-espace vectoriel de A, stable par $\times$ et vérifiant $1_A \in B$

(iii) B est un sous-anneau de A stable par combinaison linéaire.

◁ **Éléments de preuve.**

Comme toujours dans ce type de propriété, une fois la stabilité acquise, toutes les propriétés universelles sont préservées par restriction. Il faut juste s'assurer des propriétés existentielles. ▷

Corollaire 8.2.6 – Caractérisation des sous-algèbres

B est une sous-algèbre de A si et seulement si :

(i) $B \subset A$

(ii) $1_A \in B$

(iii) $\forall(\lambda, x, y) \in \mathbb{K} \times B \times B,\quad \lambda x = y \in B$

(iv) $\forall(x, y) \in B \times B,\quad xy \in B$

Exemples 8.2.7

1. $\mathbb{K}$ est une sous-algèbre de $\mathbb{K}[X]$ (le corps $\mathbb{K}$ étant identifié aux polynômes constants).
2. Plus généralement, si A est une $\mathbb{K}$-algèbre, $\mathbb{K} \cdot 1_A = \mathrm{Vect}(1_A)$ est une sous-algèbre de A
3. $\mathbb{K}[X]$ est une sous-algèbre de $\mathbb{K}[X, Y]$.
4. L'ensemble $\mathcal{D}_n(\mathbb{K})$ des matrices diagonales est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$.
5. De même, $\mathcal{T}_n^+(\mathbb{K})$ est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$.
6. $\mathcal{T}_n^{++}(\mathbb{K})$ (matrices strictement triangulaires supérieures) est-elle une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$?

Définition 8.2.8 – Morphisme d'algèbre

Soit $(A, +, \cdot, \times)$ et $(B, +, \cdot, \times)$ deux algèbres. Un morphisme d'algèbre est une application $f \in \mathcal{F}(A, B)$ telle que :

(i) f est une application linéaire de $(A, +, \cdot)$ dans $(B, +, \cdot)$;

(ii) f est un morphisme d'anneaux de $(A, +, \times)$ dans $(B, +, \times)$

Un isomorphisme d'algèbre est un morphisme d'algèbre bijectif.

Ainsi, en notant 1_A et 1_B les neutres multiplicatifs respectifs de A et B, f est un morphisme d'algèbre si et seulement si :

- $\forall(\lambda, x, y) \in \mathbb{K} \times A \times A,\ f(\lambda x + y) = \lambda f(x) + f(y)$
- $\forall(x, y) \in A^2,\ f(xy) = f(x)f(y)$
- $f(1_A) = 1_B$.

Exemples 8.2.9

1. $\mathbb{K} \cdot 1_A$ est isomorphe à $\mathbb{K}$.
2. $f : \mathbb{R}[X] \to \mathbb{C}$ définie par $f(P) = P(\mathrm{i})$ est un morphisme d'algèbre.
3. Plus généralement pour $a \in \mathbb{K}$, l'évaluation $\mathrm{ev}_a : \mathbb{K}[X] \to \mathbb{K}$ définie par $\mathrm{ev}_a(P) = P(a)$ est un morphisme d'algèbre.
4. Si a est transcendant (i.e. non racine d'un polynôme de $\mathbb{Q}[X]$), alors $\mathbb{Q}[a]$ est isomorphe à $\mathbb{Q}[X]$.

Proposition 8.2.10 – Image d'une sous-algebre

Soit $f : A \to B$ un morphisme d'algèbre, et A' une sous-algèbre de A. Alors $f(A')$ est une sous-algèbre de B.

II.2 Spécialisation d'un polynôme

Définition 8.2.11 – Spécialisation d'un polynôme sur une algèbre

Soit A une $\mathbb{K}$-algèbre, et soit $P \in \mathbb{K}[X]$, décrit par

$$P = \sum_{k=0}^{d} a_k X^k.$$

La spécialisation de P en $x \in A$ est définie par

$$P(x) = \sum_{k=0}^{d} a_k x^k,$$

Remarque 8.2.12

1. La spécialisation $P(x)$ est bien définie, puisqu'on sait :
   - faire des produits (donc prendre des puissances) dans A, y compris un produit vide (la puissance d'ordre 0), puisque A est unitaire : pour tout $x \in A$, $x^0 = 1_A$;
   - multiplier un élément de A par un scalaire $a_i \in \mathbb{K}$;
   - sommer entre eux des éléments de A.
2. Le terme « évaluation » pour désigner $P(x)$ est en général réservé au cas où $x \in \mathbb{K}$. Ainsi, on ne peut évaluer qu'en des éléments du corps de base, mais on peut spécialiser sur des objets plus complexes.
3. L'évaluation est un cas particulier de spécialisation, en considérant $A = \mathbb{K}$, qui est bien une algèbre.

Exemples 8.2.13

1. Spécialisation en une matrice carrée : si $M \in \mathcal{M}_n(\mathbb{K})$, on peut considérer $P(M)$.

   Par exemple, déterminer $P(M)$ lorsque $P = 2X^2 + X + 1$ et $M = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$.
2. Spécialisation en un endomorphisme : si $u \in \mathcal{L}(E)$, $P(u)$ est bien défini, et est un élément de $\mathcal{L}(E)$, les puissances qui interviennent correspondant à l'itération de la composition.

Par exemple si $P = 2X^2 + X + 1$, $P(u)$ est l'endomorphisme défini par :

$$P(u)(x) = 2u^2(x) + u(x) + u^0(x) = 2u \circ u(x) + u(x) + x.$$

3. Attention à ne pas confondre :
   (i) $P(u)(x)$ évaluation de l'endomorphisme $P(u)$ en x
   (ii) $P(u(x))$ spécialisation du polynôme P en $u(x)$ qui nous amène donc à prendre des puissances du vecteur $u(x)$, ce qui n'a de sens que si E est elle-même une algèbre ; ces puissances n'ont alors aucun rapport avec des compositions de u par lui-même.

Proposition 8.2.14 – Morphisme de spécialisation

Pour $x \in A$, l'application $S_x : \mathbb{K}[X] \to A$ définie par $S_x(P) = P(x)$ est un morphisme d'algèbre.

◁ **Éléments de preuve.**
Le seul point un peu délicat à montrer est le respect du produit qui nécessite une petite manipulation de sommes. ▷

Proposition/Définition 8.2.15 – Sous-algèbre engendré par un élément

Soit A une $\mathbb{K}$-algèbre, et $x \in A$. On définit

$$\mathbb{K}[x] = \{P(x),\ P \in \mathbb{K}[X]\}$$

Alors $\mathbb{K}[x]$ est une sous-algèbre commutative de A, et est minimale parmi les sous-algèbres de A contenant x. On l'appelle sous-algèbre de A engendrée par x.

Ainsi, $\mathbb{K}[x]$ est l'ensemble des éléments de A s'obtenant polynomialement à partir de x.

◁ **Éléments de preuve.**
C'est l'image par un morphisme d'algèbre d'une algèbre commutative. ▷

En particulier, même si A n'est pas commutative, deux polynômes d'un même élément $x \in A$ commutent.

Exemples 8.2.16

1. Pour $A = \mathcal{L}(E)$ et $u \in \mathcal{L}(E)$, le produit dans $\mathcal{L}(E)$ étant la composition, le respect du produit par S_u s'écrit :
$$P(u) \circ Q(u) = (PQ)(u) = (QP)(u) = Q(u) \circ P(u).$$
L'égalité du milieu exploite la commutativité dans $\mathbb{K}[X]$, qui permet donc d'obtenir la commutativité de l'algèbre $(\mathbb{K}[u], +, \cdot, \circ)$.
2. Décrire $\mathbb{K}[M]$, où $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

On remarque que cette propriété de commutation est un cas particulier d'un résultat un peu plus général :

Proposition 8.2.17 – Polynômes d'endomorphismes qui commutent

Soit u et v dans $\mathcal{L}(E)$ tels que $u \circ v = v \circ u$. Soit P et Q deux polynômes de $\mathbb{K}[X]$. Alors $P(u) \circ Q(v) = Q(v) \circ P(u)$.

◁ **Éléments de preuve.**
Par manipulation de sommes, en remarquant que $u^k v^\ell = v^\ell u^k$. ▷

Corollaire 8.2.18 – Des propriétés de stabilité par commutation

Soit u et v des endomorphismes qui commutent et P et Q des polynômes. Alors :

1. $\mathrm{Im}(u)$ et $\mathrm{Ker}(u)$ sont stables par v ;
2. $\mathrm{Im}(P(u))$ et $\mathrm{Ker}(P(u))$ sont stables par $Q(v)$

◁ **Éléments de preuve.**
Le 1 est classique et se démontre indépendamment de ce qui précède. Le 2 consiste à appliquer le 1 aux endomorphismes $P(u)$ et $Q(v)$ qui commutent. ▷

III Polynômes d'endomorphismes

III.1 Polynômes annulateurs et polynôme minimal

On rappelle que pour E un $\mathbb{K}$-espace vectoriel et $u \in \mathcal{L}(E)$, l'application $S_u : \mathbb{K}[X] \to \mathcal{L}(E)$ est un morphisme d'algèbre. En particulier, c'est un morphisme d'anneaux. On en déduit la structure de son noyau.

Proposition/Définition 8.3.1 – Idéal des polynômes annulateurs

- Le noyau $\mathrm{Ker}(S_u)$ de $S_u : P \mapsto P(u)$ est un idéal, constitué de tous les polynômes $P \in \mathbb{K}[X]$ tels que $P(u) = 0_{\mathcal{L}(E)}$.
- Un polynôme vérifiant $P(u) = 0_{\mathcal{L}(E)}$ est appelé polynôme annulateur de u.

Proposition 8.3.2 – Existence d'un polynôme annulateur non nul en dimension finie

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $u \in \mathcal{L}(E)$. Alors u admet au moins un polynôme annulateur non nul.

◁ **Éléments de preuve.**
Pour $n = \dim(E)$, la famille $(u^0, u^1, \dots, u^{n^2})$ est liée car $\dim(\mathcal{L}(E)) = n^2$. ▷

Définition 8.3.3 – Polynôme minimal

Le polynôme minimal de $u \in \mathcal{L}(E)$ est l'unique polynôme unitaire engendrant l'idéal principal $\mathrm{Ker}(S_u)$ des polynômes annulateurs de u.
Il est souvent noté μ_u ou π_u.

Proposition 8.3.4 – Base et dimension de $\mathbb{K}[u]$

Soit E un espace vectoriel de dimension finie n, et $u \in \mathcal{L}(E)$. Soit $d = \deg(\mu_u)$. Alors

$$\dim \mathbb{K}[u] = d,$$

et une base en est $(\mathrm{id}, u, \dots, u^{d-1})$.

◁ **Éléments de preuve.**
Caractère générateur par division euclidienne. Liberté par minimalité du degré de μ_u. ▷

Ainsi, le polynôme minimal μ_u est l'unique polynôme unitaire tel que pour tout polynôme annulateur P de u, $\mu_u \mid P$. C'est donc aussi, parmi les polynômes annulateurs unitaires non nuls, celui dont le degré est minimal.
On peut évidemment traduire ces définitions en termes matriciels. Pour $M \in \mathcal{M}_n(\mathbb{K})$, l'ensemble des polynômes annulateurs de M est l'idéal $\mathrm{Ker}(S_M)$, non réduit à 0 car $(M^0, \dots, M^{n^2})$ est liée, et engendrée par un unique polynôme annulateur unitaire μ_M ou π_M, appelé polynôme minimal de M.
Avant de voir des exemples de polynômes minimaux, on donne une définition :

Définition 8.3.5 – Endomorphisme nilpotent

Un endomorphisme $u \in \mathcal{L}(E)$ est nilpotent s'il existe $k \in \mathbb{N}$ tel que $u^k = 0_{\mathcal{L}(E)}$. Dans ce cas, l'entier

$$d = \min\{k \in \mathbb{N}, \quad u^k = 0_{\mathcal{L}(E)}\}$$

est appelé indice de nilpotence de u. C'est donc le plus petit entier positif tel que $u^d = 0$.

L'indice de nilpotence d'un endomorphisme nilpotent est toujours au moins égal à 1 (car $u^d = \mathrm{id}$, et même à 2, sauf si $u = 0$.

Cette définition s'adapte bien entendu aux matrices.

Exemples 8.3.6

Les exemples suivants doivent être pris comme des exemples du cours.

1. Polynôme minimal de $0_{\mathcal{L}(E)}$.
2. Polynôme minimal d'une matrice scalaire λI_n, $\lambda \neq 0$.
3. Polynôme minimal de $J_n = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$
4. Les projecteurs sont définis algébriquement par un polynôme annulateur $X^2 - X$. À quelle condition est-ce le polynôme annulteur
5. De même, $X^2 - 1$ est un polynôme annulateur de toute symétrie. À quelle condition est-ce le polynôme minimal ?

L'exemple de la matrice J_n se généralise de la sorte.

Proposition 8.3.7 – Caractérisation d'un endomorphisme nilpotent par son pol. min.

Soit $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes:

(i) u est nilpotente d'indice de nilpotence d ;

(ii) $\mu_u = X^d$.

Par ailleurs, on sait comparer le degré de μ_d à $n = \dim(E)$, dans ce cas particulier, grâce au lemme suivant.

Lemme 8.3.8 – Majoration de l'indice de nilpotence

Soit E un espace vectoriel de dimension finie n, et soit u un endomorphisme nilpotent de E, d'indice de nilpotence d. Alors $d \leqslant n$.

◁ **Éléments de preuve.**

Montrer qu'il existe $x \in E$ tel que $(x, u(x), \dots, u^{d-1}(x))$ soit libre. ▷

Ainsi, le polynôme minimal d'un endomorphisme minimal est toujours de degré inférieur à $\dim(E)$. Ce résultat se généralisera à tout polynôme minimal plus tard, grâce au théorème de Cayley-Hamilton, qui fournit un polynôpme annulateur de degré $n = \dim(E)$ (ce qui assure de facto que le polynôme minimal est de degré plus petit).

III.2 Lemme de décomposition des noyaux

On termine cette section dédiée aux polynômes d'endomorphismes et de matrices par un résultat qui sera d'une importance capitale pour les problèmes de diagonalisation, et plus généralement de réduction. Je lui donne un statut de théorème, même s'il est communément appelé lemme.

Théorème 8.3.9 – Lemme de décomposition des noyaux

Soit $r \in \mathbb{N}^*$, et $P_1, \dots, P_r$ des éléments de $\mathbb{K}[X]$ deux à deux premiers entre eux. Alors

$$\mathrm{Ker}(P(u)) = \bigoplus_{i=1}^{r} \mathrm{Ker}(P_i(u)).$$

◁ **Éléments de preuve.**

Par récurrence immédiate, la primalité 2 à 2 impliquant en particulier que P_k est premier avec $P_1 \dots P_{k-1}$, il suffit de le montrer lorsque $r = 2$.

Pour $r = 2$, se servir d'une relation de Bézout, à la fois pour montrer l'inclusion directe, et pour montrer le caractère directe de la somme. ▷

Corollaire 8.3.10 – Décomposition de E issue d'un polynôme annulateur

Soit P un polynôme annulateur de u, tel que $P = P_1 \dots P_r$, où les P_i sont 2 à 2 premiers entre eux. Alors

$$E = \bigoplus_{i=1}^{r} \mathrm{Ker}(P_i(u)).$$

Exemples 8.3.11

1. Si p est un projecteur, on retrouve $E = \mathrm{Ker}(p) \oplus \mathrm{Ker}(p - \mathrm{id})$
2. Si s est une symétrie, on retrouve $E = \mathrm{Ker}(s - \mathrm{id}) \oplus \mathrm{Ker}(s + \mathrm{id})$.
3. Dans les deux exemples précédents, trouver une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(p)$ (resp. $\mathrm{Mat}_{\mathcal{B}}(s)$) soit diagonale.

Le dernier point nous fait bien comprendre l'utilité de ce résultat dans des études de diagonalisabilité. Ainsi, on obtient par exemple le résultat important suivant (qui sera repris et complété dans le chapitre suivant) :

Théorème 8.3.12 – Caractérisation de la diagonalisabilité par polynômes annulateurs

Soit E de dimension finie, et $u \in \mathcal{L}(E)$ admettant un polynôme annulateur scindé à racines simples. Alors P est diagonalisable.

◁ **Éléments de preuve.**

Par le lemme de décomposition des noyaux, puis prendre une base de chaque terme de la somme directe.

C'est même une équivalence, comme on le justifiera plus tard. D'où la terminologie « caractérisation ». ▷

9

Réduction des endomorphismes

Le problème de la réduction des endomorphismes consiste à trouver une base dans laquelle la forme de la matrice associée sera suffisamment simple pour être propice aux calculs.
L'idéal est par exemple de trouver une base dans laquelle la matrice est diagonale. C'est le principe de la diagonalisation. Ce n'est pas toujours possible, c'est pourquoi on recherche parfois des formes un peu plus générale, comme des matrices triangulaires par exemple.
Lorsqu'on adapte la situation aux matrices, il s'agit donc de travailler sur l'endomophisme f de $\mathbb{K}^n$ canoniquement associé à une matrice M, et de trouver une base dans laquelle la nouvelle matrice N de f sera d'une forme particulière, par exemple diagonale ou triangulaire. Par la formule de changement de base, on a donc une relation du type $M = PNP^{-1}$, où $P \in \mathrm{GL}_n(\mathbb{K})$.

Définition 9.0.1 – Matrices semblables

On dit que deux matrices M et N de $\mathcal{M}_n(\mathbb{K})$ sont semblables s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $M = PNP^{-1}$. Cela définit une relation binaire sur $\mathcal{M}_n(\mathbb{K})$, appelée relation de similitude.

Cette relation est un peu plus forte que la relation d'équivalence, puisqu'on impose l'égalité des deux matrices de passage. De fait, deux matrices semblables sont équivalentes mais la réciproque est fausse.

Proposition 9.0.2 – La relation de similitude est une équivalence

La relation de similitude est une relation d'équivalence. Ses classes d'équivalence sont appelées classes de similitude.

Ainsi, le problème de la réduction des endomorphismes se ramène donc à la recherche de représentants simples dans une classe de similitude donnée. Dans l'idéal, on souhaiterait donner même un représentant privilégié dans chaque classe de similitude, comme on l'avait fait avec les matrices $J_{n,p,r}$ pour les classes d'équivalence matricielle, de sorte à pouvoir classifier les classes de similitude. Comme on le verra, c'est un problème assez dur, auquel on ne répondra pas complètement cette année.

I Élements propres d'un endomorphisme

On suppose que E est un espace vectoriel sur $\mathbb{K}$. Le corps $\mathbb{K}$ est supposé être un sous-corps de $\mathbb{C}$, conformément au programme, même si la plupart des résultats de ce chapitre restent vrais sans cette condition.

I.1 Valeurs propres, vecteurs propres, sous-espaces propres

Si l'on souhaite trouver une base $\mathcal{B}$ telle que $\mathrm{Mat}_{\mathcal{B}}(u)$ soit diagonale, en particulier $u(b) \in \mathbb{K}b$, donc $u(\mathbb{K}b) \subset \mathbb{K}b$. Les vecteurs de cette base engendrent donc des droites stables par u.

Le problème de la diagonalisation est donc très lié à la recherche des droites stables.
Réciproquement, une droite $\mathbb{K}b$ est stable si et seulement si $u(b) \in \mathbb{K}b$, donc s'il existe $\lambda \in \mathbb{K}$ tel que $u(b) = \lambda b$, c'est-à-dire $(u - \lambda \mathrm{id})(b) = 0$.

Définition 9.1.1 – Valeurs propres, vecteurs propres, sev propres

Soit $u \in \mathcal{L}(E)$.

1. Une valeur propre de u est un élément $\lambda \in \mathbb{K}$ tel que $\mathrm{Ker}(u - \lambda \mathrm{id}) \neq \{0\}$, donc tel que $u - \lambda \mathrm{id}$ ne soit pas injectif.
2. Soit λ une valeur propre de u. Le sous-espace propre associé à λ est :
$$E_\lambda(u) = \mathrm{Ker}(u - \lambda \mathrm{id}_E).$$
3. Les éléments **non nuls** de $E_\lambda(u)$ sont alors appelés vecteurs propres associés à la valeur propre λ.

Remarque 9.1.2

1. Par définition, si λ est une valeur propre, $E_\lambda(u) \neq \{0\}$, et il existe donc au moins un vecteur propre associé à λ.
   Une valeur propre λ se caractérise donc par le fait qu'il existe $x \in E \backslash \{0\}$ tel que $u(x) = \lambda x$.
2. Il arrive parfois, pour des raisons techniques (éviter des discussions) d'étendre la notation $E_\lambda(u)$ au cas où λ n'est pas une valeur propre. Dans ce cas, λ est une valeur propre ssi $E_\lambda(u) \neq \{0\}$.
3. Lorsque E est de dimension infinie, la non injectivité de $u - \lambda \mathrm{id}$ n'équivaut pas à sa non inversibilité. La non inversibilité définit ce qu'on appelle les valeurs spectrales de u (hors programme).
4. Les sev propres sont parfois aussi notés $\mathrm{SEP}_\lambda(u)$.

Exemple 9.1.3

Décrire les valeurs propres et les sous-espaces propres d'un projecteur.

Définition 9.1.4 – Spectre

Lorsque E est de dimension finie, on note $\mathrm{Sp}(u)$ l'ensemble des valeurs propres de u. L'ensemble $\mathrm{Sp}(u)$ est appelé spectre de u.

Remarque 9.1.5

En dimension infinie, le spectre est défini non pas avec les valeurs propres, mais avec les valeurs spectrales, raison pour laquelle cette définition n'est ici donnée qu'en dimension finie.

Proposition 9.1.6 – Somme directe des s.e.p.

Soit $\lambda_1, \dots, \lambda_p$ des valeurs propres. Alors la somme $\bigoplus_{i=1}^{p} E_{\lambda_p}(u)$ est directe.

◁ **Éléments de preuve.**

Par le lemme de décomposition des noyaux. Peut aussi se montrer facilement directement par récurrence, c'est un bon entraînement, mais c'est un peu plus fastidieux. ▷

Corollaire 9.1.7

Soit $(x_i)_{i\in I}$ une famille de vecteurs propres associés à des valeurs propres 2 à 2 distinctes. Alors $(x_i)_{i\in I}$ est libre.

◁ **Éléments de preuve.**

Ce sont des vecteurs non nuls de sous-espaces vectoriels en somme directe. ▷

Corollaire 9.1.8 – Majoration du nombre de valeurs propres

Soit E un espace vectoriel de dimension finie n, et $u \in \mathcal{L}(E)$. Alors $|\mathrm{Sp}(u)| \leqslant n$.

◁ **Éléments de preuve.**

Soit par l'examen des dimensions, soit par le corollaire précédent qui donne sinon une famille libre trop grande. ▷

Dans le but d'étudier des propriétés de réduction simultanée, nous aurons besoin du lemme suivant :

Proposition 9.1.9 – Stabilité des sev propres de u par v commutant avec u

Soit $u, v \in \mathcal{L}(E)$ tels que $u \circ v = v \circ u$. Alors v laisse stable les sous-espaces propres de u (et inversement).

◁ **Éléments de preuve.**

Application directe de 8.2.17. ▷

Remarque 9.1.10

En particulier, puisque u commute avec lui-même, les sous-espaces propres de u sont stables par u. Cela se comprend bien de façon directe, sans passer par ce résultat, et c'est une évidence, mais ça ne fait pas de mal que ce soit dit explicitement au moins une fois !

Enfin, remarquons qu'une façon de restreindre la recherche des valeurs propres de u est de s'intéresser à des polynômes annulateurs :

Théorème 9.1.11 – Valeurs propres et racines d'un polynôme annulateur

Soit $u \in \mathcal{L}(E)$.

1. Soit P un polynôme annulateur de u. Alors toute valeur propre est racine de P.
2. Si u admet un polynôme minimal μ_u, alors les racines de μ_u sont exactement les valeurs propres de u

◁ **Éléments de preuve.**

1. Par récurrence, si x est un vecteur propre associé à λ, $u^n(x) = \lambda^n \cdot x$, puis $P(u)(x) = P(\lambda) \cdot x$.
2. Pour le point 2, une inclusion est acquise. Si l'inclusion est stricte, on peut factoriser $P(u)$ par $u - \lambda\mathrm{id}$, où λ n'est pas une valeur propre. L'injectivité de $u - \lambda\mathrm{id}$ permet alors de contredire la minimalité de μ_u.

▷

Remarque 9.1.12

Lorsque E est de dimension finie, on peut réexprimer ces propriétés à l'aide du spectre :

1. Pour tout polynôme annulateur P de u, $\mathrm{Sp}(u) \subset \mathrm{Rac}(P)$;
2. $\mathrm{Sp}(u) = \mathrm{Rac}(\mu_u)$.

Exemples 9.1.13

1. Quelle est la seule valeur propre possible d'un endomorphisme nilpotente ?
2. Déterminer les valeurs propres dans $\mathbb{C}$ de l'endomorphisme de $\mathbb{C}^3$ canoniquement associé à $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

I.2 Éléments propres d'une matrice carrée

Le dernier exemple suggère qu'on peut adapter toutes les définitions et propriétés au cas des matrices. Comme l'espace de travail est alors de dimension finie, on n'a pas besoin de précaution particulière pour parler du spectre.

Soit $M \in \mathcal{M}_n(\mathbb{K})$.

- Un scalaire $\lambda \in \mathbb{K}$ est valeur propre de M si $\mathrm{Ker}(M - \lambda I_n) \neq \{0\}$, c'est-à-dire s'il existe $X \neq 0$ dans $\mathcal{M}_{n,1}(\mathbb{K})$ tel que $MX = \lambda X$. Cela équivaut aussi à dire que la matrice $M - \lambda I_n$ n'est pas inversible.
- L'ensemble des valeurs propres de M est noté $\mathrm{Sp}(M)$, et est appelé spectre de M.
- Un vecteur propre associé à λ est un élément non nul $X \in \mathcal{M}_{n,1}(\mathbb{K})$ tel que $MX = \lambda X$, c'est-à-dire $X \in \mathrm{Ker}(M - \lambda I_n) \backslash \{0\}$.
- Le sous-espace propre $E_\lambda(M)$ associé à la valeur propre λ est $\mathrm{Ker}(M - \lambda I_n)$. C'est donc l'ensemble des vecteurs propres, augmenté de 0.
- Les sous-espaces propres sont en somme directe, et donc en nombre au plus égal à n
- Les valeurs propres sont racines de tout polynôme annulateur P de M : $\mathrm{Sp}(M) \subset \mathrm{Rac}(P)$
- Les valeurs propres sont exactement les racines du polynôme minimal : $\mathrm{Sp}(M) = \mathrm{Rac}(\mu_M)$.

À cela on ajoute les deux propriétés suivantes.

Proposition 9.1.14 – Spectre de matrices semblables

Deux matrices semblables ont même spectre.

◁ **Éléments de preuve.**

Si M et N sont semblables, $M - \lambda I_n$ et $N - \lambda I_n$ le sont aussi, donc ont même rang. ▷

Le lien entre éléments propres d'une matrice et éléments propres d'un endomorphisme se fait via les résultats suivants.

Proposition 9.1.15 – Élements propres de u et de $\mathrm{Mat}_{\mathcal{B}}(u)$

Soit $u \in \mathcal{L}(E)$ et $\mathcal{B}$ une base de E. Alors :

1. $\mathrm{Sp}(u) = \mathrm{Sp}(\mathrm{Mat}_{\mathcal{B}}(u))$;
2. x est vecteur propre de u associé à λ ssi $[x]_{\mathcal{B}}$ est vecteur propre de $\mathrm{Mat}_{\mathcal{B}}(u)$ associé à λ ;
3. Pour $\lambda \in \mathrm{Sp}(u)$, $E_\lambda(\mathcal{M}_{\mathcal{B}}(u)) = \varphi_{\mathcal{B}}(E_\lambda(u))$, où $\varphi_{\mathcal{B}} : x \mapsto [x]_{\mathcal{B}}$.

On fait aussi la correspondance entre polynômes de l'endomorphisme u et polynômes de la matrice $\mathcal{M}_{\mathcal{B}}(u)$.

Proposition 9.1.16 – Polynômes d'endomorphismes versus polynômes de matrices

Soit $u \in \mathcal{L}(E)$ et $\mathcal{B}$ une base de E.

1. Pour tout $P \in \mathcal{L}(E)$, $\mathrm{Mat}_{\mathcal{B}}(P(u)) = P(\mathrm{Mat}_{\mathcal{B}}(u))$.
2. En particulier, P est un polynôme annulateur de u si et seulement s'il est un polynôme annulateur de $\mathrm{Mat}_{\mathcal{B}}(u)$.

3. Notamment, $\mu_u = \mu_{\mathrm{Mat}_{\mathcal{B}}(u)}$.

Sur les matrices, on dispose d'une opération d'extension des scalaires, consistant à voir les matrices de $\mathcal{M}_n(\mathbb{K})$ comme des matrices de $\mathcal{M}_n(\mathbb{K}')$, dès lors que $\mathbb{K}'$ est un corps contenant $\mathbb{K}$. C'est notamment ce qu'il se produit lorsqu'on considère une matrice à coefficients réels comme une matrice à coefficients complexes. Dans cet exemple, on peut notamment se demander si le spectre d'une matrice réelle M reste le même selon qu'on considère M comme une matrice réelle ou comme une matrice complexe. La réponse est non en général, mais une inclusion est vraie.

Proposition 9.1.17 – Extension du corps des coefficients

Soit $\mathbb{K} \subset \mathbb{K}'$ deux corps tels que $\mathbb{K}$ soit un sous-corps de $\mathbb{K}'$, et soit $M \in \mathcal{M}_n(\mathbb{K})$. On note $\mathrm{Sp}_{\mathbb{K}}(M)$ et $\mathrm{Sp}_{\mathbb{K}'}(M)$ le spectre de M vue respectivement comme matrice de $\mathcal{M}_n(\mathbb{K})$ et de $\mathcal{M}_n(\mathbb{K}')$. Alors $\mathrm{Sp}_{\mathbb{K}}(M) \subset \mathrm{Sp}_{\mathbb{K}'}(M)$.

◁ **Éléments de preuve.**
Un vecteur propre $X \in \mathcal{M}_{n,1}(\mathbb{K})$ est aussi un vecteur propre dans $\mathcal{M}_{n,1}(\mathbb{K}')$. ▷

Exemple 9.1.18

L'inclusion peut être stricte, comme le montre l'exemple de $M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

I.3 Polynôme caractéristique d'un endomorphisme en dimension finie

Parfois, la résolution de l'équation aux éléments propres $u(x) = \lambda x$ peut s'avérer laborieuse et calculatoire, et la recherche d'un polynôme annulateur un peu hasardeuse. On donne dans cette section un outil qui permet également de déterminer les valeurs propres. Il s'agit du polynôme caractéristique de u, dont les racines sont les valeurs propres.

Soit E un espace vectoriel de dimension finie. Le fait que λ soit valeur propre se traduit par le fait que $\lambda \mathrm{id} - u$ ne soit pas bijective, c'est-à-dire que le déterminant de sa matrice relativement à une base quelconque est nul.

Or, si $M \in \mathcal{M}_n(\mathbb{K})$, $XI_n - M$ est une matrice à coefficients dans $\mathbb{K}[X]$. Le déterminant ne faisant intervenir que des produits et des sommes des coefficients de la matrice, $\det(XI_n - M)$ est bien défini dans $\mathbb{K}[X]$.

Définition 9.1.19 – Polynôme caractéristique d'une matrice

Soit $M \in \mathcal{M}_n(\mathbb{K})$. Le polynôme caractéristique χ_M de M est le polynôme défini par

$$\chi_M(X) = \det(XI_n - M) \in \mathbb{K}[X]$$

Propriétés 9.1.20 – Coefficients du polynôme caractéristique

Soit $M \in \mathcal{M}_n(\mathbb{K})$. Alors

(i) $\deg(\chi_M) = n$.

En écrivant $\chi_M = \sum_{k=0}^{n} a_k X^k$, on a alors :

(ii) $a_n = 1$ (donc χ_M est unitaire)

(iii) $a_{n-1} = -\mathrm{tr}(M)$

(iv) $a_0 = (-1)^n \det(M)$.

Ainsi, $\chi_M = X^n - \mathrm{tr}(M)X^{n-1} + \cdots + (-1)^n \det(M)$.

◅ **Éléments de preuve.**
Par la formule de développement, tous les termes sont de degré au plus $n-2$, sauf celui indexé par id. En effet, une permutation de $\mathfrak{S}_n$ ne peut pas avoir exactement $n-1$ points fixes.
On obtient donc les termes de degré n et $n-1$ en développant $(X-a_{1,1})\cdots(X-a_{n,n})$, ce qui se fait avec les relations de Viète.
Le coefficient a_0 s'obtient par évaluation. ▻

Proposition 9.1.21 – Racines du polynôme caractéristique

$\mathrm{Rac}(\chi_M) = \mathrm{Sp}(M)$.

◅ **Éléments de preuve.**
λ est racine ssi $\det(\lambda I_n - M) = 0$. Qu'est-ce que cela signifie sur l'inversibilité de la matrice ? ▻

Voici une situation dans laquelle le polynôme caractéristique est simple à calculer :

Proposition 9.1.22 – Polynôme caractéristique d'une matrice triangulaire

Soit T une matrice triangulaire, de coefficients diagaunaux $t_{1,1},\dots,t_{n,n}$. Alors

$$\chi_T(X) = (X-t_{1,1})\cdots(X-t_{n,n}).$$

Ainsi, les valeurs propres de T sont les coefficients diagonaux.

Proposition 9.1.23 – Polynôme caractéristique de matrices semblables

Soit M et N deux matrices semblables. Alors $\chi_M = \chi_N$.

Par conséquent, vu les deux derniers résultats obtenus, si on recherche une matrice triangulaire T semblable à une matrice M, les coefficients diagonaux de T sont les valeurs propres de M, leur multiplicité sur la diagonale de T correspondant à leur multiplicité dans le polynôme caractéristique de M. C'est ce qu'on appelle la multiplicité algébrique d'une valeur propre.
Le dernier résultat permet aussi de définir le polynôme caractéristique d'un endomorphisme u d'un espace E de dimension finie. En effet, si $\mathcal{B}$ et $\mathcal{C}$ sont deux bases de E, $\mathrm{Mat}_{\mathcal{B}}(u)$ et $\mathrm{Mat}_{\mathcal{C}}(u)$ sont semblables.

Définition 9.1.24 – Polynôme caractéristique de $u \in \mathcal{L}(E)$

Soit u un endomorphisme d'un espace vectoriel E de dimension finie. Alors le polynôme caractéristique de u est défini par

$$\chi_u = \chi_{\mathrm{Mat}_{\mathcal{B}}(u)},$$

où $\mathcal{B}$ est une base de E. Cette définition est indépendante du choix de la base.

Les propriétés du polynôme caractéristique de u sont alors similaires à celles d'une matrice représentant u. Par exemple, comme la trace de u est également définie comme étant la trace commune des matrices représentant u dans une base quelconque, et comme le déterminant vérifie aussi cette propriété, on en déduit que

$$\chi_u(X) = X^n - \mathrm{tr}(u)X^{n-1} + \cdots + (-1)^n \det(u).$$

Proposition 9.1.25 – Polynôme caractéristique d'un endomorphisme induit

Soit F stable par $u \in \mathcal{L}(E)$, et u_F l'endomorphisme induit sur F par u. Alors $\chi_{u_F} \mid \chi_u$.

◅ **Éléments de preuve.**
Dans une base convenable, la matrice de u est triangulaire par bloc, l'un des blocs étant une matrice de u_F. Utiliser le calcul du déterminant de matrices triangulaires par blocs. ▻

I.4 Recherche de valeurs propres

On a donc essentiellement 3 méthodes pour la recherche des valeurs propres, plus ou moins efficaces selon la situation

Méthode 9.1.26 – Recherche de valeurs propres

1. Résoudre l'équation aux éléments propres $u(x) = \lambda x$, en essayant de voir à quelle condition sur λ cette équation admet une solution autre que la solution triviale. En dimension infinie, sauf s'il existe un polynôme annulateur de petit degré facile à deviner, c'est la seule solution vraiment envisageable.
2. Rechercher un polynôme annulateur, si possible de petit degré. En dimension finie, on sait qu'il en existe un, et on montrera qu'il en existe un de degré au plus $n = \dim(E)$. En se plaçant dans le cadre matriciel, on peut trouver un polynôme annulateur en calculant les premières puissances M^k, et en trouvant une relation entre ces puissances. Les valeurs propres sont à trouver parmi les racines de ce polynôme.

   C'est parfois peu commode et hasardeux. De plus, il faut un peu d'habitude pour trouver les polynômes annulateurs après avoir calculé les premières puissances. Enfin, comme on ne devine pas toujours le polynôme minimal, on peut trouver des racines qui ne sont pas valeurs propres. Ça peut en revanche être efficace s'il y a peu de valeurs propres par rapport à la dimension, ou si le contexte fournit un polynôme annulateur naturel.
3. Calculer le polynôme caractéristique, par un calcul de déterminant, et chercher les racines de ce polynôme.

Méthode 9.1.27 – Recherche d'espaces propres

Une fois qu'on a déterminé les valeurs propres λ, ou au moins, un ensemble fini qui contient toutes les valeurs propres, résoudre $u(x) = \lambda x$, de l'inconnue x. Si on est en dimension finie, cela se ramène à la résolution d'un système linéaire d'équations.

Exemples 9.1.28

1. Déterminer les valeurs propres et les espaces propres associés de l'endomorphisme $D : f \mapsto f'$ du $\mathbb{C}$-ev $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$.
2. Déterminer les valeurs propres et les sous-espaces propres de $M \in \mathcal{M}_{2n}(\mathbb{C})$ définie par

$$M = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 & -1 \\ \vdots & & & \iddots & \iddots & 0 \\ \vdots & & \iddots & -1 & \iddots & \vdots \\ \vdots & \iddots & 1 & \iddots & & \vdots \\ 0 & \iddots & \iddots & & & \vdots \\ 1 & 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

3. Déterminer les valeurs propres et les sous-espaces propres de $\begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}$ et de $\begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$.

I.5 Multiplicité géométrique et algébrique d'une valeur propre

On donne les définitions dans le cadre d'un endomorphisme, mais cela s'adapte évidemment aussi au cas des matrices.

Définition 9.1.29 – Multiplicités d'une valeur propre

Soit E un espace vectoriel de dimension finie, et $u \in \mathcal{L}(E)$. Soit $\lambda \in \mathrm{Sp}(u)$

1. La multiplicité géométrique de λ est la dimension de $E_\lambda(u)$.
2. La multiplicité algébrique de λ est la multiplicité de λ en tant que racine de χ_u.

Remarque 9.1.30

Les deux notions peuvent différer, comme le montre le dernier exemple du 9.1.28.

En revanche, on a toujours une inégalité qui est satisfaite :

Proposition 9.1.31 – Comparaison des multiplicités géométrique et algébrique

Soit E un espace vectoriel de dimension finie, et $u \in \mathcal{L}(E)$. Alors la multiplicité géométrique est inférieure à la multiplicité algébrique.

En d'autres termes, la dimension du sous-espace propre $E_\lambda(u)$ est majorée par la multiplicité de λ en tant que racine de χ_u.

$\triangleleft$ **Éléments de preuve.**

Considérer le polynôme caractéristique de $u_{E_\lambda(u)}$, l'endomorphisme induit par u sur le sous-espace stable $E_\lambda(u)$. $\triangleright$

Faisons pour terminer un bilan des résultats obtenus pour les endomorphismes induits.

Proposition 9.1.32 – Multiplicités des valeurs propres d'un endomorphisme induit

Soit E un espace vectoriel de dimension fini, F un sev de E, et $u \in \mathcal{L}(E)$. On suppose que F est stable par u. Alors :

1. $\chi_{u_F} | \chi_u$;
2. $\mathrm{Sp}(u_F) \subset \mathrm{Sp}(u)$;
3. pour tout $\lambda \in \mathrm{Sp}(u_F)$, la multiplicité algébrique de λ en tant que valeur propre u_F est inférieure à sa multiplicité algébrique en tant que valeur propre de u ;
4. pour tout $\lambda \in \mathrm{Sp}(u_F)$, la multiplicité géométrique de λ en tant que valeur propre u_F est inférieure à sa multiplicité géométrique en tant que valeur propre de u ;

I.6 Théorème de Cayley-Hamilton

On a constaté que les polynômes annulateurs d'une matrice (ou d'un endomorphisme) ont un lien très fort avec le spectre de cette matrice, puisque le spectre est inclus dans l'ensemble des racines. On a aussi vu que les racines du polynôme caractéristique sont exactement les valeurs propres. On peut donc raisonnablement se demander si le polynôme caractéristique est un polynôme annulateur. Le fait que ses racines soient des valeurs propres n'est pas suffisant pour conclure, car on n'a pas de contrôle suffisant sur la multiplicité. Si la multiplicité est trop faible (inférieure à la multiplicité dans le polynôme minimal), la réponse est négative. Mais cette situation ne peut en fait pas se produire. C'est ce qu'affirme le théorème de Cayley-Hamilton.

Théorème 9.1.33 – Cayley-Hamilton

1. Soit M une matrice de $\mathcal{M}_n(\mathbb{K})$. Alors χ_M est un polynôme annulateur de M.
2. Soit u un endomorphisme d'un espace vectoriel E de dimension finie. Alors χ_u est un polynôme annulateur de u.

◁ **Éléments de preuve.**
Démonstration non exigible.
Les deux énoncés sont évidemment équivalents. On montre le cas matriciel.
Écrire $\chi_A(X) = \sum_{k=0}^{n} a_k X^k$, et $\mathrm{Com}(XI_n - A)^\top = \sum_{k=0}^{n-1} X^k C_k$.
Avec la formule de Cayley, trouver une relation matricielle reliant C_k, C_{k-1}, A et a_k. Former alors $\chi_A(A)$ par télescopage. ▷

Corollaire 9.1.34 – Relation entre χ_u et μ_u

1. Le polynôme minimal μ_u divise le polynôme caractéristique χ_u.
2. En particulier, $\deg(\mu_u) \leqslant n$.

II Diagonalisation

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie.

II.1 Diagonalisation d'un endomorphisme, point de vue géométrique

Définition 9.2.1 – Endomorphisme diagonalisable

Un endomorphisme $u \in \mathcal{L}(E)$ est diagonalisable s'il existe une base $\mathcal{B}$ telle que $\mathrm{Mat}_{\mathcal{B}}(u)$ soit diagonale.

Théorème 9.2.2 – Caractérisation de la diagonalisabilité par les sev propres

Soit $u \in \mathcal{L}(E)$ et χ_u son polynôme caractéristique. Les propositions suivantes sont équivalentes:

(i) u est diagonalisable ;

(ii) $\bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda(u) = E$;

(iii) $\sum_{\lambda \in \mathrm{Sp}(u)} \dim(E_\lambda(u)) = \dim(E)$;

(iv) χ_u est scindé, et pour tout $\lambda \in \mathrm{Sp}(E)$, la multiplicité algébrique de λ est égale à la multiplicité géométrique de λ.

◁ **Éléments de preuve.**

- $(i) \Longrightarrow (ii)$: toute base de diagonalisation est une famille libre de vecteurs propres.
- $(ii) \Longrightarrow (iii)$ par propriété de la dimension d'une somme directe.
- $(iii) \Longrightarrow (iv)$ car la multiplicité géométrique est toujours inférieure à la multiplicité algébrique, et car le degré de χ_u est connu.
- $(iv) \Longrightarrow (i)$: construire une base de diagonalisation à partir de bases des $E_\lambda(u)$.

▷

Corollaire 9.2.3 – Diagonalisabilité d'un endomorphisme ayant $\dim(E)$ valeurs propres

Soit $u \in \mathcal{L}(E)$ et soit $n = \dim(E)$.

1. Si u admet n valeurs propres distinctes, alors u est diagonalisable.
2. Si χ_u est simplement scindé, alors u est diagonalisable.

◁ **Éléments de preuve.**

1. Dans ce cas, on a nécessairement $\sum \dim E_\lambda(u) \geqslant n$, et trivialement dans l'autre sens aussi.
2. C'est un corollaire du point 1. On peut aussi remarquer que cela implique que la multiplicité algébrique et la multiplicité géométrique de toutes les valeurs propres sont égales.

▻

II.2 Diagonalisation d'un endomorphisme, point de vue algébrique

La caractérisation de la diagonalisabilité par le polynôme caractéristique, qui est lui-même un polynôme annulateur de u, suggère qu'on peut adopter également un point de vue purement algébrique, basé sur l'étude des polynômes annulateurs de u.
On rappelle qu'un polynôme simplement scindé est un polynôme sans racine multiple, ayant autant de racines que son degré (ou de façon équivalente, dont les seuls diviseurs irréductibles sont de degré 1).

Théorème 9.2.4 – Caractérisation de la diagonalisabilité par polynômes annulateurs

Soit u un endomorphisme de E. Les propositions suivantes sont équivalentes:

(i) u est diagonalisable ;

(ii) u admet un polynôme annulateur (non nul) simplement scindé ;

(iii) le polynôme minimal de u est simplement scindé.

◅ **Éléments de preuve.**

- $(i) \Longrightarrow (ii)$: Considérer $\prod_{\lambda \in \mathrm{Sp}(E)} (X - \lambda)$, et utiliser le décomposition de E en somme directe de sev propres.
- $(ii) \Longrightarrow (iii)$: Dans ce cas, le polynôme minimal est un diviseur d'un polynôme scindé.
- $(iii) \Longrightarrow (i)$: par le lemme de décomposition des noyaux.

▻

Corollaire 9.2.5 – Polynôme minimal d'un endomorphisme diagonalisable

u est diagonalisable si et seulement si $\mu_u = \prod_{\lambda \in \mathrm{Sp}(u)} (X - \lambda)$.

Lemme 9.2.6 – Polynôme minimal d'un endomorphisme induit

Soit u un endomorphisme de E, et F un sous-espace de E stable par u. Soit $u_F \in \mathcal{L}(F)$ l'endomorphisme induit par u sur F.

1. Tout polynôme annulateur de u est un polynôme annulateur de u_F.
2. En particulier, μ_{u_F} divise μ_u.

Proposition 9.2.7 – Diagonalisabilité d'un induit

Soit u un endomorphisme diagonalisable, et F un sous-espace de E stable par u. Alors l'endomorphisme $u_F \in \mathcal{L}(F)$ induit par u sur F est diagonalisable.

◅ **Éléments de preuve.**

Le lemme implique qu'alors u_F est simplement scindé. ▻

On en déduit le résultat suivant, hors-programme mais archi-classique en concours.

Théorème 9.2.8 – Codiagonalisabilité, HP

Soit u et v deux endomorphismes de E tels que $u \circ v = v \circ u$. Si u et v sont diagonalisable, alors il existe une base commune $\mathcal{B}$ de diagonalisation (c'est-à-dire une base dont les éléments sont tous à la fois des vecteurs propres de u et des vecteurs propres de v).

◁ **Éléments de preuve.**

Les $E_{\lambda_i}(u)$ sont stables par v. L'endomorphisme induit v_i est alors diagonalisable. Il existe donc une base de $E_{\lambda_i}(u)$ formée de vecteurs propres de v. Faire cela pour chaque sev $E_{\lambda_i}(u)$ et concaténer toutes ces bases. ▷

II.3 Diagonalisation d'une matrice

On traduit sans surprise ces définitions et propriétés dans le contexte matriciel.

Proposition/Définition 9.2.9 – Diagonalisabilité d'une matrice

Soit $M \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. Les propositions suivantes sont équivalentes:

- L'endomorphisme canonique associé à M est diagonalisable ;
- Il existe une matrice D diagonale et une matrice $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $M = PDP^{-1}$;
- M est semblable à une matrice diagonale ;

On dit qu'une matrice vérifiant ces propriétés équivalentes est diagonalisable.

◁ **Éléments de preuve.**

- $(i) \Longleftrightarrow (ii)$ par la formule de changement de base.
- $(ii) \Longleftrightarrow (iii)$ est la définition de matrices semblables.

▷

Proposition 9.2.10 – Caractérisation matricielle de la diagonalisabilité de u

Soit $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes:

(i) u est diagonalisable ;

(ii) pour toute base $\mathcal{B}$ de E, $\mathrm{Mat}_{\mathcal{B}}(u)$ est diagonalisable ;

(iii) il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(u)$ soit diagonalisable.

Du fait même de cette correspondance très forte entre diagonalisabilité d'un endomorphisme et diagonalisabilité d'une matrice, ainsi que d'après les propositions 9.1.15 et 9.1.16, toutes les propriétés vues dans le paragraphe précédent reste valides pour les matrices, et en particulier :

- la caractérisation géométrique 9.2.2 de la diagonalisabilité par la décomposition de E en somme directe de sous-espaces propres, ou par l'égalité entre multiplicité algébrique et géométrique des valeurs propres ;
- la diagonalisabilité d'une matrice de $\mathcal{M}_n(\mathbb{K})$ ayant n valeurs propres distinctes (corollaire 9.2.3) ;
- la caractérisation algébrique 9.2.4 de la diagonalisabilité par les polynômes simplement scindés.
- la propriété (HP) de codiagonalisabilité.

Définition 9.2.11 – Diagonaliser M

Diagonaliser une matrice M consiste à expliciter une matrice $P \in \mathrm{GL}_n(\mathbb{K})$ et une matrice diagonale $D \in \mathcal{D}_n(\mathbb{K})$ telles que $M = PDP^{-1}$.

Méthode 9.2.12 – Diagonaliser M

1. Rechercher les valeurs propres de M, soit en calculant le polynôme caractéristique, soit en trouvant un polynôme annulateur simple.
2. Pour tout $\lambda \in \mathrm{Sp}(M)$, résoudre le système linéaire $MX - \lambda X = 0$, et donner une base de $E_\lambda(M)$
3. Justifier (ou infirmer) la diagonalisabilité par un argument de dimension ou par les multiplicités des racines d'un polynôme annulateur.
4. Le cas échéant, construire une base de diagonalisation par concaténation de bases des espaces propres.
5. Définir P comme la matrice de passage de la base canonique à cette base, et écrire la formule de changement de base pour conclure.

Remarque 9.2.13

La diagonalisation permet par exemple de calculer les puissances d'une matrice $A = PDP^{-1}$. En effet, par simplification des termes intermédiaires, $A^n = PD^nP^{-1}$.
On peut donc aussi notamment calculer ainsi des exponentielles matricielles.

Exemples 9.2.14

Diagonalisabilité, et le cas échéant, diagonalisation des matrices suivantes :

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix} \qquad \text{et} \qquad B = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}.$$

III Trigonalisation

Toutes les matrices ne sont donc pas diagonalisables. On peut déterminer, sur $\mathbb{C}$, des formes réduites privilégiées (sous forme de matrices diagonales par blocs du type $\lambda I + J$, où J est une matrice de Jordan), valide de façon plus gérénale dès que le polynôme caractéristique est scindé sur $\mathbb{K}$ (réduction de Jordan).
Cette décomposition n'est pas simple à obtenir (même si cela reste accessible avec les outils de classe préparatoire). Nous nous contentons d'étudier une propriété intermédiaire un peu dégradée, qui est l'existence d'une base $\mathcal{B}$ de E dans laquelle la matrice de u est triangulaire (c'est le cas de la réduction de Jordan).
D'autres types de réduction existent, valides sous des conditions sur $\mathbb{K}$ plus générales, par exemple la décomposition de Forbenius qui donne une forme réduite diagonale par blocs, avec des blocs diagonaux égaux à des « matrices compagnons » de certains polynômes.
E désigne toujours un $\mathbb{K}$-espace vectoriel de dimension finie.

III.1 Trigonalisabilité et caractérisations algébriques

Définition 9.3.1 – Trigonalisabilité

1. On dit qu'un endomorphisme $u \in \mathcal{L}(E)$ s'il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(u) \in \mathcal{T}^+(\mathbb{K})$.
2. On dit qu'une matrice $M \in \mathcal{M}_n(\mathbb{K})$ est trigonalisable s'il existe une matrice $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $P^{-1}MP \in \mathcal{T}^+(\mathbb{K})$. En notant T cette matrice, on a alors $M = PTP^{-1}$.

Les deux définitions (vectorielle et matricielle) se ramènent l'une à l'autre, comme dans le cas de la diagonalisabilité.

Proposition 9.3.2 – Trigonalisabilité d'un endomorphisme versus matrice

Soit $u \in \mathcal{L}(E)$ et $\mathcal{B}$ une base de E

1. u est trigonalisable si et seulement si $\mathrm{Mat}_{\mathcal{B}}(u)$ est trigonalisable.
2. En particulier, M est trigonalisable si et seulement si l'endomorphisme $f_M \in \mathcal{L}(\mathbb{K}^n)$ canoniquement associé à M est trigonalisable.

Proposition 9.3.3 – Caractérisation de la trigonalisabilité par stabilisation d'un drapeau

Un endomorphisme $u \in \mathcal{L}(E)$ est trigonalisable si et seulement s'il existe une suite

$$\{0\} = E_0 \subset E_1 \subset E_2 \subset \cdots \subset E_n = E$$

de sous-espaces vectoriels, où $n = \dim(E)$, et vérifiant :

- pour tout $k \in [\![0, n]\!]$, $\dim(E_k) = k$ (ainsi la suite est strictement croissante).
- E_k est stable par u

Une telle suite d'espaces vectoriels est parfois appelée drapeau de l'espace E.

La pratique de la trigonalisation est quelque chose de technique et compliqué, vous serez nécessairement guidé dans la démarche. En revanche, il existe une condition très simple de trigonalisabilité. Cette CNS a évidemment un analogue matriciel.

Théorème 9.3.4 – CNS de trigonalisabilité

Soit $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes:

(i) u est trigonalisable ;

(ii) Le polynôme caractéristique χ_u de u est scindé ;

(iii) u admet un polynôme annulateur (non nul) scindé ;

(iv) le polynôme minimal de u est scindé.

◁ **Éléments de preuve.**

- $(i) \implies (ii)$: exprimer le polynôme caractéristique à partir d'une base de trigonalisation.
- $(ii) \implies (iii)$: d'après Cayley-Hamilton
- $(iii) \implies (iv)$: car μ_u divise alors un polynôme scindé
- $(iv) \implies (i)$: faire une récurrence, en distinguant suivant que μ_u est constant ou non. S'il n'est pas constant, il a une racine, qui est alors une valeur propre de u. Considérer une base commençant par un vecteur propre associé.

▷

Remarque 9.3.5

On peut se dispenser de l'utilisation du théorème de Cayley-Hamilton ici, en remarquant que $(ii) \implies (i)$ peut se faire comme $(iv) \implies (i)$, et que $(i) \implies (iii)$ peut se faire directement, en considérant le polynôme P simplement scindé dont les racines sont les coefficients diagonaux d'une matrice triangulaire T représentant u, et en remarquant que $P(T)$ est strictement triangulaire supérieure. On a alors $(P(T))^n = 0$, donc P^n est un polynôme annulateur scindé de T.

Corollaire 9.3.6

Toute matrice de $\mathcal{M}_n(\mathbb{C})$ est trigonalisable

◁ **Éléments de preuve.**

C'est d'Alembert-Gauss ▷

De la même façon, toutes les propriétés qui suivent, utilisant une hypothèse de scindement du polynôme caractéristique, sont systématiquement vraies sur $\mathbb{C}$.

III.2 Endomorphismes nilpotents

Nous avons déjà eu l'occasion de justifier que les endomorphismes nilpotents u sont caractérisés par leur polynôme minimal, égal à X^d, où d est l'indice de nilpotence de u.
Cela se traduit également sur le polynôme caractéristique, et sur les propriétés de trigonalisation.
L'espace E est toujours un $\mathbb{K}$-ev de dimension finie égale à n.

Théorème 9.3.7 – Caractérisation des endomorphismes nilpotents

Soit $u \in \mathcal{L}(E)$ un endomorphisme. Les propositions suivantes sont équivalentes:

(i) u est nilpotent

(ii) u admet un polynôme annulateur de la forme X^k

(iii) μ_u est de la forme X^d, d étant alors l'indice de nilpotence

(iv) u est trigonalisable, et $\mathrm{Sp}(u) = \{0\}$

(v) il existe une base $\mathcal{B}$ de E dans laquelle $\mathrm{Mat}_{\mathcal{B}}(u)$ est triangulaire supérieure stricte.

(vi) $\chi_u = X^n$, où $n = \dim(E)$

Corollaire 9.3.8 – Majoration de l'indice de nilpotence revisité

Soit $u \in \mathcal{L}(E)$ nilpotent. L'indice de nilpotence d de u vérifie $d \leqslant \dim(E)$.

III.3 Sous-espaces caractéristiques

Nous allons maintenant obtenir une forme un peu plus précise de la trigonalisation obtenue pour les endomorphismes à polynôme caractéristique scindé. Nous allons pour cela nous ramener à des endomorphismes nilpotents grâce à une factorisation adéquate de l'espace.

Proposition 9.3.9 – Décomposition de E en somme de sous-espaces caractéristiques

Soit $u \in \mathcal{L}(E)$ tel que χ_u soit scindé. On note

$$\chi_u = \prod_{\lambda \in \mathrm{Sp}(u)} (X - \lambda)^{m_\lambda},$$

où m_λ est la multiplicité algébrique de la valeur propre λ. Alors,

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} N_\lambda(u), \qquad \text{où} \qquad N_\lambda(u) = \mathrm{Ker}((u - \lambda \mathrm{id})^{m_\lambda}).$$

Le sous-espace $N_\lambda(u)$ (parfois aussi noté $\mathrm{SEC}_\lambda(u)$) est appelé sous-espace caractéristique associé à la valeur propre λ de u.

Proposition 9.3.10 – Endomorphismes induits sur les $N_\lambda(u)$

Avec les notations de la proposition précédentes :

1. Les sous-espaces caractéristiques $N_\lambda(u)$ sont stables par u ;
2. En notant u_λ l'endomorphisme induit par u sur $N_\lambda(u)$, $u_\lambda - \lambda \mathrm{id}$ est nilpotent.
3. Pour tout $\lambda \in \mathrm{Sp}(u)$, $\chi_{u_\lambda} = (X - \lambda)^{m_\lambda}$
4. En particulier, $\dim N_\lambda(u) = m_\lambda$.

◁ **Éléments de preuve.**
Les points 1 et 2 sont immédiats. On déduit du 2 que le polynôme caractéristique de $(u_\lambda - \lambda \mathrm{id})$ est $(X-\lambda)^{d_\lambda}$, où $d_\lambda = \dim N_\lambda(u)$.
Utiliser la divisibilité de χ_u par χ_{u_λ} et la décomposition de E en somme d'espaces caractéristiques pour conclure simultanément les points 3 et 4. ▷

Corollaire 9.3.11

Soit u un endomorphisme tel que $\chi_u = \prod_{i=1}^{k}(X-\lambda_i)^{m_i}$.
Alors il existe une base $\mathcal{B}$ telle que $\mathrm{Mat}_{\mathcal{B}}(u)$ soit diagonale par blocs, de blocs diagonaux $A_1, \dots, A_k$ vérifiant :

(i) la matrice A_i est de taille m_i ;

(ii) la matrice A_i est triangulaire à coefficients diagonaux tous égaux à λ_i.

On remarquera que cette matrice est en particulier triangulaire, et cela donne une forme possible un peu plus précise de la trigonalisation de u.

Corollaire 9.3.12

Soit $M \in \mathcal{M}_n(\mathbb{K})$ à polynôme caractéristique scindé $\chi_M = \prod_{i=1}^{k}(X-\lambda_i)^{m_i}$.
Alors M est semblable à une matrice T diagonale par blocs, de blocs diagonaux $A_1, \dots, A_k$ vérifiant :

(i) la matrice A_i est de taille m_i ;

(ii) la matrice A_i est triangulaire à coefficients diagonaux tous égaux à λ_i.

' On termine cette section par un théorème, grand classique des concours, mais à la marge du programme.

Théorème 9.3.13 – Décomposition de Dunford, HP

Soit $u \in \mathcal{L}(E)$ un endomorphisme à polynôme caractéristique scindé. Alors il existe deux endomorphismes n et d de E tels que :

(i) $u = n + d$

(ii) n est nilpotent

(iii) d est diagonalisable

(iv) $n \circ d = d \circ n$.

Évidemment, ce théorème se décline aussi en version matricielle.
Le théorème de Dunford assure aussi l'unicité de cette décomposition, ainsi que le fait que cela caractérise le scindement du polynôme caractéristique.

10

Suites et séries de fonctions

Dans ce chapitre, on étudie les suites et séries de fonctions. De façon naturelle, on peut définir, si elle existe, la valeur point par point (c'est-à-dire la limite de $f_n(x)$) à x fixé, pour une suite (f_n)). On peut alors se poser la question de savoir si certaines propriétés des suites f_n passent bien à la limite (par exemple la continuité, la dérivabilité etc) et si cela revient au même de faire certaines opérations terme à terme ou sur la somme totale, par exemple si la somme des intégrales est égale à l'intégrale de la somme (ou de même pour les limites). Il s'agit donc de propriété d'interversion de certains symboles, définis à base de limites. En général ces interversions ne seront pas possibles et il faudra des modes de convergence plus contraignants.

I Modes de convergences

I.1 Convergence simple

Soit E et F deux espaces vectoriels de dimension finie. On munit F de sa topologie naturelle d'e.v.n., associée à une norme quelconque $\|\cdot\|_F$. Soit $A \subset E$. On considère une suite d'applications $(f_n)_{n\in\mathbb{N}}$ de A dans E.

Dans certains énoncés, E devra également être muni de sa structure topologique naturelle d'e.v.n., notamment quand il sera question de continuité. Nous noterons dans ce cas $\|\cdot\|_E$ une norme (arbitrairement choisie) de E.

Comme on est en dimension finie, le choix des normes sur E et F n'a pas d'influence sur les propriétés de convergence et de continuité.

Définition 10.1.1 – Convergence simple

La suite (f_n) converge simplement vers une fonction $f : A \to F$ si pour tout $x \in A$, la suite $(f_n(x))_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ converge vers $f(x)$. Autrement dit :

$$\forall x \in A,\ \forall \varepsilon > 0,\quad \exists N_x \in \mathbb{N},\ \forall n \in \mathbb{N},\ n \geqslant N_x \Longrightarrow \|f_n(x) - f(x)\|_F \leqslant \varepsilon.$$

Proposition 10.1.2 – Unicité de la limite simple

En cas d'existence, la fonction f est unique.

Remarque 10.1.3

1. La propriété de convergence simple, ainsi que la valeur de la limite simple f, ne dépend par de la norme sur F
2. L'espace E n'a pas besoin d'être muni d'une norme pour définir la convergence simple.

3. Les règles usuelles sur les limites de suites à valeurs dans F s'appliquent (sommes, produits si pertinent etc). Par exemple, si (f_n) et (g_n) admettent des limites simples f et g, alors $(f_n + \lambda g_n)$ converge simplement vers $f + \lambda g$.
4. L'indice x pour désigner le rang N_x dans la définition est indiqué pour insister sur la dépendance de N par rapport à x (il dépend aussi de ε, mais ce point est plus « évident », alors que la dépendance ou non par rapport à x est justement la subtilité qui différentie la convergence simple de la convergence uniforme)

Exemples 10.1.4

Étudier la convergence simple de la suite (f_n) où, pour tout $n \in \mathbb{N}$, $f_n : \mathbb{R}_+ \to \mathbb{R}$ est définie par :

1. $f_n(x) = \mathrm{e}^{-nx}$;
2. $f_n(x) = nx\mathrm{e}^{-nx}$.
3. $f_n(x) = n\mathrm{e}^{-nx}$

Remarque 10.1.5

Comme on l'a fait remarquer dans le préambule, certaines propriétés ne se conservent pas. Ici, dans l'exemple 1, la continuité des f_n n'est par exemple pas préservée au passage à la limite simple f. Il faut, pour que ce soit le cas, un mode de convergence plus fort

I.2 Convergence uniforme

Définition 10.1.6 – Convergence uniforme

1. On dit que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f (sur A) si :
$$\forall \varepsilon > 0, \quad \exists N \in \mathbb{N}, \quad \forall x \in A, \ \forall n \in \mathbb{N} n \geqslant N \Longrightarrow \|f_n(x) - f(x)\|_F \leqslant \varepsilon.$$
2. Soit $B \subset A$. On dit que (f_n) converge uniformément sur B si $((f_n)_{|B})$ converge uniformément, donc si
$$\forall \varepsilon > 0, \quad \exists N \in \mathbb{N}, \quad \forall x \in B, \ \forall n \in \mathbb{N}, \ n \geqslant N \Longrightarrow \|f_n(x) - f(x)\|_F \leqslant \varepsilon.$$

Remarque 10.1.7

1. La différence essentielle avec la convergence simple est l'indépendance de N par rapport à x. La convergence est donc « aussi rapide » en chaque point.
2. Graphiquement, pour des fonctions réelles, cela signifie qu'àpcr, la courbe de f_n est dans un « tube » de rayon ε autour de la courbe de f.
3. La quantité $\|f_n - f\|_\infty$ étant bien définie dans $\overline{\mathbb{R}}$ (égale à $+\infty$ si $f - f_n$ n'est pas bornée), la définition peut se réexprimer de la façon suivante :
$$\|f_n - f\|_\infty \underset{n\to 0}{\longrightarrow} 0,$$

Proposition 10.1.8 – CVU par majoration uniforme

S'il existe une suite $(M_n)_{n\in\mathbb{N}}$ (indépendante de x) telle que pour tout $x \in A$,
$$\|f_n(x) - f(x)\|_F \leqslant M_n,$$
et telle que $\lim M_n = 0$, alors (f_n) converge uniformément vers f.

Proposition 10.1.9 – Limite simple / limite uniforme

Si (f_n) converge uniformément vers f, alors elle converge aussi simplement vers f.

Ainsi, la limite uniforme, si elle existe, est égale à la limite simple.

Proposition 10.1.10 – Caractérisation de la CVU par convergence dans un e.v.n.

Soit $f_n, f : A \to F$. Les propositions suivantes sont équivalentes:

(i) (f_n) converge uniformément vers f sur A

(ii) $(f_n - f)$ est bornée à partir d'un certain rang n_0, et $(f_n - f)_{n \geqslant n_0}$ converge vers 0 dans $\mathcal{B}(A, F)$.

Méthode 10.1.11 – Étudier la convergence uniforme

- Étudier la limite simple pour trouver le seul candidat f à être la limite uniforme.
- Majorer $\|f_n(x) - f(x)\|$ indépendamment de x par une suite numérique qui tend vers 0.

Exemples 10.1.12

1. Parmi les exemples précédents, y en a-t-il qui CVU sur $\mathbb{R}_+$? sur $[a, +\infty[$, a>0 ?
2. Étudier la convergence uniforme sur $\mathbb{R}$ des suites définies par :
   - $f_n(x) = \mathrm{e}^{-(t+n)^2}$
   - $f_n(x) = \mathrm{e}^{-\left(t+\frac{1}{n}\right)^2}$.
3. Étudier la convergence uniforme sur $B(0, r)$ de la suite $(A(z)^n)_{n \in \mathbb{N}}$, où $r \in]0, 1[$ et $A : \mathbb{C} \to \mathcal{M}_n(\mathbb{C})$ est définie par

$$A(z) = \begin{pmatrix} z & 1 & 0 & \dots & 0 \\ 0 & z & 1 & & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & & \ddots & z & 1 \\ 0 & 0 & \dots & 0 & z \end{pmatrix}$$

Proposition 10.1.13 – CN de CVU

Supposons que (f_n) converge uniformément vers f sur A. Alors pour tout $(x_n)_{n \in A} \in A^{\mathbb{N}}$, tel que $x_n \longrightarrow a \in A$, $f_n(x_n) - f(x_n) \longrightarrow 0$

Méthode 10.1.14 – Montrer qu'une suite ne converge pas absolument

Utiliser la contraposée de la propriété précédente : pour montrer que (f_n) ne converge pas uniformément vers f, il suffit de trouver $x_n \in A^{\mathbb{N}}$ tel que $f_n(x_n) - f(x_n)$ ne converge pas vers 0

Exemples 10.1.15

1. Reprendre les exemples précédents.
2. Étudier la CVU sur $\mathbb{R}$ de (f_n) définie par $f_n(x) = \cos\left(\frac{x}{n}\right)$.

I.3 Convergence uniforme locale

Comme on le verra un peu plus loin, la convergence uniforme permet de conserver les propriétés de continuité. Comme la continuité n'est qu'une propriété locale, et non globale, il n'est pas nécessaire d'avoir la convergence uniforme sur tout le domaine, mais de l'avoir au voisinage de tout point en lequel on veut établir la continuité.

Définition 10.1.16 – Convergence uniforme locale

Soit $a \in A$. On dit que (f_n) converge uniformément vers f au voisinage du point a s'il existe un voisinage $V \in \mathcal{V}(a)$ tel que (f_n) converge uniformément vers f sur $V \cap A$

Exemple 10.1.17

Montrer que les exemples 10.1.4 sont uniformément convergents au voisinage de tout point de $\mathbb{R}_+^*$, mais par uniformément continue sur $\mathbb{R}_+^*$.

Attention au fait que la convergence uniforme sur une partie B de A n'implique pas nécessairement la convergence uniforme locale au voisinage de tout point de B :

Exemple 10.1.18

Soit $f_n : x \mapsto \begin{cases} 1 & \text{si } x < 0 \\ \mathrm{e}^{-nt} & \text{si } x \geqslant 0 \end{cases}$.
Montrer que (f_n) CVU sur $]-\infty, 0]$, mais ne converge pas uniformément au voisinage de 0.

On peut préciser un peu les choses :

Proposition 10.1.19 – CVU locale versus CVU globale

Soit $B \subset A$. On suppose que (f_n) converge uniformément vers f sur B.

1. Alors (f_n) converge uniformément au voisinage de tout point de $\mathring{B}$, et plus généralement, de l'intérieur relatif de B dans A (*i.e.* le plus grand ouvert relatif de A inclus dans B).
2. En particulier, si B est un ouvert relatif de A, alors (f_n) converge uniformément au voisinage de tout point de B.

Dans le cas d'une fonction définie sur un intervalle I de $\mathbb{R}$, on peut se contenter d'étudier la convergence uniforme sur tout segment inclus dans I.

Proposition 10.1.20 – CVU sur un intervalle de $\mathbb{R}$

Soit $f_n : I \to F$. Les propositions suivantes sont équivalentes:

(i) (f_n) converge uniformément au voisinage de tout point de I.

(ii) (f_n) converge uniformément sur tout segment $[a, b] \subset I$;

II Régularité des limites de suites de fonctions

Cette section traite essentiellement du passage à la limite d'un certain nombre de propriétés et constructions : la continuité, l'intégration, la dérivation. Ces propriétés et constructions s'expriment ou se définissent toutes en fonction de limites (ou bornes supérieures). Dans tous les cas, il s'agit donc ici d'établir un certain nombre de propriété d'interversion de limites.

II.1 Continuité d'une limite uniforme

Théorème 10.2.1 – Préservation de la continuité par CVU

Soit (f_n) une suite de fonctions $A \to F$.

1. Soit $a \in A$. Si (f_n) converge uniformément vers f au voisinage de a, et si les (f_n) sont continues en a, alors f est continue en a.
2. Soit B un ouvert relatif de A. Si (f_n) converge uniformément vers f sur B et si les f_n sont continues sur B, alors f est continue sur B

3. En particulier si (f_n) converge uniformément vers f (sur A entier) et si les f_n sont continues, alors f est continue (sur A).

◁ **Éléments de preuve.**

Le premier point implique facilement les deux autres. Pour montrer 1, poser un ε et décomposer

$$f(x) - f(a) = f(x) - f_n(x) + f_n(x) - f_n(a) + f_n(a) - f(a),$$

sur une boule sur laquelle exploiter la convergence uniforme. ▷

Exemple 10.2.2 – Fonction de Bolzano

On considère la suite (f_n) de fonctions $[0,1] \to \mathbb{R}$ affines par morceaux, définie par $f_0 = \mathrm{id}_{[0,1]}$, et, f_n étant définie continue, affine sur chaque intervalle $[\frac{k}{3^n}, \frac{k+1}{3^n}]$, on définit f_{n+1} continue et affine sur chaque intervalle $[\frac{k}{3^{n+1}}, \frac{k+1}{3^{n+1}}]$, prenant les mêmes valeurs que f_n sur les $\frac{3k}{3^{n+1}}$, et telle que

$$f_{n+1}\left(\frac{3k+1}{3^{n+1}}\right) = f_n\left(\frac{3k+2}{3^{n+1}}\right) \qquad \text{et} \qquad f_{n+1}\left(\frac{3k+2}{3^{n+1}}\right) = f_n\left(\frac{3k+1}{3^{n+1}}\right).$$

Montrer que (f_n) converge vers une fonction f et que f est continue.
En représentant les fonctions (f_n), on se doute bien que les irrégularités de plus en plus nombreuses vont être un frein à la dérivabilité. De fait, on peut montrer que f est une fonction continue sur $[0,1]$, mais dérivable en aucun point de $[0,1]$. C'est historiquement le premier exemple d'une fonction continue nulle part dérivable, découvert par Bolzano vers 1830.

Avertissement 10.2.3

Si la propriété de CVU ne porte pas sur A tout entier, attention à l'hypothèse d'ouverture. En particulier, si on travaille dans $\mathbb{R}$, et que A et B sont des intervalles, il faut faire attention aux bornes.

Exemple 10.2.4

La suite définie par

$$f_n(x) = \begin{cases} 0 & \text{si } x \leqslant -\frac{1}{n} \\ nx+1 & \text{si } x \in]-\frac{1}{n}, 0[\\ 1 & \text{si } x \geqslant 0 \end{cases}$$

est uniformément convergente sur $[0, +\infty[$, et les f_n sont toutes continues sur cet intervalle. Pourtant, la limite $\mathbb{1}_{\mathbb{R}_+}$ n'est pas continue en 0.
On doit se ramener à une convergence uniforme sur un ouvert (relatif, mais ici, $A = \mathbb{R}$ tout entier donc c'est la même chose). Donc tout ce qu'on peut obtenir ici, c'est la continuité sur $\mathbb{R}_+^*$, qui est bien vérifiée.

Méthode 10.2.5 – Étudier la continuité d'une limite

Pour étudier la continuité d'une limite d'une suite de fonctions définies sur A, on étudiera d'abord la CVU sur A. S'il n'y a pas CVU sur A, on pourra s'intéresser à la CV locale, en décomposant A comme union d'ouverts relatifs, ou en se ramenant à des segments si A est un intervalle réel.

II.2 Double-limite

La propriété de continuité correspond à une interversion $\lim_{x\to a}$ / $\lim_{n\to+\infty}$:

Si (f_n) CVU au voisinage de $a \in A$ en lequel les $f_n)$ sont tous continus, alors

$$\lim_{x\to a}\lim_{n\to+\infty} f_n(x) = \lim_{n\to+\infty}\lim_{x\to a} f_n(x).$$

On peut se demander si une telle propriété subsiste au bord du domaine, lorsque $a \in \overline{A}\backslash A$. Cela reste valide, mais un peu plus délicat, car la valeur de la double-limite est un peu moins accessible.

Théorème 10.2.6 – Théorème de la double-limite

Soit $f_n : A \to F$ convergeant uniformément vers f, et $a \in \overline{A}$ (éventuellement infini si $E = \mathbb{R}$). On suppose que pour tout $n \in \mathbb{N}$, $f_n(x) \underset{x\to a}{\longrightarrow} \ell_n$.
Alors (ℓ_n) admet une limite $\ell \in F$, et

$$\lim_{x\to a} f(x) = \ell.$$

◁ **Éléments de preuve.**

La démonstration est hors-programme. Le point délicat est de prouver la convergence de (ℓ_n), car une fois celle-ci acquise, on peut se ramener au théorème de continuité par prolongement.

On peut en fait assez facilement montrer d'abord que (ℓ_n) est bornée (en commençant par majorer $\|f_n - f_{n_0}\|_\infty$, pour un n_0 convenable), pour trouver une valeur d'adhérence, puis par ε, montrer que $f(x) \longrightarrow \ell$. La convergence de la suite (ℓ_n) s'en déduit par une propriété d'unicité de la limite de f, amenant l'unicité de la valeur d'adhérence de (ℓ_n). est ▷

Ainsi, l'existence de limites (finies) ℓ_n pour tous les f_n assure l'existence d'une limite pour f, égale à la limite des ell_n. Ou encore, sous ces hypothèses :

$$\lim_{n\to+\infty}\lim_{x\to a} f_n(x) = \lim_{x\to a}\lim_{n\to+\infty} f_n(x).$$

C'est donc un théorème d'interversion de limites, comme le théorème de continuité.

Avertissement 10.2.7

Sans hypothèse de CVU, une telle interversion peut être fausse, même si toutes les limites existent.

Exemple 10.2.8

La suite définie par $f_n(x) = \mathrm{e}^{-nx}$ sur $\mathbb{R}_+^*$ vérifie :

$$\lim_{x\to 0^+}\lim_{n\to+\infty} f_n(x) = 0 \qquad \text{et} \qquad \lim_{n\to+\infty}\lim_{x\to 0^+} f_n(x) = 1.$$

II.3 Interversion limite-intégrale

Nous savons déjà que la limite uniforme d'une suite de fonctions f_n continues sur un segment $[a,b]$ est continue, donc intégrable au sens de Riemann. On peut légitimement se demander si son intégrale est obtenue en passant les intégrales des f_n à la limite. Nous montrons dans un premier temps un résultat un peu plus fort.

Théorème 10.2.9 – Primitivation d'une limite uniforme de fonctions continues

Soit (f_n) une suite de fonctions d'un segment $[a,b]$ dans un e.v.n. E de dimension finie. On suppose que :

- les f_n sont continues ;
- (f_n) converge uniformément vers f sur $[a,b]$

On définit la suite

$$F_n : x \mapsto \int_a^x f_n(x)\,\mathrm{d}x, \qquad \text{et} \qquad F : x \mapsto \int_a^x f(x)\,\mathrm{d}x.$$

Alors (F_n) converge uniformément vers F sur $[a,b]$.

◁ **Éléments de preuve.**

Remarquer que $\left\| \int_a^x f_n(t) - f(t)\ \mathrm{d}t \right\|_E \leqslant |b-a| \cdot \|f_n - f\|_\infty$ ▷

Corollaire 10.2.10 – Interversion limite uniforme/intégrale

Soit (f_n) une suite de fonctions d'un segment $[a,b]$ dans un e.v.n. E de dimension finie. On suppose que :

- les f_n sont continues ;
- (f_n) converge uniformément vers f sur $[a,b]$

Alors, f est intégrable sur $[a,b]$ et

$$\int_a^b f(t)\ \mathrm{d}t = \lim_{n\to+\infty} \int_a^b f_n(t)\ \mathrm{d}t.$$

◁ **Éléments de preuve.**

Prendre la limite simple en $x = b$ dans le théorème précédent. ▷

Remarque 10.2.11

Ce théorème affirme donc que sous les hypothèses de continuité et de convergence uniforme, on peut intervertir lim et $\displaystyle\int$, ou, en d'autres termes, « passer à la limite sous l'intégrale » :

$$\int_a^b \lim_{n\to+\infty} f_n(t)\ \mathrm{d}t = \lim_{n\to+\infty} \int_a^b f_n(t)\ \mathrm{d}t.$$

Exemple 10.2.12

Soit f la fonction de Bolzano définie dans un exemple précédent. Calculer $\displaystyle\int_0^1 f(x)\ \mathrm{d}x$.

Avertissement 10.2.13

- L'hypothèse de convergence uniforme est importante.
- L'hypothèse de continuité est moins importante, le résultat restant vrai sous la seule hypothèse d'intégrabilité (mais l'intégrabilité de la limite est un peu plus délicate à démontrer)

Exemples 10.2.14

Soit f_n définie sur $[0,1]$, pour $n \geqslant 1$, par

$$f_n(x) = \begin{cases} n^2 x & \text{si } x \in [0, \frac{1}{2n}] \\ n - n^2 x & \text{si } x \in]\frac{1}{2n}, \frac{1}{n}] \\ 0 & \text{si } x \in]\frac{1}{n}, 1]. \end{cases}$$

Comparer $\displaystyle\lim_{n\to+\infty} \int_0^1 f_n(t)\ \mathrm{d}t$ et $\displaystyle\int_0^1 \lim_{n\to+\infty} f_n(t)\ \mathrm{d}t$.

Remarque 10.2.15

Le théorème d'interversion limite/intégrale en cas de convergence uniforme peut être vu comme un cas particulier assez simple du théorème de convergence dominé, appliqué sur un segment à la suite $(f_n - f)$, dominée à partir d'un certain rang par une fonction constante, intégrable sur un segment.

II.4 Dérivabilité et dérivée d'une limite uniforme

Comme pour la continuité, la dérivabilité d'une limite de suite de fonctions (f_n) dérivables n'est pas nécessairement dérivable. Elle peut aussi être dérivable, mais de dérivée différente de la limite (simple) de la suite (f'_n) des fonctions dérivées.

Exemples 10.2.16

1. Soit f_n définie pour tout $n \in \mathbb{N}^*$ sur $[0,1]$ par
$$f_n(x) = \begin{cases} x^{1+\frac{1}{n}} \sin\left(\frac{1}{x}\right) & \text{si } x \in]0,1] \\ 0 & \text{sinon.} \end{cases}.$$
Montrer que (f_n) converge uniformément vers une fonction f à déterminer. Montrer que pour tout $n \in \mathbb{N}^*$, f_n est dérivable sur $[0,1]$, mais pas f
2. Soit f_n définie pour tout $n \in \mathbb{N}^*$ sur $[0,1]$ par $f_n(x) = \frac{x^n}{n}$.
Montrer que (f_n) converge uniformément vers une limite f, dérivable, mais que (f'_n) ne converge pas vers f' sur $[0,1]$.

Ainsi, même une hypothèse de convergence uniforme, portant sur la suite (f_n), n'est pas suffisante. De fait, il faut bien une hypothèse de convergence uniforme pour pouvoir intervertir limite et dérivation, mais elle doit porter sur la suite des dérivées.

Théorème 10.2.17 – Dérivation d'une suite de fonctions

Soit (f_n) une suite de fonctions définies d'un intervalle $I \subset \mathbb{R}$ dans un e.v.n. F de dimension finie. On suppose que :

(i) les f_n sont de classe $\mathcal{C}^1$ sur I ;

(ii) la suite (f_n) converge simplement vers f

(iii) la suite (f'_n) converge uniformément sur I vers une application $g : I \to F$.

Alors (f_n) converge uniformément vers f sur I, est de classe $\mathcal{C}^1$ sur I, et pour tout $x \in I$,
$$f'(x) = g(x) = \lim_{n \to +\infty} f'_n(x).$$

◁ **Éléments de preuve.**

Appliquer le théorème de primitivation à la suite (f'_n), pour comparer f et $x \mapsto f(a) + \int_a^x g(x)$. ▷

Par récurrence, on obtient facilement la version pour la classe $\mathcal{C}^p$

Théorème 10.2.18 – Classe $\mathcal{C}^p$ d'une limite

Soit (f_n) une suite de fonctions définies d'un intervalle $I \subset \mathbb{R}$ dans un e.v.n. F de dimension finie, et $p \in \mathbb{N}$. On suppose que :

(i) les f_n sont de classe $\mathcal{C}^p$ sur I ;

(ii) pour tout $k \in [\![0, p-1]\!]$ la suite $(f_n^{(k)})$ converge simplement ;

(iii) la suite $(f_n^{(p)})$ converge uniformément sur I.

Alors :

- La limite f de (f_n) est de classe $\mathcal{C}^p$, (resp sur tout $[a,b] \subset I$) ;
- pour tout $k \in [\![0,p]\!]$, $(f_n^{(k)})$ converge uniformément sur I vers $f^{(k)}$.

En particulier, pour tout $k \in [\![0,p]\!]$, pour tout $x \in I$,

$$f^{(k)}(x) = \lim_{n\to+\infty} f_n^{(k)}(x).$$

III Séries de fonctions

Le but de cette section est d'adapter les résultats obtenus pour les suites de fonctions au cas des séries de fonctions. Comme dans le cas numérique, on peut obtenir certains critères de convergence simple ou uniforme ne nécessitant pas de connaître la limite. Cela donne une bien meilleure efficacité que le point de vue des suites.
On pourra en bénéficier pour l'étude des suites de fonctions, puisque, comme dans le cas numérique, l'étude d'une limite de suite de fonctions (f_n) se ramène à l'étude de la série télescopique $\sum f_n - f_{n-1}$.

III.1 Modes de convergence

On se donne E un espace vectoriel de dimension finie, F un e.v.n. de dimension finie, et A une partie de E, ainsi qu'une suite $(f_n)_{n\in\mathbb{N}}$ de fonctions de A dans F.
On note, pour tout $x \in A$, et tout $n \in \mathbb{N}$,

$$S_n(x) = \sum_{k=0}^{n} f_k(x)$$

les sommes partielles de la série $\sum f_n$. Soit également $B \subset A$.

Définition 10.3.1 – Convergence simple, convergence uniforme

- La série $\sum f_n$ converge simplement si la suite des sommes partielles $(S_n)_{n\in\mathbb{N}}$ converge simplement.
- La série $\sum f_n$ converge uniformément sur B si la suite des sommes partielles $(S_n)_{n\in\mathbb{N}}$ converge uniformément sur B.

En cas de convergence simple sur B, on notera dans la suite du chapitre

$$\forall n \in \mathbb{N},\ \forall x \in B,\quad R_n(x) = \sum_{k=n+1}^{+\infty} f_k(x).$$

les restes de la série $\sum f_n$.

Proposition 10.3.2 – Caractérisation de la CVU par le reste

On suppose que $\sum f_n$ converge simplement sur B. Alors $\sum f_n$ converge uniformément sur B si et seulement si R_n converge uniformément vers 0 sur B.

Il suffit donc de majorer $R_n(x)$ par une suite indépendante de x, et tendant vers 0. Le plus naturel est d'essayer de majorer la série en normes, ce qui nous ramnène à la définition suivante

Définition 10.3.3 – Convergence normale

La série $\sum f_n$ converge normalement (sur A) si pour tout $n \in \mathbb{N}$, f_n est bornée, et si $\sum \|f_n\|_\infty$ converge.

Remarque 10.3.4

1. Autrement dit, la convergence normale correspond à la convergence absolue de la série dans $(\mathcal{B}(A,F), \|\cdot\|_\infty)$.
2. En pratique, la convergence uniforme, ou la convergence normale, seront souvent obtenues

sur des sous-parties B adéquates, comme dans le cas des suites de fonctions. Ce sera en général suffisant pour obtenir les propriétés de continuité.

Théorème 10.3.5 – CVN implique CVU

On suppose que les f_n sont bornées sur A. Alors si $\sum f_n$ converge normalement sur A, elle converge uniformément sur A.

◁ **Éléments de preuve.**

Le reste de la série des normes est indépendant de x, tend vers 0, et majore $\|R_n(x)\|$. ▷

Remarque 10.3.6

Il suffit, pour utiliser ce théorème, que les f_n soient bornées à partie d'un certain rang n_0, et que $\sum\limits_{n \geqslant n_0} f_n$ soit normalement convergente.

Exemples 10.3.7

1. Montrer que $\sum \frac{x^n}{n!}$ converge uniformément sur $[-A, A]$, pour tout $A \in \mathbb{R}_+$.
2. Montrer que $\sum\limits_{n \geqslant 1} \frac{1}{x^2+n^2}$ converge uniformément sur $\mathbb{R}$.
3. Soit ζ la fonction de Riemann, définie par

$$\zeta(z) = \sum_{n=1}^{+\infty} \frac{1}{n^z}.$$

Montrer que ζ est définie sur $D = \{z \in \mathbb{C} \mid |z| > 1\}$, et que la somme définissant ζ est uniformément convergente sur $D_a = \{\{z \in \mathbb{C} \mid |z| > 1\}$, pour tout $a > 1$.

Méthode 10.3.8 – Montrer qu'une CV n'est pas normale, ou pas uniforme

1. Pour montrer que $\sum f_n$ ne CV pas normalement sur A,
   - D'abord, regarder si les f_n sont bornées (au moins àpcr) ;
   - si c'est le cas, trouver $(x_n) \in A^{\mathbb{N}}$ telle que $\sum f_n(x_n)$ diverge.
   - On peut aussi calculer explicitement $\|f_n\|_\infty$ (c'est dans certains cas évident, dans d'autres, on peut faire une étude de fonctions) pour se ramener à une série numérique.
2. Pour montrer que $\sum f_n$ ne CV pas uniformément, c'est un peu plus délicat.
   - Trouver $(x_n)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ tel que $(R_n(x_n))$ ne tend pas vers 0.
   - On peut aussi contredire l'une des propriétés qui devrait être vérifiée en cas de CVU (par exemple la continuité de la somme).

Exemples 10.3.9

1. $\sum \frac{x^n}{n!}$ ne converge pas normalement sur $\mathbb{R}$.
2. $\sum \frac{x^n}{n}$ ne converge pas normalement sur $]-1, 1[$.
3. $\sum x e^{-nx}$ converge simplement sur $\mathbb{R}_+$, mais pas normalement.
4. Les trois séries précédentes ne sont pas non plus uniformément convergentes sur les intervalles donnés.
5. La somme définissant ζ n'est pas uniformément sur $]1, +\infty[$.

Avertissement 10.3.10

La CVN n'est pas une condition nécessaire à la CVU. Une série peut converger uniformément sur un ensemble A sans y converger normalement. Nous donnons des exemples, après avoir donné des méthodes pour gérer cette situation.

Méthode 10.3.11 – Justifier une CVU sans CVN

1. Majorer $R_n(x)$ par les moyens du bord (éventuellement voir des compensations entre des termes consécutifs)
2. Un cas particulier important est le cas des séries alternées : Si pour tout $x \in \mathbb{N}$, $\sum f_n(x)$ est une série alternée, on est ramené à l'étude de la CVU de f_n vers 0.
3. Une autre méthode consiste à faire une transformation d'Abel, ce qui s'apparente à une intégration par parties discrète. Cela permet de gérer le cas de $\sum a_n(x)b_n(x)$, lorsque on connaît la CVU de $\sum b_n(x)$. En notant $R_n(x)$ le reste de cette série, on peut réécrire $b_n = R_{n-1} - R_n$, et réindexer de sorte à regrouper les R_n.
4. On fait aussi parfois des transformations d'Abel avec les sommes partielles en écrivant $b_n = B_n - B_{n-1}$. Cela peut être efficace lorsque les sommes partielles (B_n) sont (uniformément) bornées, sans être convergentes.

Exemples 10.3.12

1. Montrer que $\sum \frac{(-1)^n x^n}{n}$ converge uniformément sur $[0,1]$, mais pas normalement.
2. Montrer que $\sum \cos(n)\frac{x^n}{\sqrt{n}}$ est uniformément convergente sur $[0,1]$.
3. Monter que $\sum \frac{\cos(nx)}{n}$ converge uniformément sur $[\varepsilon, 2\pi - \varepsilon]$.

III.2 Régularité et interversions sur des sommes uniformes

On adapte dans cette section les résultats démontrés dans le cadre des suites. Dans la pratique en général, on tentera d'abord de démontrer la convergence normale sur des sous-parties adéquates de A, et sinon, on se ramènera aux études précédentes pour l'étude de la converge uniforme.

Théorème 10.3.13 – Continuité d'une somme infinie

Soit $\sum f_n$ une série uniformément convergente sur A. Si pour tout $n \in \mathbb{N}$, f_n est continue sur A, alors $\sum\limits_{n=0}^{+\infty} f_n$ est continue sur A.

Le théorème de la double limite s'adapte en un théorème d'interversion lim / $\sum$.

Théorème 10.3.14 – Passage à la limite sous $\sum$

Soit $\sum f_n$ une série uniformément convergente sur A et $f = \sum\limits_{n=0}^{+\infty} f_n$. Soit $a \in \overline{A}$. On suppose que pour tout $n \in \mathbb{N}$, $f_n(x) \to \ell_n \in F$. Alors $\sum \ell_n$ converge, et f admet une limite en a, égale à

$$\lim_{x\to a} f(x) = \sum_{n=0}^{+\infty} f_n(x).$$

En d'autres termes,

$$\lim_{x\to a} \sum_{n=0}^{+\infty} f_n(x) = \sum_{n=0}^{+\infty} \lim_{x\to a} f_n(x).$$

Exemples 10.3.15

1. Déterminer $\displaystyle\lim_{x\to+\infty} \frac{1}{x^2+n^2}$.
2. Déterminer $\displaystyle\lim_{x\to+\infty} \zeta(x)$.

Théorème 10.3.16 – Intégration terme à terme sur un segment

Soit (f_n) une suite de fonctions d'un intervalle I à valeurs dans un e.v.n. F de dimension finie, et soit $a \in I$.

1. Si $\sum f_n$ est uniformément convergente sur I, alors

$$x \mapsto \sum_{n=0}^{+\infty} \int_a^x f_n(t)\ \mathrm{d}t$$

   converge uniformément sur I vers $x \mapsto \displaystyle\int_a^x \sum_{n=0}^{+\infty} f_n(t)\ \mathrm{d}t$.
2. En particulier, si $\sum f_n$ est uniformément convergente sur $[a, b]$,

$$\int_a^b \sum_{n=0}^{+\infty} f_n(t)\ \mathrm{d}t = \sum_{n=0}^{+\infty} \int_a^b f_n(t)\ \mathrm{d}t.$$

Exemples 10.3.17

1. Montrer que pour tout $x \in [-1, 1[$,

$$\ln(1+x) = \sum_{k=1}^{+\infty} \frac{(-1)^{k+1}x^k}{k}.$$

2. À l'aide de l'exemple précédent, calculer $\displaystyle\int_0^1 \frac{\ln(1-x)}{x}\ \mathrm{d}x$. On admettra que $\zeta(2) = \dfrac{\pi^2}{6}$.

Théorème 10.3.18 – Dérivation terme à terme

Soit (f_n) une suite de fonctions d'un intervalle I dans F, telles que :

(i) les f_n sont de classe $\mathcal{C}^1$ sur I ;

(ii) la série $\sum f_n$ converge simplement ;

(iii) la série $\sum f_n'$ converge uniformément sur I.

Alors $S = \sum f_n$ est de classe $\mathcal{C}^1$ sur I, et pour tout $x \in I$,

$$S'(x) = \sum_{n=0}^{+\infty} f_n'(x).$$

L'hypothèse (iii) peut être remplacée par :

(iii') la série $\sum f_n'$ converge uniformément sur tout segment inclus dans I.

Ce théorème permet donc de dériver une série terme à terme (sous les hypothèses idoines).

Théorème 10.3.19 – Classe $\mathcal{C}^p$ d'une série

Soit (f_n) une suite de fonctions d'un intervalle I dans F, et $p \in \mathbb{N}$. On suppose que :

(i) les f_n sont de classe $\mathcal{C}^p$ sur I ;

(ii) la série $\sum f_n$ converge simplement ;

(iii) la série $\sum f_n^{(p)}$ converge uniformément sur I.

Alors $S = \sum f_n$ est de classe $\mathcal{C}^p$ sur I, et pour tout $x \in I$, et tout $k \in [\![0, p]\!]$,

$$S^{(k)}(x) = \sum_{n=0}^{+\infty} f_n^{(k)}(x).$$

L'hypothèse (iii) peut être remplacée par :

(iii') la série $\sum f_n^{(p)}$ converge uniformément sur tout segment inclus dans I.

Exemples 10.3.20

1. Montrer que ζ est dérivable à tout ordre sur $\mathbb{R}_+^*$, et exprimer sa dérivée à tout ordre sous forme d'une somme.
2. Montrer que pour tout $x \in]-1, 1[$, et tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{+\infty} (k+n)\dots(k+1)x^k = \frac{n!}{(1-x)^{n+1}}.$$

On retrouve ainsi la formule du binôme négatif, déjà établie par récurrence en utilisant des produits de Cauchy.

III.3 Techniques asymptotiques

Dans ce paragraphe, nous donnons quelques pistes, non exhaustives, pour l'étude asymptotique des sommes de séries de fonctions, au bord de leur domaine.

Méthode 10.3.21 – Déterminer une limite

1. L'outil à privilégier, si possible, est le théorème de la double limite (interversion limite / somme). Attention au fait qu'on ne l'a donné que dans le cas où toutes les limites sont dans F (pas de limite infinie).
2. Pour montrer qu'on a une limite infinie au bord du domaine, procéder par minoration. Dans le cas où les séries sont positives, on pourra commencer par minorer une somme partielle puis passer à la limite.
3. On pourra aussi minorer terme à terme de sorte à se ramener à des séries dont l'étude est plus simple.

Exemple 10.3.22

Déterminer la limite de ζ quand $x \to 1^+$.

La méthode utilisée peut s'exprimer sous forme d'une proposition, mais la démonstration doit être remise en oeuvre à chaque fois.

Proposition 10.3.23 – Limite infinie au bord du domaine, HP

Soit (f_n) une suite de fonctions définies de $[a, b]$ dans $\mathbb{R}$. On suppose que :

(i) pour tout $n \in \mathbb{N}$, $f_n \geqslant 0$,

(ii) $\sum f_n$ converge (simplement) sur $]a, b]$,

(iii) $\sum f_n(a)$ diverge.

Alors $\displaystyle\lim_{x \to a^+} \sum_{n=0}^{+\infty} f_n(x) = +\infty$.

Cela s'adapte bien sûr aussi pour la borne supérieure de l'intervalle.

Méthode 10.3.24 – Trouver un équivalent

1. On peut deviner un équivalent. Par exemple en remplaçant le terme général par un terme qui lui est proche asymptotiquement. Dans ce cas,
   - on peut diviser par l'équivalent suspecté (ou son ordre de grandeur sans la constante) et essayer de se ramener au théorème de la double limite ;
   - Si les ordres de grandeurs des différents termes de la série sont très différents au voisinage du point considéré, (si on a une décroissance rapide de ces ordres de grandeur), il se peut que la somme partielle fournisse un développement asymptotique. C'est ce qui se passe par exemple avec les séries entières $\sum a_n x^n$ en 0, comme on le verra. Cela permet de deviner les équivalents successifs, et donc de se ramener au premier point ci-dessus.
2. Dans le cas où pour tout x, $f_n(x) = g(n, x)$ où g est définie sur $[a, +\infty[\times I$, et est monotone par rapport à sa première variable (la deuxième étant fixée), on pourra faire une comparaison avec une intégrale.
3. Dans certains cas, on peut essayer, à l'aide de développements asymptotiques, de se ramener à des séries classiques qu'on sait expliciter.

Exemples 10.3.25

1. Équivalent en $+\infty$ de $\zeta(x) - 1$. Développement asymptotique à 3 termes de ζ en $+\infty$
2. Équivalent en 0 de $\zeta(x)$.
3. Équivalent en $+\infty$ de $\displaystyle\sum_{k=0}^{+\infty} \frac{1}{x^2 + t^2}$.
4. Équivalent en $+\infty$ de $\displaystyle\sum_{n=0}^{+\infty} \frac{(a_n x)^n}{n!}$, où $a_n = (2^n + 1)^{\frac{1}{n+1}}$. On pourra faire un développement asymptotique de a_n.

IV Approximations uniformes

IV.1 Principes généraux

Définition 10.4.1 – Approximation uniforme à ε près

Soit $f : A \to F$ une application d'une partie A de E dans l'e.v.n. de dimension finie F. On dit que $g : A \to F$ est une approximation uniforme de f à ε près si $\|g - f\|_\infty \leqslant \varepsilon$.
On dira que g est une ε-approximation uniforme de f.

Remarque 10.4.2

Cela implique implicitement que $g - f$ doit être borné ; mais g et f ne le sont pas nécessairement.

Exemple 10.4.3

La fonction $x \mapsto x^2$ est une approximation uniforme à $\varepsilon = \frac{1}{100}$ près de $x \mapsto x^2 + \frac{1}{100}\sin(x)$ sur $\mathbb{R}$.

On cherche à savoir si une fonction f donnée peut être approchée uniformément arbitrairement proche par des fonctions d'un certain type.

Définition 10.4.4 – Approximation uniforme par un ensemble de fonctions

Soit $\mathcal{G}$ une partie de E^A. On dit que $f : A \to E$ peut être approchée uniformément par des fonctions de $\mathcal{G}$ si et seulement si pour tout $\varepsilon > 0$, il existe $g \in \mathcal{G}$ telle que g soit une ε-approximation uniforme de f.

Exemple 10.4.5

On dit que $f : \mathbb{R} \to \mathbb{R}$ est en escalier si elle est en escalier sur tout $[a, b]$. On remarquera par exemple que $x \mapsto \lfloor x \rfloor$ est alors en escalier sur $\mathbb{R}$ (mais non bornée). En effet, contrairement au cas d'une fonction en escalier sur un segment, on peut avoir une infinité de paliers.
Soit $L > 0$. Montrer que toute fonction L-lipschitzienne $f : \mathbb{R} \to \mathbb{R}$ peut être approchée uniformément par des fonctions en escalier.

Remarque 10.4.6

1. Si $\mathcal{G} \subset \mathcal{B}(A, \mathbb{R})$, alors les propositions suivantes sont équivalentes:
   (i) f peut être approchée uniformément par des fonctions de $\mathcal{G}$
   (ii) $f \in \overline{\mathcal{G}}$ (pour la norme $\|\cdot\|_\infty$).

   Dans ce cas, f est nécessairement aussi bornée.
2. Ainsi, si $\mathcal{F}, \mathcal{G} \subset \mathcal{B}$, les propositions suivantes sont équivalentes:
   (i) tout f de $\mathcal{F}$ peut être approché uniformément par des fonctions de $\mathcal{G}$;
   (ii) $\mathcal{G}$ est dense dans $\mathcal{F}$ (pour la norme $\|\cdot\|_\infty$).

Proposition 10.4.7 – Caractérisation séquentielle

Une application f peut être approchée uniformément par des applications de $\mathcal{G}$ ssi il existe une suite $(g_n) \in \mathcal{G}^{\mathbb{N}}$ telle que (g_n) converge uniformément vers f.

IV.2 Approximations uniformes classiques

On généralise la définition des fonctions en escalier, vues dans le cadre de fonctions à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. On rappelle que F est un e.v.n. de dimension finie.

Définition 10.4.8 – Fonctions en escalier

La fonction $f : [a, b] \to F$ est une fonction en escalier s'il existe une subdivision

$$\alpha = \sigma_0 < \cdots < \sigma_n = b$$

de $[a, b]$ telle que f est constante sur chaque intervalle $]\sigma_i, \sigma_{i+1}[$ de la subdivision.
On note $\mathrm{Esc}([a, b], F)$ l'ensemble des fonctions en escalier sur $[a, b]$, à valeurs dans F.

Un premier théorème d'approximation uniforme a déjà été vu dans le cas réel au moment de définir l'intégrale de Riemann.

Théorème 10.4.9 – Approximation uniforme de $\mathcal{C}^0$ par Esc

Les fonctions continues sur un segment $[a, b]$ à valeurs dans un e.v.n. F de dimension fini peuvent être approchées uniformément par des fonctions en escalier sur $[a, b]$.

Autrement dit, $\mathrm{Esc}([a, b], F)$ est dense dans $\mathcal{C}^0([a, b], F)$.

◁ **Éléments de preuve.**
Cela provient de la continuité uniforme de f, issue du théorème de Heine. On peut alors considérer une subdivision régulière suffisamment fine. ▷

On note $\mathcal{C}^0_M(I, \mathbb{C})$ l'ensemble des applications continues par morceaux de I dans $\mathbb{R}$.

Corollaire 10.4.10 – Approximation uniforme de $\mathcal{C}^0_M$ par Esc

Les fonctions continues par morceaux sur $[a, b]$ peuvent être approchées uniformément par des fonctions en escalier sur $[a, b]$.

Autrement dit, $\mathrm{Esc}([a, b], F)$ est dense dans $\mathcal{C}^0_M([a, b], F)$.

◁ **Éléments de preuve.**
Approcher uniformément f sur chaque part d'une subdivision associée (on peut à chaque fois prolonger $f_{]\sigma_i,\sigma_{i+1}[}$ par continuité pour se ramener à une fonction continue sur un segment). ▷

Un autre théorème est nettement plus délicat à obtenir. Sa démonstration est de fait hors-programme. On rappelle que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Pour S une partie de $\mathbb{R}$, on note $\mathbb{K}_S[x]$ l'ensemble des fonctions polynomiales à coefficients dans $\mathbb{K}$, restreintes à S.

Théorème 10.4.11 – théorème d'approximation uniforme de Weierstrass

Toute fonction f continue sur un segment $[a, b]$ peut être approchée uniformément par des fonctions polynomiales à coefficients dans $\mathbb{K}$, définies sur ce segment.

Autrement dit, en notant $S = [a, b]$, $\mathbb{K}_S[x]$ est dense dans $\mathcal{C}^0(S, \mathbb{K})$.

◁ **Éléments de preuve.**
Hors-programme. On le verra peut-être en DM. ▷

Corollaire 10.4.12 – Approximation par des fonctions $\mathcal{C}^\infty$

Toute fonction $f : [a, b] \to \mathbb{K}$ peut être approchée uniformément par des fonctions de classe $\mathcal{C}^\infty$.

◁ **Éléments de preuve.**
Les fonctions polynomiales sont de classe $\mathcal{C}^\infty$. ▷

11

Intégrales à paramètres

Le but de ce chapitre est d'étudier le comportement d'intégrales à paramètres.
Il s'agit donc d'étudier une famille de fonctions $(f_\lambda)_{\lambda\in\Lambda}$ intégrables sur un intervalle I, à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Leurs intégrales définissent une famille $\left(\int_I f_\lambda(t)\ \mathrm{d}t\right)_{\lambda\in\Lambda}$, donc on cherche à étudier le comportement par rapport au paramètre λ.
Le contexte initial (qui est à la source des autres études) concerne un paramètre entier (*i.e.* $\Lambda = \mathbb{N}$). On dispose donc d'une suite de fonctions $(f_n)_{n\in\mathbb{N}}$ intégrables sur $\mathbb{N}$, et on s'intéresse à la limite de la suite $\left(\int_I f_n(t)\ \mathrm{d}t\right)_{n\in\mathbb{N}}$.
Nous nous intéresserons ensuite au cas où le paramètre λ est pris dans une partie Λ d'un e.v.n. de dimension finie. Nous adopterons dans ce contexte plutôt une représentation fonctionnelle de la famille, qui peut alors être décrite par une application :

$$\begin{array}{rcl} f : \Lambda\times I & \to & \mathbb{K} \\ x & \mapsto & f(x,t). \end{array}$$

Leurs intégrales définissent donc une application de Λ dans $\mathbb{K}$

$$\begin{array}{rcl} g : \Lambda & \to & \mathbb{K} \\ x & \mapsto & \displaystyle\int_I f(x,t)\ \mathrm{d}t \end{array}$$

On s'intéressera dans ce contexte au propriétés relatifs aux limites et à la continuité.
Enfin, dans un contexte un peu plus restreint où Λ est un intervalle de $\mathbb{R}$, on s'intéressera aux problèmes de dérivabilité de g.

I Suites et séries d'intégrales

Dans toute cette section, I désigne un intervalle de $\mathbb{R}$, E désigne un e.v.n. de dimension finie, et (f_n) désigne une suite d'applications de I dans E. Les premiers résultats ont déjà été vus dans des contextes divers, et le but est ici de les regrouper afin d'avoir une meilleure vue d'ensemble sur les différents outils à disposition pour l'étude des intégrales de suites ou séries de fonctions.

I.1 Interversion limite/intégrale

Lorsque I est un segment, on dispose d'un outil particulier, déjà vu dans un chapitre antérieur, lié à la convergence uniforme.

Théorème 11.1.1 – interversion lim/intégrale en cas de CVU

On suppose que :

(i) I est un segment ;

(ii) les applications (f_n) sont continues ;

(iii) la suite (f_n) converge uniformément vers f sur I.

Alors,

$$\lim_{n\to+\infty}\int_I f_n(t)\ \mathrm{d}t = \int_I \lim_{n\to+\infty} f_n(t)\ \mathrm{d}t.$$

Nous avions montré un résultat plus général, concernant la convergence uniforme d'une suite de primitives.

Exemple 11.1.2

Calculer $\lim\limits_{n\to+\infty}\int_0^1 \tan(x)^{\ln n}\ \mathrm{d}x$.

Proposition 11.1.3 – Extension du théorème d'interversion limite/intégrale

Le théorème d'interversion limite/intégrale par CVU reste vrai sous les hypothèses suivantes :

(i) I est un segment ;

(ii) les (f_n) sont c.p.m. (ou plus généralement Riemann-intégrable)

(iii) (f_n) converge uniformément vers une fonction f

(iv) la fonction f est c.p.m. (ou plus généralement Riemann-intégrable)

La quatrième hypothèse remplace les hypothèses plus fortes de l'énoncé principal, qui assuraient la continuité de f.

Cela permet d'englober le cas des fonctions en escaliers.

Méthode 11.1.4

Si certains points empêchent la CVU, ou si l'intégrale n'est pas définie sur tout le segment on peut soit se rabattre sur le théorème de convergence dominé, soit, si ce n'est pas possible, couper autour de la borne.

Exemple 11.1.5

Calculer $\lim\limits_{n\to+\infty}\int_0^{\frac{pi}{4}} \tan(x)^n\ \mathrm{d}x$

L'outil le plus efficace, et plus général que le précédent (puisque valide sur des intervalles quelconques) reste le suivant, qu'on a déjà donné, sans le démontrer/

Théorème 11.1.6 – Théorème de convergence dominée, TCD

Soit (f_n) une suite de fonctions de I dans $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$. On suppose que :

(i) pour tout $n\in\mathbb{N}$, f_n est c.p.m.

(ii) (f_n) converge simplement vers une fonction f

(iii) (hyp. de domination) il existe $\varphi : I\to\mathbb{K}$ intégrable telle que : $\forall n\in\mathbb{N}, \forall x\in I,\ |f_n(x)|\leqslant\varphi(x)$.

Alors les (f_n) et f sont intégrables sur I et $\displaystyle\int_I f_n \underset{n\to+\infty}{\longrightarrow}\int_I f$.

Exemples 11.1.7

1. Reprendre les exemples précédents, y compris celui où on n'avait pas CVU sur tout l'intervalle.
2. Déterminer $\lim_{n\to+\infty} \int_0^n \left(1-\frac{x}{n}\right)^n \, \mathrm{d}x$.
3. Déterminer $\lim_{n\to+\infty} \int_0^1 (x+n)\mathrm{e}^{-\sqrt{n}x} \, \mathrm{d}x$

Remarques 11.1.8

1. Lorsque la suite (f_n) est décroissante (ce qui est le cas des premiers exemples), alors on pourra toujours dominer la suite par f_0, à condition qu'elle soit intégrable.
2. Lorsque la suite est croissante, et que la limite f est intégrable, alors on pourra dominer (f_n) par f. C'est le cas du deuxième exemple.
3. Le dernier exemple montre un cas où (f_n) n'est ni croissante ni décroissante.

I.2 Intégration terme à terme

Nous avons vu que le théorème d'interversion limite/intégrale dans de cas CVU s'adapte bien au cas des séries :

Théorème 11.1.9 – interversion $\sum$ / $\int$ par CVU sur un segment

Soit (f_n) une suite de fonctions d'un intervalle I à valeurs dans un e.v.n. F de dimension finie, et soit $a \in I$.

1. Si $\sum f_n$ est uniformément convergente sur I, alors
$$x \mapsto \sum_{n=0}^{+\infty} \int_a^x f_n(t) \, \mathrm{d}t$$
converge uniformément sur I vers $x \mapsto \int_a^x \sum_{n=0}^{+\infty} f_n(t) \, \mathrm{d}t$.
2. En particulier, si $\sum f_n$ est uniformément convergente sur $[a,b]$,
$$\int_a^b \sum_{n=0}^{+\infty} f_n(t) \, \mathrm{d}t = \sum_{n=0}^{+\infty} \int_a^b f_n(t) \, \mathrm{d}t.$$

L'intégrabilité sur un segment quelconque I n'ayant été définie que pour des fonctions à valeurs dans $\mathbb{K}$, les autres énoncés se placent dans ce contexte ; ils peuvent s'adapter sans réelle difficulté au cadre d'e.v.n. de dimension finie, en regardant coordonnée par coordonnée dans une base.

Théorème 11.1.10 – Intégration terme à terme, cas positif

Soit, pour tout $n \in \mathbb{N}$, $f_n : I \to \mathbb{R}$. On suppose que :

(i) pour tout $n \in \mathbb{N}$, $f_n \geqslant 0$;

(ii) $\sum_{n\geqslant 0} f_n$ converge simplement vers une fonction S

(iii) S est continue par morceaux sur I

Alors, l'égalité suivante est toujours vérifiée dans $\overline{\mathbb{R}}_+$:
$$\int_I \sum_{n=0}^{+\infty} f_n(t) \, \mathrm{d}t = \sum_{n=0}^{+\infty} \int_I f_n(t) \, \mathrm{d}t.$$

◁ **Éléments de preuve.**

Démonstration hors programme. Elle reste abordable avec le TCD.

1. Si $\sum f_n$ est intégrable, dominer la somme partielle par la somme totale et utiliser le TCD
2. Sinon, on a divergence sur une des deux bornes de I, disons la borne supérieure. Soi $a \in I$. Justifier qu'il existe x_0 tel que pour tout $x \geqslant x_0$,

$$\int_a^x \sum_{n=0}^{+\infty} f_n(x)\ \mathrm{d}x \geqslant A,$$

A étant posé arbitrairement. Utiliser encore le TCD sur $[a, x]$, pour montrer que $\sum_{n=0}^{+\infty} \int_I f_n$ est minoré par tout A réel.

▷

Remarque 11.1.11

En particulier, si les f_n sont positifs, on peut justifier l'intégrabilité de $\sum_{n=0} f_n$ en majorant $\sum_{n=0}^{+\infty} \int_I f_n(t)\ \mathrm{d}t$.

Exemple 11.1.12

1. Intégrabilité de $x \mapsto \sum_{n=1}^{+\infty} \frac{1}{x^3 + n^3}$ sur $[0, +\infty[$?
2. Intégrabilité de $x \mapsto \sum_{n=1}^{+\infty} \frac{1}{1 + n^2 t^2}$ sur $[0, +\infty[$.
3. Calculer $\int_0^1 \sum_{n=1}^{+\infty} n^2 \mathbb{1}_{[0, \frac{1}{n^4}]}$
4. Calculer $\int_0^{+\infty} \frac{t}{\mathrm{e}^t - 1}\ \mathrm{d}t$, en fonction de $\zeta(2)$.

Remarques 11.1.13

1. Le TITT positif peut s'utiliser dans des cas où on n'a pas convergence uniforme, et permet donc de gérer des cas plus variés que le théorème d'interversion $\sum / \int$ par CVU. C'est par exemple le cas du troisième exemple.
2. Le dernier exemple montre une situation typique dans laquelle on combine un développement en série et une intégration terme à terme pour calculer une intégrale. Les exemples de ce type sont nombreux, pensez-y !
3. Le point le plus délicat à justifier dans le cas positif est souvent la c.p.m. de la somme. Le programme stipule explicitement de ne pas trop s'attarder sur ces justifications.

Théorème 11.1.14 – Intégration terme à terme, cas général

Soit, pour tout $n \in \mathbb{N}$, $f_n : I \to \mathbb{K}$. On suppose que :

(i) pour tout $n \in \mathbb{N}$, f_n est intégrable

(ii) $\sum_{n \geqslant 0} f_n$ converge simplement vers une fonction S ;

(iii) S est c.p.m. sur I ;

(iv) $\sum_{n=0}^{+\infty} \int_I |f_n(t)|\ \mathrm{d}t < +\infty$.

Alors S est intégrable sur I, et

$$\int_I \sum_{n=0}^{+\infty} f_n(t)\ \mathrm{d}t = \sum_{n=0}^{+\infty} \int_I f_n(t)\ \mathrm{d}t.$$

◁ **Éléments de preuve.**

Démonstration hors-programme. On peut le démontrer un peu plus facilement sous une hypothèse supplémentaire, en supposant de plus $\sum |f_n|$ c.p.m.. ▷

Exemple 11.1.15

Calculer $\displaystyle\int_0^1 \frac{\ln(1+t)}{t}\ \mathrm{d}t$ sous forme d'une somme de Riemann alternée.

Remarque 11.1.16

Remarquez la similitude entre ces résultats d'intégration terme à terme, et les théorèmes de Fubini pour les familles sommables (version positive dans $\overline{\mathbb{R}}$, et version générale). D'ailleurs, les théorèmes de Fubini peuvent se retrouver à partir des théorèmes d'ITT, en considérant des fonctions en escalier f_n prenant la valeur $a_{n,p}$ sur l'intervalle $[p, p+1[$. Ces théorèmes peuvent être vus comme une version continue/discrète des théorèmes de Fubini discret/discret. Il en existe aussi une version pour les intégrales doubles (version continue/continue), mais celle-ci n'est pas au programme.

Méthode 11.1.17 – Si $\sum \int_I |f_n| = +\infty$

Dans le cas où les conditions du TITT ne sont pas vérifiées, on peut parfois utiliser le TCD sur la somme partielle (S_n). C'est notamment assez fréquent avec des séries alternées, puisqu'on parvient dans ce cas à encadrer la somme S entre S_0 et S_1. L'hypothèse de domination est donc assez facile à obtenir.

Exemple 11.1.18

Justifier l'intégrabilité, et exprimer sous forme d'une somme la valeur de $\displaystyle\int_0^{+\infty} \sum_{n=1}^{+\infty} (-1)^n \mathrm{e}^{\sqrt{n}\cdot t}\ \mathrm{d}t$

Remarque 11.1.19

Le caractère alterné permet aussi l'étude facile de la CVU, donc de la continuité sur les compacts, ce qui nous assure la continuité par morceaux (au moins sur l'intervalle $\mathring{I}$, ce qui est suffisant).

Méthode 11.1.20 – Récupérer les bornes par CVU

Si on ne peut faire qu'une interversion partielle (par CVU ou par TITT), on peut faire tendre une borne vers la borne voulue en étudiant la convergence uniforme de $\displaystyle\sum_{n=0}^{+\infty} \int_a^x f_n(t)\ \mathrm{d}t$, pour pouvoir utiliser le théorème de la double limite.

Exemple 11.1.21

En considérant $x \mapsto \sum_{n=0}^{+\infty} (-x)^n$, montrer que $\sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} = \ln(2)$.

II Limite et continuité d'intégrales à paramètre dans un e.v.n.

Le point de départ de l'étude des intégrales à paramètres est la version continue du TCD. Elle est donnée dans le programme sous la forme suivante.

Théorème 11.2.1 – Théorème de convergence dominée, version continue réelle

Soit I et J deux intervalles de $\mathbb{R}$ et $(f_\lambda)_{\lambda\in J}$ une famille de fonctions de I dans $\mathbb{K}$. Soit $f \in \mathbb{K}^I$ et $a \in \overline{J}$. On suppose que :

(i) pour tout $\lambda \in J$, f_λ est c.p.m. ;

(ii) pour tout $t \in I$, $f_\lambda(t) \underset{\lambda\to a}{\longrightarrow} f(t)$, et f est c.p.m. ;

(iii) il existe φ telle que pour tout $\lambda \in J$, $|f_\lambda| \leqslant \varphi$, et φ est intégrable.

Alors
$$\lim_{\lambda\to a}\int_I f_\lambda(t)\,\mathrm{d}t = \int_I f(t)\,\mathrm{d}t = \int_I \lim_{\lambda\to a} f_\lambda(t)\,\mathrm{d}t.$$

◁ **Éléments de preuve.**

On montre directement la version plus générale ci-dessous. ▷

Une version un peu plus générale, pas explicitement au programme (mais pouvant se déduire du théorème de continuité qui suit, après prolongement de chaque $f(x,\cdot)$), et exprimée, comme on le fait souvent, en notation fonctionnelle plutôt que familiale, permet de considérer un paramètre dans un e.v.n..
Dans cet énoncé, f étant une fonction définie sur un produit cartésien $A \times I$, $f(\cdot,t)$ désigne la fonction $x \mapsto f(x,t)$ définie sur A, obtenue en fixant la deuxième variable égale à t. De même $f(x,\cdot)$ désigne la fonction d'une variable obtenue en faixant la première variable égale à x.

Théorème 11.2.2 – Théorème de convergence dominée, version continue vectorielle

Soit I un intervalle de $\mathbb{R}$, E un e.v.n. de dimension finie, et A une partie de E. Soit $f : A\times I \to \mathbb{K}$, $a \in \overline{A}$, et $g : I \to \mathbb{K}$. On suppose que :

(i) pour tout $x \in A$, $f(x,\cdot)$ est c.p.m. ;

(ii) pour tout $t \in I$, $f(x,t) \underset{x\to a}{\longrightarrow} g(t)$, et g est c.p.m. ;

(iii) (domination) il existe φ telle que pour tout $x \in A$, $|f(x,\cdot)| \leqslant \varphi$, et φ est intégrable.

Alors
$$\lim_{x\to a}\int_I f(x,t)\,\mathrm{d}t = \int_I g(t)\,\mathrm{d}t = \int_I \lim_{x\to a} f(x,t)\,\mathrm{d}t.$$

◁ **Éléments de preuve.**

Par critère séquentiel, on se ramène au TCD discret. ▷

Exemple 11.2.3

Déterminer $\displaystyle\lim_{x\to+\infty}\int_0^{+\infty}\frac{x+\ln(t)}{1+x^2t^{\frac{3}{2}}}\,\mathrm{d}t$.

En considérant $a \in I$, on obtient une propriété de continuité

Théorème 11.2.4 – Continuité ponctuelle d'une intégrale à paramètre dans un e.v.n.

Soit I un intervalle de $\mathbb{R}$, E un e.v.n. de dimension finie, et A une partie de E. Soit $f : A\times I \to \mathbb{K}$ et $a \in A$. On suppose que :

(i) pour tout $x \in A$, $f(x,\cdot)$ est c.p.m. ;

(ii) pour tout $t \in I$, $f(\cdot,t)$ est continue en a ;

(iii) (domination) il existe φ telle que pour tout $x \in A$, $|f(x,\cdot)| \leqslant \varphi$, et φ est intégrable.

Alors la fonction

$$F : x \mapsto \int_I f(x,t)\ \mathrm{d}t$$

est continue en a.

◁ **Éléments de preuve.**
C'est l'application directe du TCD, dans sa version continue vectorielle, en un point du domaine. ▷

On en déduit de façon immédiate :

Théorème 11.2.5 – Continuité globale d'une intégrale à paramètre dans un e.v.n.

Soit I un intervalle de $\mathbb{R}$, E un e.v.n. de dimension finie, et A une partie de E. Soit $f : A \times I \to \mathbb{K}$. On suppose que :

(i) pour tout $x \in a$, $f(x,\cdot)$ est c.p.m. ;

(ii) pour tout $t \in I$, $f(\cdot,t)$ est continue en a ;

(iii) (domination) il existe φ telle que pour tout $x \in A$, $|f(x,\cdot)| \leqslant \varphi$, et φ est intégrable.

Alors la fonction

$$F : x \mapsto \int_I f(x,t)\ \mathrm{d}t$$

est continue sur A.

Exemple 11.2.6

1. Étudier la continuité de $x \mapsto \displaystyle\int_0^{+\infty} \sin(xt)\mathrm{e}^{-t^2}\ \mathrm{d}t$ et déterminer sa limite en $+\infty$.
2. Transformée de Laplace :
   Soit f une fonction intégrable sur $]0,+\infty[$. Soit $A = \{p \in \mathbb{C} \mid \mathrm{Re}(p) \geqslant 0\}$.
   Étudier l'existence et la continuité sur A de la transformée de Laplace $F(p) = \displaystyle\int_0^{+\infty} f(t)\mathrm{e}^{pt}\ \mathrm{d}t$.

Remarque 11.2.7

En pratique, il est fréquent de ne pas vérifier l'hypothèse de domination sur A tout entier, mais de le faire sur une famille $(B_i)_{i\in I}$ de sous-ensembles de A tel que tout point de A soit dans l'intérieur d'un des B_i. Par exemple, si A est un intervalle de $\mathbb{R}$, on peut se contenter de faire la domination sur tous les segments.

Exemple 11.2.8

Étudier la continuité sur $\mathbb{R}_+^*$ de $\Gamma : x \mapsto \displaystyle\int_0^{+\infty} t^{x-1}\mathrm{e}^{-t}$.

III Dérivation d'intégrales à paramètre réel

Dans cette section, afin de pouvoir dériver, on considère le cas où le paramètre est réel, de sorte à obtenir après intégration une application d'une variable réelle.

Théorème 11.3.1 – Dérivation d'une intégrale à paramètre réel

Soit I et A deux intervalles de $\mathbb{R}$ et $f : A \times I \to \mathbb{K}$ une application telle que :

(i) pour tout $t \in I$, $f(\cdot,t)$ est de classe $\mathcal{C}^1$ sur I ;

(ii) pour tout $x \in A$, $f(x,\cdot)$ est intégrable sur I ;

(iii) pour tout $x \in A$, $\dfrac{\partial f}{\partial x}(x,\cdot)$ est c.p.m. sur I ;

(iv) (domination de la dérivée) il existe φ intégrable sur I telle que

$$\forall x \in A, \quad \left|\frac{\partial f}{\partial x}(x, \cdot)\right| \leqslant \varphi.$$

Alors la fonction $g : x \mapsto \int_I f(x,t)\ \mathrm{d}t$ est de classe $\mathcal{C}^1$ sur A, et :

$$\forall x \in A, \quad g'(x) = \int_I \frac{\partial f}{\partial x}(x,t)\ \mathrm{d}t.$$

◁ **Éléments de preuve.**

Former le taux d'accroissement. L'IAF et les hypothèses permettent l'utilisation du TCD continu. ▷

Exemple 11.3.2

Soit $F : x \mapsto \left(\int_0^x \mathrm{e}^{-t^2}\ \mathrm{d}t\right)^2$. Exprimer la dérivée de F sous forme de la dérivée d'une intégrale à paramètres, et en déduire la valeur de $\int_0^{+\infty} \mathrm{e}^{-t^2}\ \mathrm{d}t$.

Par symétrisation par parité, on vient de démontrer le résultat classique suivant :

Proposition 11.3.3 – Intégrale de Gauss, HP classique

$\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2}\ \mathrm{d}t = \sqrt{\pi}$.

Par récurrence à partir du théorème de dérivation, on obtient :

Théorème 11.3.4 – Classe $\mathcal{C}^k$ d'une intégrale à paramètre réel

Soit I et A deux intervalles de $\mathbb{R}$ et $f : A \times I \to \mathbb{K}$ une application telle que :

(i) pour tout $t \in I$, $f(\cdot, t)$ est de classe $\mathcal{C}^k$ sur I ;

(ii) pour tout $x \in A$ et tout $j \in [\![0, k-1]\!]$, $\frac{\partial^j f}{\partial x^j}(x, \cdot)$ est intégrable sur I ;

(iii) pour tout $x \in A$, $\frac{\partial^k f}{\partial x^k}(x, \cdot)$ est c.p.m. sur I ;

(iv) (domination de la dérivée k-ième) il existe φ intégrable sur I telle que

$$\forall x \in A, \quad \left|\frac{\partial^k f}{\partial x^k}(x, \cdot)\right| \leqslant \varphi.$$

Alors la fonction $g : x \mapsto \int_I f(x,t)\ \mathrm{d}t$ est de classe $\mathcal{C}^k$ sur A, et

$$\forall j \in [\![0, k]\!], \forall x \in A, \quad g^{(k)}(x) = \int_I \frac{\partial^j f}{\partial x^j}(x,t)\ \mathrm{d}t.$$

Exemple 11.3.5

Montrer que la fonction $\Gamma : x \mapsto \int_0^{+\infty} t^{x-1}\mathrm{e}^{-t}\ \mathrm{d}t$ est de classe $\mathcal{C}^\infty$ et exprimer sa dérivée sous forme intégrale.

IV La fonction Γ (HP mais classique)

Nous faisons dans ce paragraphe un rapide bilan des propriétés importantes de la fonction Γ d'Euler, pour la plupart déjà vues en exemple ou en exercice.

Définition 11.4.1 – Fonction Γ, Euler

La fonction Γ est définie, pour tout $z \in \mathbb{C}$ telle que l'intégrale converge, par :

$$\Gamma(z) = \int_0^{+\infty} t^{z-1} \mathrm{e}^{-t} \, \mathrm{d}t.$$

Proposition 11.4.2 – Domaine de définition

La fonction Γ est définie sur $P = \{z \in \mathbb{C} \mid \mathrm{Re}(z) > 0\}$.

Proposition 11.4.3 – Identité remarquable

Pour tout $z \in P$, $\Gamma(z+1) = z\Gamma(z)$.

Proposition 11.4.4 – Valeurs remarquables

1. Pour tout $n \in \mathbb{N}^*$, $\Gamma(n) = (n-1)!$.
2. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

◁ **Éléments de preuve.**

Pour le 2, se ramener à l'intégrale de Gauss par changement de variable. ▷

Proposition 11.4.5 – Continuité sur P

La fonction Γ est continue sur P.

Proposition 11.4.6 – Régularité sur $\mathbb{R}_+^*$

La fonction Γ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}_+^*$, et pour tout $n \in \mathbb{N}$, et tout $x \in \mathbb{R}_+^*$,

$$\Gamma^{(n)}(x) = \int_0^{+\infty} (\ln(t))^n t^{x-1} \mathrm{e}^{-t} \, \mathrm{d}t.$$

Proposition 11.4.7 – Convexité et variations

Γ est convexe, décroissante sur $]0,1]$ puis croissante sur $[1,+\infty[$.

◁ **Éléments de preuve.**

C'est immédiat pour les dérivées d'ordre pair. Constater ensuite par un argument direct que les dérivées d'ordre pair sont croissantes, donc admettent une dérivée positive. ▷

En particulier, cela implique la croissance et la convexité de Γ et de toutes ses dérivées.

Proposition 11.4.8 – Comportement asymptotique

(i) $\lim\limits_{x \to 0} \Gamma(x) = +\infty$

(ii) $\lim\limits_{x \to +\infty} \Gamma(x) = +\infty$

(iii) $\Gamma(x) \underset{0}{\sim} \dfrac{1}{x}$.

◁ **Éléments de preuve.**

La dernière propriété entraînant la première, on peut la montrer d'abord, avec l'identité remarquable. La deuxième s'obtient par croissance et les valeurs particulières déjà trouvées. ▷

La fonction Γ possède bien d'autres propriétés, et est d'une grande importance, notamment car elle intervient dans le calcul de nombreuses transformées de Laplace. Nous aurons peut-être l'occasion de prolonger notre étude de Γ en DM.

12

Endomorphismes d'un espace euclidien

Le but de ce chapitre est d'étudier certains types d'endomorphismes d'un espace euclidien. On s'intéresse en particulier aux isométries vectorielles (les endomorphismes conservant les distances) et aux endomorphismes autoadjoint, dont la matrice dans une base orthonormale est symétrique.
On montrera notamment que, dans une base orthonormale bien choisie, les isométries f ont une matrice diagonale par blocs, chaque bloc étant de taille 1 ou 2, et correspondant à une direction stable sur laquelle f agit comme une symétrie ou l'identité (pour les blocs de taille 1), ou a un plan stable sur lequel f agit comme une rotation.
On montrera également que les endomorphismes autoadjoints (ou symétriques) sont diagonalisables en base orthonormale, ce qui se réexprime en disant que toute matrice symétrique réelle est diagonalisable. C'est le théorème spectral. Enfin, nous caractériserons, pour un endomorphisme autoadjoint a, la positivité de la forme quadratique $x \mapsto \langle x, a(x)\rangle$ (ou matriciellement $X \mapsto X^\top AX$) par la positivité du spectre. Cette propriété sera utile en particulier pour l'étude des points critiques des fonctions de plusieurs variables, en vue de faire de l'optimisation (recherche d'extrema).

I Rappels sur les espaces euclidiens

Ce survol rapide des définitions et propriétés essentielles du cours de première année ne remplace pas une révision complète et en profondeur de votre cours de première année. Il n'a pas vocation à être complet, et ne fait que réintroduire rapidement les notions que nous aurons à utiliser par la suite.

I.1 Formes bilinéaires et produit scalaire

Définition 12.1.1 – Forme bilinéaire

Soit E un espace vectoriel sur $\mathbb{R}$. Une forme bilinéaire φ sur E est une application $\varphi : E \times E \to \mathbb{R}$, linéaire par rapport à chaque facteur, l'autre étant fixé, c'est-à-dire :

(i) $\forall (x,y,z) \in E^3, \quad \forall \lambda \in \mathbb{R}, \quad \varphi(\lambda x + y, z) = \lambda\varphi(x,z) + \varphi(y,z)$

(ii) $\forall (x,y,z) \in E^3, \quad \forall \lambda \in \mathbb{R}, \quad \varphi(x, \lambda y + z) = \lambda\varphi(x,y) + \varphi(x,z)$.

On rappelle la propriété de bilinéarité généralisée

Lemme 12.1.2 – Bilinéarité généralisée

Soit φ une forme bilinéaire sur E. Soit $(k,\ell) \in (\mathbb{N}^*)^2$, et soit $(x_1,\dots,x_k,y_1,\dots,y_\ell) \in E^{k+\ell}$, et

$(\lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_\ell) \in \mathbb{K}^{k+\ell}$. Alors

$$\varphi\left(\sum_{i=1}^{k} \lambda_i x_i, \sum_{j=1}^{\ell} \mu_j y_j\right) = \sum_{i=1}^{k}\sum_{j=1}^{\ell} \lambda_i \mu_j \varphi(x_i, y_j).$$

Définition 12.1.3 – Matrice associée à une forme bilinéaire

Soit φ une forme bilinéaire sur E de dimension finie, et $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Alors on définit la matrice de φ relativement à la base $\mathcal{B}$ par :

$$\mathrm{Mat}_{\mathcal{B}}(\varphi) = (\varphi(e_i, e_j))_{1 \leqslant i,j \leqslant n}.$$

Théorème 12.1.4 – Caractérisation de $\mathrm{Mat}_{\mathcal{B}}(\varphi)$ par la relation $\varphi(x, y) = X^\top MY$

Soit $M \in \mathcal{M}_n(\mathbb{K})$. La matrice M est la matrice de φ relativement à la base $\mathcal{B}$ si et seulement si pour tout $(x, y) \in E^2$, dont les matrices colonnes coordonnées dans $\mathcal{B}$ sont X et Y,

$$\varphi(x, y) = X^\top MY.$$

Proposition 12.1.5 – Formule de changement de base

Soit E un espace vectoriel de dimension finien et φ une forme bilinéaire. Soit $\mathcal{B}$ et $\mathcal{C}$ deux bases de E, et $P = P_{\mathcal{B}}^{\mathcal{C}}$ la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$. Alors

$$\mathrm{Mat}_{\mathcal{C}}(\varphi) = P^\top \mathrm{Mat}_{\mathcal{B}}(\varphi) P.$$

Définition 12.1.6 – Produit scalaire

Un produit scalaire φ sur E est une forme bilinéaire telle que :

(i) (positivité) : pour tout $x \in E$, $\varphi(x, x) \geqslant 0$

(ii) (caractère défini) : pour tout $x \in E$, $\varphi(x, x) = 0 \iff x = 0$

(iii) (Symétrie) : pour tout $(x, y) \in E^2$, $\varphi(x, y) = \varphi(y, x)$.

On note souvant $\langle x, y\rangle$ ou $(x|y)$.

L'expression matricielle d'une forme bilinéaire s'applique aussi aux produits scalaires, et donc, en notant M la matrice du produit scalaire $\langle \cdot, \cdot \rangle$ dans une base $\mathcal{B}$ et X et Y les vecteurs-colonne coordonnées de x et y, le produit scalaire s'exprime :

$$\langle x, y\rangle = X^\top MY.$$

Proposition 12.1.7 – Matrice d'un produit scalaire

La matrice d'un produit scalaire dans une base quelconque est symétrique et inversible. La réciproque est fausse.

Définition 12.1.8 – Espace préhilbertien réel

Un espace préhilbertien réel $(E, \langle \cdot, \cdot \rangle)$ est un espace vectoriel E sur $\mathbb{R}$, muni d'un produit scalaire $\langle \cdot, \cdot \rangle$.

S'il n'y a pas d'ambiguïté sur le produit scalaire, on parlera plus simplement de l'espace préhilbertien E, au lieu de $(E, \langle \cdot, \cdot \rangle)$.

Définition 12.1.9 – Espace euclidien

Un espace euclidien est un espace préhilbertien réel de dimension finie.

I.2 Orthogonalité

On suppose que E est un espace euclidien. La plupart des notions se généralise au cas préhilbertien réel, mais certains propriétés peuvent être fausses en revanche.
On rappelle que deux vecteurs x et y sont orthogonaux si, par définition, $\langle x, y\rangle = 0$.

Définition 12.1.10 – Famille orthogonale, orthonormale

1. Une famille $(x_i)_{i\in I}$ est orthogonale si pour tout $i \neq j$ de I, $x_i \perp x_j$.
2. On dit que la famille $(x_i)_{i\in I}$ est orthonormale (ou orthonormée) si et seulement si elle est orthogonale, et que pour tout $i \in I$, $\|x_i\| = 1$.

Une propriété important de ces familles est leur liberté :

Proposition 12.1.11 – Liberté des familles orthogonales

1. Une famille orthogonale ne contenant pas le vecteur nul est libre.
2. Une famille orthonormale est libre.

Le procédé de Gram-Schmidt permet, à partir d'une famille libre $(e_1, \dots, e_n)$, de construire une famille orthonormale $(f_1, \dots, f_n)$ telle que pour tout $k \in [\![1, n]\!]$

$$\mathrm{Vect}(e_1, \dots, e_k) = \mathrm{Vect}(f_1, \dots, f_k), \qquad \text{et} \qquad \langle e_k, f_k\rangle = 1.$$

En particulier, en partant d'une base d'un espace euclidien, on en déduit :

Théorème 12.1.12 – Existence d'une base orthonormale

Soit E un espace euclidien. Alors E admet une b.o.n.

Plus précisément, en appliquant le procédé de Gram-Schmidt à une base obtenue en complétant une famille orthonormale :

Théorème 12.1.13 – Théorème de la base orthonormale incomplète

Soit E un espace euclidien. Alors toute famille orthonormale peut être complétée en une base orthonormale.

Définition 12.1.14 – Sous-espaces orthogonaux

Les sous-espaces F et G de E sont orthogonaux si pour tout $(x, y) \in F \times G$, $x \perp y$.

Proposition 12.1.15 – Orthogonalité et somme directe

Soit $(F_i)_{i\in I}$ une famille de sous-espaces de E, 2 à 2 orthogonaux. Alors la somme $\bigoplus_{i\in I} F_i$ est directe.

Proposition/Définition 12.1.16 – Supplémentaire orthogonal

Soit E un espace euclidien, et F un sous-espace vectoriel de E. L'ensemble $F^\perp = \{y \in E,\ y \perp F\}$ est un supplémentaire de E, orthogonal à F. Il est appelé supplémentaire orthogonal de F.

Avertissement 12.1.17

C'est faux si E n'est pas de dimension finie (cadre préhilbertien réel au lieu d'euclidien).

I.3 Projection orthogonale

Définition 12.1.18 – Projeté orthogonal sur un sev

Soit F un sous-espace vectoriel de E, et $y \in E$. On dit que $z \in E$ est le projeté orthogonal de y sur F si et seulement si :

(i) $z \in F$

(ii) $(y - z) \perp F$.

Théorème 12.1.19 – Existence du projeté orthogonal sur un sev de dimension finie

Soit $y \in E$, et F un sous-espace vectoriel **de dimension finie** de E, et $\mathcal{B} = (b_1, \dots, b_m)$ une b.o.n. de F. Alors le projeté orthogonal de y sur F existe, est unique, et vaut :

$$z = \sum_{i=1}^{m} \langle y, b_i \rangle b_i.$$

Proposition 12.1.20 – Description de la projection orthogonale

Soit E un espace euclidien et F un sous-espace de E. L'application p qui à x associe son projeté orthogonal sur F est un projecteur de E. Plus précisément, c'est le projecteur sur F parallèlement à $F^\perp$.

I.4 Coordonnées en base orthonormale

Théorème 12.1.21 – Coordonnées d'un vecteur dans une b.o.n. et norme

Soit E un espace euclidien, et $\mathcal{B} = (b_1, \dots, b_n)$ une b.o.n. de E. Soit $x, y \in E$. Alors :

(i) $x = \sum_{i=1}^{n} \langle x, b_i \rangle b_i$, c'est-à-dire $[x]_{\mathcal{B}} = \begin{pmatrix} \langle x, b_1 \rangle \\ \vdots \\ \langle x, b_n \rangle \end{pmatrix}$.

(ii) $\langle x, y \rangle = \sum_{i=1}^{n} \langle x, b_i \rangle \langle y, b_i \rangle$

(iii) $\|x\|^2 = \sum_{i=1}^{n} \langle x, b_i \rangle^2$

Ainsi, si $\mathcal{B}$ est une b.o.n. et si

$$X = [x]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{et} \qquad Y = [y]_{\mathcal{B}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

sont les coordonnées de x et y dans la b.o.n. $\mathcal{B}$, le produit scalaire et la norme s'expriment par les formules usuelles :

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = X^\top Y \qquad \text{et} \qquad \|x\|^2 = \sum_{i=1}^{n} x_i^2 = X^\top X.$$

Ces relations traduisent le fait que la matrice du produit scalaire dans une base orthonormale pour ce produit scalaire est égale à I_n.

Théorème 12.1.22 – Matrice d'un endomorphisme relativement à une b.o.n.

Soit E un espace euclidien muni d'une b.o.n. $\mathcal{B} = (b_1, \dots, b_n)$. Soit $u \in \mathcal{L}(E)$. Alors :

$$\mathrm{Mat}_{\mathcal{B}}(u) = (\langle b_i, u(b_j)\rangle)_{1 \leqslant i,j \leqslant n} = \begin{pmatrix} \langle b_1, u(b_1)\rangle & \cdots & \langle b_1, u(b_n)\rangle \\ \vdots & & \vdots \\ \langle b_n, u(b_1)\rangle & \cdots & \langle b_n, u(b_n)\rangle \end{pmatrix}.$$

I.5 Changements de base orthonormales, matrices orthogonales

Théorème 12.1.23 – propriété des matrices de passage d'une b.o.n. à une autre

Soit $\mathcal{B}$ et $\mathcal{C}$ deux bases de E. Soit $P = P_{\mathcal{B}}^{\mathcal{C}}$ la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{C}$. Alors :

$$P^\top P = I_n = PP^\top, \qquad \text{donc:} \qquad P^{-1} = P^\top.$$

◁ **Éléments de preuve.**
Utiliser la formule de changement de base pour la matrice du ps dans les bases $\mathcal{B}$ et $\mathcal{C}$. On peut aussi le retrouver à l'aide de la desciption explicite de $P_{\mathcal{B}}^{\mathcal{C}}$ et des expressions des $\langle c_i, c_j\rangle$ en b.o.n. $\mathcal{B}$. ▷

Définition 12.1.24 – Matrice orthogonale

Soit $P \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée d'ordre n. On dit que P est une matrice orthogonale si et seulement si $P^\top P = I_n$.

Exemples 12.1.25

1. Dans $\mathcal{M}_n(\mathbb{R})$, les matrices $R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ et $S(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$
2. La matrice I_n, la matrice $-I_n$, et plus généralement, toute matrice obtenue de I_n en changeant les signes de certains de ses coefficients diagonaux, et en permutant les colonnes.

De la définition même découle de façon immédiate :

Proposition 12.1.26 – Inverse d'une matrice orthogonale

Soit P une matrice orthogonale. Alors P est inversible, et $P^{-1} = P^\top$.

Une matrice orthogonale peut se caractériser également par l'orthonormalité de ses colonnes ou ses lignes :

Théorème 12.1.27 – Caractérisation d'une matrice orthogonale par ses colonnes ou lignes

Soit $P \in \mathcal{M}_n(\mathbb{R})$. Les propositions suivantes sont équivalentes:

(i) P est orthogonale

(ii) $P^\top$ est orthogonale

(iii) Les colonnes de P forment une base orthonormale de $\mathbb{R}^n \simeq \mathcal{M}_{n,1}(\mathbb{R})$

(iv) Les lignes de P forment une base orthonormale de $\mathbb{R}^n \simeq \mathcal{M}_{1,n}(\mathbb{R})$.

Théorème 12.1.28 – Caractérisation des matrices de passage entre b.o.n. par orthogonalité

Soit E un espace euclidien.

1. Toute matrice de passage d'une b.o.n. de E à une autre b.o.n. de E est une matrice orthogonale.
2. Réciproquement, soit $\mathcal{B}$ une b.o.n. de E, et P une matrice orthogonale. Alors il existe une

unique base $\mathcal{C}$ telle que P soit la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, et cette base $\mathcal{C}$ est une b.o.n. de E.

3. De même, si $\mathcal{C}$ est une b.o.n. de E, et P une martice orthogonale, il existe une unique base $\mathcal{B}$ telle que P soit la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, et cette base $\mathcal{B}$ est une b.o.n. de E.

◁ **Éléments de preuve.**

1. Résulte de la définition.
2. Remarquer que si x et y sont deux vecteurs de coordonnées X et Y dans une b.o.n, et si φ est le psc, $\langle x, y\rangle = \varphi(X, Y)$. Justifier que ceci permet de définir une b.o.n. dont les coordonnées dans $\mathcal{B}$ sont les colonnes de P.
3. Appliquer ce qui précède à P^{-1} qui est encore orthogonal.

▷

Définition 12.1.29 – Groupe orthogonal

On note $\mathrm{O}_n(\mathbb{R})$, ou $\mathrm{O}(n)$ l'ensemble des matrices orthogonales de $\mathcal{M}_n(\mathbb{R})$. Cet ensemble $\mathrm{O}_n(\mathbb{R})$ est appelé groupe orthogonal.

Théorème 12.1.30 – Structure de $\mathrm{O}_n(\mathbb{R})$

L'ensemble $\mathrm{O}_n(\mathbb{R})$ est un groupe multiplicatif, ce qui est cohérent avec la terminologie introduite dans la définition précédente.

◁ **Éléments de preuve.**

Montrer que c'est un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$; les vérifications sont simples. ▷

Théorème 12.1.31 – Déterminant d'une matrice orthogonale

Soit $P \in \mathrm{O}_n(\mathbb{R})$. Alors $\det(P) \in \{-1, 1\}$. Plus précisément, det est un morphisme de groupe de $\mathrm{O}_n(\mathbb{R})$ dans $(\{-1, 1\}, \times)$.

◁ **Éléments de preuve.**

Appliquer le det à l'égalité $P^\top P = I_n$. ▷

Le noyau de ce morphisme est donc un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Définition 12.1.32 – Groupe spécial orthogonal

Le noyau du déterminant défini sur $\mathrm{O}_n(\mathbb{R})$ est appelé groupe spécial orthogonal, et noté $\mathrm{SO}_n(\mathbb{R})$ ou $\mathrm{SO}(n)$. Ainsi, les éléments de $\mathrm{SO}_n(\mathbb{R})$ sont les matrices orthogonales P telles que $\det(P) = 1$.

Remarque 12.1.33

Si $n \in \mathbb{N}^*$, le sous-groupe $\mathrm{SO}_n(\mathbb{R})$ n'est pas égal à $\mathrm{O}(n)$ tout entier. En effet, la matrice $\mathrm{Diag}(-1, 1, \cdots, 1)$ est dans $\mathrm{O}(n)\backslash\mathrm{SO}(n)$.
Plus précisément, on peut alors montrer qu'il y a « autant » d'éléments dans $\mathrm{SO}(n)$ que dans $\mathrm{O}(n)\backslash\mathrm{SO}(n)$.

Proposition 12.1.34 – Bijection entre $\mathrm{O}(n)$ et $\mathrm{SO}(n)$

Soit $P \in \mathrm{O}(n)\backslash\mathrm{SO}(n)$. L'application

$$Q \mapsto PQ$$

est une bijection de $\mathrm{SO}(n)$ dans $\mathrm{O}(n)\backslash\mathrm{SO}(n)$.

Définition 12.1.35 – Matrices orthogonales directes, indirectes

Soit $P \in \mathrm{O}_n(\mathbb{R})$. On dit que :

- P est une matrice orthogonale positive (ou directe) si $P \in \mathrm{SO}_n(\mathbb{R})$;
- P est une matrice orthogonale négative (ou indirecte) si $P \in \mathrm{O}_n(\mathbb{R}) \backslash \mathrm{SO}_n(\mathbb{R})$.

Terminologie 12.1.36 – Matrices orthogonalement semblables

Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{R})$.

1. On dit que A et B sont orthogonalement semblables s'il existe $P \in \mathcal{O}(n)$ telle que $B = P^{-1}AP = P^\top AP$.
2. On dit que A est orthogonalement diagonalisable si A est orthogonalement semblable à une matrice diagonale.
3. On dit que A est orthogonalement trigonalisable si A est semblable à une matrice triangulaire supérieure.

I.6 Orientation d'un espace

Proposition/Définition 12.1.37 – Orientation d'un espace

Soit E un espace vectoriel sur $\mathbb{R}$.

1. La relation $\mathcal{B}\mathcal{R}\mathcal{B}' \iff \det_{\mathcal{B}}(\mathcal{B}') > 0$ est une relation d'équivalence sur l'ensemble des bases de E
2. Cette relation d'équivalence possède exactement deux classes d'équivalence
3. On dira que $\mathcal{B}$ et $\mathcal{B}'$ ont même orientation si elles appartiennent à la même classe d'équivalence.
4. Une orientation de E est alors le choix arbitraire d'une de ces deux classes d'équivalence, qu'on fait généralement en se donnant une base de référence $\mathcal{B}_0$.
5. Ainsi, étant donné une base de référence $\mathcal{B}_0$, une base $\mathcal{B}$ est directe si $\det_{\mathcal{B}_0}(\mathcal{B}) > 0$, et indirecte si $\det_{\mathcal{B}_0}(\mathcal{B}) > 0$.

Ainsi, orienter l'espace E consiste à choisir une base de référence $\mathcal{B}_0$. Toute base de la même orientation que $\mathcal{B}_0$ sera alors appelée base directe.

Proposition 12.1.38 – Caractérisation de l'orientation pour les b.o.n.

Soit $\mathcal{B}$ et $\mathcal{B}'$ deux b.o.n. d'un espace euclidien E. Les propositions suivantes sont équivalentes:

(i) $\mathcal{B}$ et $\mathcal{B}'$ ont même orientation ;

(ii) $\mathrm{Mat}_{\mathcal{B}}(\mathcal{B}') \in SO(n)$

(iii) $\det_{\mathcal{B}}(\mathcal{B}') = 1$

(iv) $\det_{\mathcal{B}} = \det_{\mathcal{B}'}$.

Définition 12.1.39 – Produit mixte dans un espace euclidien orienté

Soit E un espace euclidien orienté, et $\mathcal{B}$ une b.o.n. directe. Soit $(u_1, \dots, u_n) \in E^n$. Le produit mixte de $(u_1, \dots, u_n)$ est défini par :

$$[u_1, \dots, u_n] = \det_{\mathcal{B}}(u_1, \dots, u_n).$$

Cette définition est indépendante du choix de la b.o.n. directe $\mathcal{B}$.

II Endomorphismes autoadjoints

II.1 Adjoint d'un endomorphisme

Théorème 12.2.1 – Théorème de représentation de Riesz

Soit E un espace euclidien, et $\varphi \in E^*$ une forme linéaire sur E. Il existe un unique vecteur $a \in E$ tel que $\varphi = \langle a, \cdot \rangle$.
Plus précisément, $a \mapsto \varphi_a = \langle a, \cdot \rangle$ est un isomorphisme de E dans E^*.

Proposition/Définition 12.2.2 – Adjoint d'un endomorphisme

Soit E un espace euclidien, et $u \in \mathcal{L}(E)$.

1. Pour tout $x \in E$:
   - l'application $y \mapsto \langle x, u(y) \rangle$ est une forme linéaire ;
   - il existe un unique vecteur $u^*(x) \in E$ tel que pour tout $y \in E$, $\langle x, u(y) \rangle = \langle u^*(x), y \rangle$.
2. L'application $u^* : x \mapsto u^*(x)$ est un endomorphisme de E.
3. L'endomorphisme $u^* \in \mathcal{L}(E)$ est appelé adjoint de l'endomorphisme u.

◁ **Éléments de preuve.**
Le premier point est une vérification élémentaire. Le second provient du théorème de représentation de Riesz.
Le fait que u^* soit un endomorphisme provient de l'unicité dans le théorème de représentation de Riesz, en constatant que $\lambda u^*(x) + u^*(x')$ répond au problème de la représentation de la forme linéaire $\langle \lambda x + x', \cdot \rangle$. ▷

Proposition 12.2.3 – Adjoint d'une composée

Soit E un espace euclidien, et u et v deux endomorphismes de E. Alors

$$(v \circ u)^* = u^* \circ v^*.$$

◁ **Éléments de preuve.**
Vérifier que $\langle u^* \circ v^*(x), y \rangle = \langle x, v \circ u \rangle$, égalité qui caractérise l'adjoint $(v \circ u)^*$. ▷

Proposition 12.2.4 – Double-adjoint

L'application $u \mapsto u^*$ de $\mathcal{L}(E)$ dans lui-même est involutif. En d'autres termes, pour tout $u \in \mathcal{L}(E)$,

$$(u^*)^* = u.$$

◁ **Éléments de preuve.**
Simplifier $\langle (u^*)^*(x), y \rangle$. La symétrie du produit scalaire pourra être utile pour se ramener de façon précise à la définition. ▷

Proposition 12.2.5 – Stabilité par u^* de l'orthogonal d'un sous-espace stable

Soit E un espace euclidien, et $u \in \mathcal{L}(E)$. Soit F un sous-espace de E. Les propositions suivantes sont équivalentes:

(i) F est stable par u

(ii) $F^\perp$ est stable par u^*.

◁ **Éléments de preuve.**

Sens direct : calculer $\langle u^*(x), y\rangle$ pour $x \in F^\perp$ et $y \in F$.

Sens réciproque : cela se ramène au sens direct, par involutivité du passage à l'adjoint et du passage au supplémentaire orthogonal. ▷

Proposition 12.2.6 – Relations entre images et noyaux, HP

Soit E un espace euclidien et $u \in \mathcal{L}(E)$. Alors

$$\mathrm{Ker}(u^*) = (\mathrm{Im}(u))^\perp \qquad \text{et} \qquad \mathrm{Im}(u^*) = \mathrm{Ker}(u)^\perp.$$

◁ **Éléments de preuve.**

La première égalité provient de l'équivalence $x \in \mathrm{Ker}(u^*)$ ssi $\forall y \in E$, $\langle u^*(x), y\rangle = 0$, puis utiliser la définition de l'adjoint.

La seconde égalité se ramène à la première en remplaçant u par u^*. ▷

Corollaire 12.2.7 – Rang d'un adjoint

Soit E un espace euclidien, et u un endomorphisme de E. Alors

$$\mathrm{rg}(u^*) = \mathrm{rg}(u).$$

Une interprétation matricielle permet de relier ce résultat à un résultat déjà connu, démontré de façon algorithmique.

Proposition 12.2.8 – Représentation matricielle d'un adjoint

Soit E un espace euclidien, muni d'une base orthonormale $\mathcal{B}$, et $u \in \mathcal{L}(E)$. Alors

$$\mathrm{Mat}_{\mathcal{B}}(u^*) = \mathrm{Mat}_{\mathcal{B}}(u)^\top.$$

◁ **Éléments de preuve.**

Les formes bilinéaires $(x, y) \mapsto \langle u^*(x), y\rangle$ et $(x, y) \mapsto \langle x, u(y)\rangle$ ont des matrices dans $\mathcal{B}$ respectivement associées à $\mathrm{Mat}_{\mathcal{B}}(u^*)^\top$ et à $\mathrm{Mat}_{\mathcal{B}}(u)$, et sont égales. ▷

L'égalité des rangs de u et u^* provient alors de l'invariance du rang par transposition. Ce résultat est plus général que celui qu'on a prouvé ci-dessus (car valide pour des matrices de tout format), mais la démonstration ci-dessus donne un éclairage géométrique à un résultat qui sinon reste purement algorithmique. D'ailleurs, on pourrait adapter cette démonstration à la situation générale, en considérant des applications linéaires entre 2 espaces euclidiens E et F.

Le déterminant, la trace, le polynôme minimal et le polynôme caractéristique sont aussi invariants par transposition. On obtient donc :

Proposition 12.2.9 – Rang, trace, déterminant, spectre d'un adjoint

Soit E un espace euclidien, et $u \in \mathcal{L}(E)$. Alors :

1. $\mathrm{rg}(u^*) = \mathrm{rg}(u)$
2. $\mathrm{tr}(u^*) = \mathrm{tr}(u)$
3. $\det(u^*) = \det(u)$
4. $\mu_{u^*} = \mu_u$
5. $\chi_{u^*} = \chi_u$

6. $\mathrm{Sp}(u^*) = \mathrm{Sp}(u)$, avec égalité des multiplicités (algébriques).

II.2 Endomorphismes autoadjoints

Une situation particulière importante est le cas des endomorphismes égaux à leur adjoint.

Définition 12.2.10 – Endomorphisme autoadjoint

Soit E un espace euclidien, et $u \in \mathcal{L}(E)$. On dit que u est un endomorphisme autoadjoint (ou endomorphisme symétrique) si $u = u^*$
On note $\mathcal{S}(E)$ l'ensemble des automorphismes auto-adjoints de E.

Remarque 12.2.11

La notation utilisée se réfère à la terminologie « endomorphisme syémtrique », qui s'explique par l'aspect matriciel évoqué ci-dessous. Mais le programmme officiel stipule explicitement de privilégier la terminologie « endomorphisme autoadjoint ».

La propriété de stabilité des supplémentaires orthogonaux se réexprime dans ce contexte de la façon suivante :

Proposition 12.2.12 – Orthogonal d'un sous-espace stable par un endomorphisme autoadjoint

Soit E un espace euclidien, et u en endomorphisme autoadjoint de E. Soit F un sous-espace de E. Les propositions suivantes sont équivalentes:

(i) F est stable par u

(ii) $F^\perp$ est stable par u.

De plus, si cette propriété est vérifiée, les endomorphismes induits u_F et $u_{F^\perp}$ sont eux-même autoadjoints.

Proposition 12.2.13 – Caractérisation des endomorphismes auto-adjoints

Soit E un espace euclidien et $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes:

(i) u est un endomorphisme autoadjoint ;

(ii) $\forall (x,y) \in E^2,\quad \langle u(x), y\rangle = \langle x, u(y)\rangle$;

(iii) pour tout b.o.n. $\mathcal{B}$ de E, $\mathrm{Mat}_{\mathcal{B}}(u) \in \mathcal{S}_n(\mathbb{R})$;

(iv) il existe une b.o.n. $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(u) \in \mathcal{S}_n(\mathbb{R})$.

$\lhd$ **Éléments de preuve.**

- $(i) \iff (ii)$ par définition
- $(i) \implies (iii)$ par description matricielle de l'adjoint.
- $(iii) \implies (iv)$ par existence d'une b.o.n.
- $(iv) \implies (i)$ aussi par description matricielle de l'adjoint.

$\rhd$

Ainsi, les endormophismes autoadjoints sont exactement les endomorphismes dont la matrice est symétrique (réelle) en base orthonormale.
Vous avez peut-être déjà remarqué au cours d'exercices que lorsque vous exprimez la matrice d'un projecteur orthogonal en base orthonormale, vous obtenez une matrice symétrique. Cela n'est pas un hasard, et résulte de la proposition suivante.

Proposition 12.2.14 – Caractérisation des projecteurs orthogonaux parmi les projecteurs

Soit E un espace euclidien, et $p \in \mathcal{L}(E)$ un projecteur de E. Les propositions suivantes sont équivalentes:

(i) p est un projecteur orthogonal

(ii) p est un endomorphisme autoadjoint (et un projecteur par hypothèse)

II.3 Théorème spectral

Dans cette section, nous nous intéressons à la réduction des endomorphismes autoadjoints. Le théorème principal est la diagonalisabilité des endomorphismes autoadjoints, en base orthonormale.
Dans tout ce paragraphe, E est un espace euclidien. Le corps de base est donc $\mathbb{R}$, et le spectre est donc un sous-ensemble de $\mathbb{R}$. Même lors des traductions matricielles, le spectre sera toujours considéré dans $\mathbb{R}$.
Le point clé permettant la diagonalisation des endomorphismes autoadjoints est l'existence d'au moins une valeur propre réelle d'une matrice symétrique réelle.

Lemme 12.2.15 – Spectre d'une matrice symétrique réelle

Soit $M \in S_n(\mathbb{R})$. Alors $\mathrm{Sp}(M) \neq \varnothing$.

◁ **Éléments de preuve.**

Considérer (exceptionnellement) une valeur propre complexe λ (pourquoi en existe-t-il une ?) et un vecteur propre complexe X associé. Considérer $\overline{X}^\top MX = \lambda \overline{X}^\top X$, transposer et conjuguer, pour conclure que $\lambda = \overline{\lambda}$. Qu'en déduire sur les racines du polynôme caractéristique de M ?

C'est la seule incursion dans les complexes, désormais, on travaille exclusivement dans $\mathbb{R}$. ▷

Corollaire 12.2.16 – Spectre d'un endomorphisme autoadjoint

Soit $u \in \mathcal{S}(E)$. Alors $\mathrm{Sp}(u) \neq \varnothing$.

Théorème 12.2.17 – Théorème spectral

Soit E un espace euclidien, et $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes:

(i) $u \in \mathcal{S}(E)$

(ii) $E = \overset{\perp}{\bigoplus\limits_{\lambda \in \mathrm{Sp}(u)}} E_\lambda(u)$

(iii) u est diagonalisable en b.o.n.

(iv) il existe une b.o.n. $\mathcal{B}$ telle que $\mathrm{Mat}_{\mathcal{B}}(u) \in D_n(\mathbb{R})$.

Corollaire 12.2.18 – Diagonalisabilité des matrices symétriques réelles

Soit $M \in \mathcal{M}_n(\mathbb{R})$. Les propositions suivantes sont équivalentes:

(i) $M \in \mathcal{S}_n(\mathbb{R})$

(ii) M est orthogonalement semblable à une matrice orthogonale.

Avertissement 12.2.19

Une matrice symétrique de $\mathcal{M}_n(\mathbb{C})$ n'est pas nécessairement diagonalisable !

II.4 Endomorphismes autoadjoints positifs

Au voisinage d'un point critique, une application de classe $\mathcal{C}^2$ sue un ouvert U de $\mathbb{R}^n$ pourra être approchée localement au second ordre par une application $f(X_0) + H^T SH$, où $H = X - X_0$ et S est une matrice symétrique, appelée matrice hessienne de f. On s'intéressera alors au signe de $H^\top SH$, afin de déterminer si la fonction présente en X_0 un extremum local ou non. Cette expression s'apparente à une expression du type $\langle x, u(x)\rangle$, où u est un endomorphisme autoadjoint. Pour cette raison, l'étude du signe de telles expressions revêt une grande importance.

Définition 12.2.20 – Endomorphismes autoadjoints (définis) positifs

Soit E un espace euclidien, et $u \in \mathcal{S}(E)$.

1. On dit que u est un endomorphisme autoadjoint positif si $\forall x \in E,\ \langle x, u(x)\rangle \geqslant 0$.
   On note $\mathcal{S}^+(E)$ l'ensemble des automorphismes autoadjoints positifs.
2. On dit que u est un endomorphsime autoadjoint défini positif si $\forall x \in E \backslash \{0\},\ \langle x, u(x)\rangle > 0$.
   On note $\mathcal{S}^{++}(E)$ l'ensemble des automorphismes autoadjoints définis positifs.

Proposition 12.2.21 – Caract. spectrale des endomorphismes autoadjoints positifs

Soit $u \in \mathcal{S}(E)$. Les propositions suivantes sont équivalentes:

(i) u est autoadjoint positif

(ii) $\mathrm{Sp}(u) \subset \mathbb{R}_+$

Proposition 12.2.22 – Caract. spectrale des endomorphismes autoadjoints définis positifs

Soit $u \in \mathcal{S}(E)$. Les propositions suivantes sont équivalentes:

(i) u est autoadjoint défini positif

(ii) $\mathrm{Sp}(u) \subset \mathbb{R}_+^*$

Ces définitions et propriétés se traduisent bien matriciellement.

Définition 12.2.23 – Matrices symétriques (définies) positifs

$M \in \mathcal{S}_n(\mathbb{R})$.

1. On dit que M est une matrice syémtrique positive si $\forall X \in \mathcal{M}_{n,1}(\mathbb{R}),\quad X^\top MX \geqslant 0$.
   On note $\mathcal{S}_n^+(\mathbb{R})$ l'ensemble des matrices symétriques positives. M est une matrice syémtrique positive si $\forall X \in \mathcal{M}_{n,1}(\mathbb{R})\backslash\{0\},\quad X^\top MX > 0$.
   On note $\mathcal{S}_n^{++}(\mathbb{R})$ l'ensemble des matrices symétriques définies positives.

Proposition 12.2.24 – Caract. spectrale des matrices syémtriques positives

Soit $M \in \mathcal{S}_n(\mathbb{R})$. Les propositions suivantes sont équivalentes:

(i) M est symétrique positive

(ii) $\mathrm{Sp}(M) \subset \mathbb{R}_+$

Proposition 12.2.25 – Caract. spectrale des matrices syémtriques définies positives

Soit $u \in \mathcal{S}_n(E)$. Les propositions suivantes sont équivalentes:

(i) M est symétrique définie positive

(ii) $\mathrm{Sp}(M) \subset \mathbb{R}_+^*$

Exemples 12.2.26

1. Pour toute matrice $M \in \mathcal{M}_n(\mathbb{R})$, $M^\top$ est symétrique positive. Elle est définie positive ssi $M \in \mathrm{GL}_n(\mathbb{R})$.
2. Une matrice M est symétrique définie positive ssi c'est la matrice d'un produit scalaire relativement à une base $\mathcal{B}$ arbitraire.
3. Une matrice M est symétrique définie positive si et seulement s'il existe $P \in \mathrm{GL}_n(\mathbb{R})$ telle que $M = P^\top P$.
4. Les matrices suivantes sont-elles positives, définies positives ?

$$M_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}, \qquad M_3 = \begin{pmatrix} 1 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 1 \end{pmatrix}.$$

III Isométries vectorielles

III.1 Définitions

Définition 12.3.1 – Isométrie vectorielle

- Soit E un espace euclidien. Une isométrie vectorielle est un endomorphismes $u \in \mathcal{L}(E)$ conservant la norme, c'est-à-dire tel que

$$\forall x \in E, \quad \|u(x)\| = \|x\|.$$

  Une isométrie vectorielle est parfois aussi appelée endomorphisme orthogonal.
- On note $\mathrm{O}(E)$ l'ensemble des isométries vectorielles de E.

Proposition 12.3.2 – Structure de $\mathrm{O}(n)$; groupe orthogonal

$(\mathrm{O}(n), \circ)$ est un sous-groupe de $\mathrm{GL}(E)$. Il est appelé groupe orthogonal de E

◁ **Éléments de preuve.**

- La conservation de la norme et la séparation des normes permet de décrire facilement $\mathrm{Ker}(u)$, pour $u \in \mathrm{O}(n)$. Un argument de dimension montre alors l'inclusion dans $\mathrm{GL}(E)$.
- La structure de sous-groupe se vérifie facilement ensuite.

▷

Exemples 12.3.3

Les endomorphismes de $\mathbb{R}^2$ (muni du ps usuel) canoniquement associés à $R(\theta)$ et $S(\theta)$ sont des isométries. Pouvez-vous les décrire géométriquement en analysant l'image de la base canonique ?

L'exemple de $S(\theta)$ se généralise facilement à tout espace

Définition 12.3.4 – Symétries orthogonales, réflexions

Soit E un espace euclidien, et $s \in \mathcal{L}(E)$.

- s est une symétrie orthogonale ssi il existe un sous-espace F de E tel que s soit la symétrie par rapport à F, parallèlement à $F^\perp$.
- s est une réflexion si et seulement si s est une symétrie orthogonal par rapport à un hyperplan P.

Remarque 12.3.5

- Une symétrie orthogonale par rapport à une droite D est appelée symétrie d'axe D.
- Une symétrie orthogonale (mais il n'y a pas trop le choix...) par rapport à $\{0\}$ est appelée symétrie centrale.

Proposition 12.3.6 – Les symétries orthogonales sont des isométries

Yep !

◁ **Éléments de preuve.**

Décomposer $x = x_F + x_{F^\perp}$, exprimer $s(x)$ à l'aide de cette décomposition, et comparer $\|x\|$ et $\|s(x)\|$ grâce au théorème de Pythagore. ▷

Proposition 12.3.7 – Caractérisations des isométries

Soit $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes:

(i) $u \in \mathrm{O}(n)$

(ii) $\forall (x,y) \in E^2$, $\langle u(x), u(y)\rangle = \langle x, y\rangle$
(caractérisation par conservation du produit scalaire)

(iii) pour toute b.o.n. $\mathcal{B} = (b_1, \dots, b_n)$, la famille $u(\mathcal{B}) = (u(b_1), \dots, u(b_n))$ est encore une b.o.n.
(caractérisation par l'image des bases orthonormales) ;

(iv) il existe une b.o.n. $\mathcal{B} = (b_1, \dots, b_n)$ telle que $u(\mathcal{B}) = (u(b_1), \dots, u(b_n))$ soit encore une b.o.n.
(caractérisation par l'image d'une base orthonormale) ;

(v) pour toute b.o.n. $\mathcal{B} = (b_1, \dots, b_n)$, $\mathrm{Mat}_{\mathcal{B}}(u) \in \mathrm{O}_n(\mathbb{R})$;
(caractérisation par sa matrice dans toute base orthonormale)

(vi) il existe une b.o.n. $\mathcal{B} = (b_1, \dots, b_n)$, telle que $\mathrm{Mat}_{\mathcal{B}}(u) \in \mathrm{O}_n(\mathbb{R})$;
(caractérisation par sa matrice dans une base orthonormale)

(vii) $u^* = u^{-1}$
(caractérisation par l'adjoint)

◁ **Éléments de preuve.**

- $(i) \implies (ii)$ en considérant $\|u(x+y)\|$. La réciproque est évidente.
- $(ii) \implies (iii) \implies (iv) \implies (vi) \implies (v) \implies (vii) \implies (ii)$; ce n'est évidemment pas le seul parcours possible.

▷

Remarques 12.3.8

1. En particulier, le point (ii) assure qu'une isométrie conserve l'orthogonalité. Mais la conservation de l'orthogonalité n'est pas suffisante pour caractériser les isométries.
2. Les points (v) et (vi) justifient la terminologie « endomorphisme orthogonal. Le programme mentionne les deux terminologies, mais stipule de privilégier « isométrie vectorielle ».
3. Les points (v) et (vi) justifient également la ressemblance de notation et la correspondance des terminologies utilisées pour désigner le groupe des isométries et le groupe des matrices orthogonales. Plus précisément, $u \mapsto \mathrm{Mat}_{\mathcal{B}}(u)$ est un isomorphisme entre $\mathrm{O}(E)$ et $\mathrm{O}_n(\mathbb{R})$.

Exemple 12.3.9

- Soit s une symétrie orthogonale. Trouver une b.o.n. dans laquelle $\mathrm{Mat}_{\mathcal{B}}(s)$ est orthogonale. Remarquez que cela propose une alternative à la preuve précédente du caractère isométrique des symétries orthogonales.
- Une symétrie $s \in \mathcal{L}(E)$ est orthogonale ssi $s^* = s$

Ce dernier point permet de caractériser les symétries qui sont des isométries.

Proposition 12.3.10 – Caractérisation des symétries isométriques

Soit s une symétrie de E. Alors s est une isométrie ssi s est une symétrie orthogonale.

Les points (v) et (vi) donnent accès au déterminant d'une isométrie vectorielle

Proposition 12.3.11 – Isométries directes et indirectes

Soit $u \in \mathrm{O}(E)$. Alors $\det(u) \in \{-1, 1\}$.
- Si $\det(u) = 1$, on dit que u est une isométrie directe.
- Si $\det(u) = -1$, on dit que u est une isométrie indirecte.

Proposition/Définition 12.3.12 – Groupe spécial orthogonal

On note $\mathrm{SO}(E)$ l'ensemble des isométries indirectes de E
C'est un sous-groupe de $\mathrm{O}(E)$, appelé groupe spécial orthogonal.

Remarques 12.3.13

- L'égalité $|\det(u)| = 1$ est assez naturelle, puisque le déterminant calcule l'effet sur les volumes de l'endomorphisme u : $\det(u)$ est le coefficient de proportionnalité entre le volume orienté du parallélotope engendré par $(x_1, \dots, x_n)$ et le volume orienté du parallélotope engendré par $(u(x_1), \dots, u(x_n))$.
  Ainsi, l'égalité $|\det(u)| = 1$ ne fait que traduire le fait qu'un isométrie conserve les configurations spaciales, et donc les volumes.
- Les isométries directes sont celles qui conservent l'orientation des b.o.n., les isométries indirectes sont celles qui inverse l'orientation des b.o.n.

Exemple 12.3.14

- Caractériser les syémtries orthogonales qui sont dans $\mathrm{SO}(E)$.
- Une réflexion est-elle une isométrie directe ?
- À quelle condition sur E une symétrie d'axe D est-elle une isométrie directe ?

III.2 Matrices orthogonales de $\mathcal{M}_2(\mathbb{R})$

Dans ce paragraphe, nous donnons la description complète des isométries d'un espace de dimension 2. Au passage, nous définierons la notion d'angle dans un espace euclidien de dimension 2 quelconque, en supposant connue la notion d'angle dans le plan euclidien usuel.
Nous donnons dans cette section une description complète des isométries en dimension 2. Pour cela, nous commençons par déterminer les matrices orthogonales de $\mathcal{M}_2(\mathbb{R})$.

Proposition 12.3.15 – Matrices orthogonales de $\mathcal{M}_2(\mathbb{R})$

(i) Soit $M \in \mathrm{SO}(2)$, alors il existe $\theta \in \mathbb{R}$ tel que $M = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$;

(ii) Soit $M \in \mathrm{O}(2) \backslash \mathrm{SO}(2)$, alors il existe $\theta \in \mathbb{R}$ tel que $M = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$;

Ces descriptions sont uniques modulo 2π.

◁ **Éléments de preuve.**

En notant C_1 et C_2 les colonnes de M, C_1 est de norme 1 et peut donc s'écrire $\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$. Le vecteur C_2 doit être orthogonal à C_1 et unitaire aussi. Cela ne laisse que deux possibilités. ▷

On note :

$$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \qquad \text{et} \qquad S(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}.$$

Proposition 12.3.16 – produits dans $\mathrm{O}(2)$

Soit θ, θ' dans $\mathbb{R}$.

(i) $R(0) = I_2$

(ii) $R(\theta)R(\theta') = R(\theta + \theta') = R(\theta')R(\theta)$

(iii) $R(\theta)^{-1} = R(-\theta)$.

(iv) $S(\theta)S(\theta') = R(\theta - \theta')$

(v) $S(\theta)R(\theta') = S(\theta - \theta')$

(vi) $R(\theta)S(\theta') = S(\theta + \theta')$.

◁ **Éléments de preuve.**

Calcul direct et peu intéressant, basé sur les formules de trigonométrie ▷

En particulier, on reconnaît en $\mathrm{SO}(2)$ un groupe qu'on a déjà rencontré.

Théorème 12.3.17 – $\mathrm{SO}(2)$ est isomorphe à $\mathbb{U}$

L'application qui à $R(\theta)$ associe $e^{i\theta}$ est un isomorphisme de groupe entre $\mathrm{SO}(2)$ et $\mathbb{U}$.
En particulier, $\mathrm{SO}(2)$ est commutatif.

◁ **Éléments de preuve.**

Montrer qu'on peut passer $\theta \mapsto R(\theta)$ au quotient (de $\mathbb{R}$ par $2\pi\mathbb{Z}$). ▷

III.3 Isométries en dimension 2

De l'étude précédente, et de la caractérisation matricielle des isométries, il découle de façon immédiate que l'on sait décrire explicitement toutes les isométries d'un espace euclidien de dimension 2.

Lemme 12.3.18 – Matrices orthogonalement semblables à $R(\theta)$

La seule matrice orthogonalement semblable à $R(\theta)$ est $R(\theta)$ elle-même.

Définition 12.3.19 – Rotation d'angle θ

Soit E un espace euclidien de dimension 2. On dit que $u \in \mathcal{L}(E)$ est une rotation d'angle θ s'il existe une b.o.n. $\mathcal{B}$ telle que $\mathrm{Mat}_{\mathcal{B}}(u) = R(theta)$.
En vertu du lemme précédent, cette définition est indépendante du choix de la b.o.n. $\mathcal{B}$ et on aura, pour toute b.o.n. $\mathcal{C}$, $\mathrm{Mat}_{\mathcal{C}}(u) = R(\theta)$.
On notera ρ_θ la rotation d'angle θ.

D'après les résultats précédents sur les matrices $R(\theta)$, on obtient $\rho_\theta = \rho_{theta'}$ ssi $\theta \equiv \theta'$ $[2\pi]$.

Théorème 12.3.20 – Classification des isométries en dimension 2

Soit E un espace euclidien de dimension 2.

1. Les éléments de $\mathrm{SO}(2)$ sont exactement les rotations
2. Les éléments de $\mathrm{O}(2)\backslash\mathrm{SO}(2)$ sont exactement les réflexions.

◁ **Éléments de preuve.**

Le premier point résulte de la description de $\mathrm{SO}_2(RR)$ et de la définition des rotations.

Pour le second point, une inclusion est déjà acquise. Pour la seconde, constater que $S(\theta)^2 = I_2$, donc les éléments de $\mathrm{O}(2)\backslash\mathrm{SO}(2)$ sont des symétries. Une propriété précédente caractérise les symétries isométriques parmi toutes les symétries. ▷

Les règles de produit matriciel dans $\mathrm{SO}(2)$ amènent directement les règles de composition des rotations, dont l'interprétation géométrique est assez intuitive :

Proposition 12.3.21 – Inverse et composée de deux rotations

(i) $\rho_0 = \mathrm{id}_E$

(ii) Pour tout $(\theta, \theta') \in \mathbb{R}$, $\rho_\theta \circ \rho_{\theta'} = \rho_{\theta+\theta'} = \rho_{\theta'} \circ \rho_\theta$

(iii) Pour tout $\theta \in \mathbb{R}$, $\rho_\theta^{-1} = \rho_{-\theta}$.

Nous pouvons définir la notion d'angle orienté entre deux vecteurs grâce aux rotations. Pour cela, nous utilisons le lemme suivant :

Lemme 12.3.22

Soit E un espace euclidien orienté, et soit x et y deux vecteurs de norme 1 de E. Il existe une unique rotation ρ telle que $\rho(x) = y$.

◁ **Éléments de preuve.**

Si on se fixe une orientation, il existe une unique b.o.n. directe $\mathcal{B}$ de premier vecteur x et une unique b.o.n. directe $\mathcal{C}$ de premier vecteur y. La rotation ρ doit nécessairement envoyer $\mathcal{B}$ sur $\mathcal{C}$. ▷

Définition 12.3.23 – Angle orienté entre deux vecteurs

Soit E un espace vectoriel orienté, et x et y deux vecteurs non nuls de E. Alors l'angle orienté $\widehat{(x,y)}$ est l'angle θ, défini modulo 2π, de l'unique rotation ρ telle que

$$\rho\left(\frac{x}{\|x\|}\right) = \frac{y}{\|y\|}.$$

Proposition 12.3.24 – Réflexion de matrice $S(\theta)$

Soit E un espace euclidien de dimension 2, $s \in \mathrm{O}(E)\backslash\mathrm{SO}(E)$, et $\mathcal{B} = (b_1, b_2)$ une b.o.n. de E. Il existe donc $\theta \in \mathbb{R}$ tel que $\mathrm{Mat}_{\mathcal{B}}(s) = S(\theta)$.
Alors, s est la réflexion par rapport à la droite dirigée par un vecteur unitaire u tel que $\widehat{(b_1, u)} = \frac{\theta}{2}$.

◁ **Éléments de preuve.**

Faire un changement de base pour exprimer s dans l'unique base (u, v) telle que (u, v) soit orthonormale de même orientation que $\mathcal{B}$. Que peut-on dire de la matrice de passage ?

▷

III.4 Réduction des isométries

On cherche maintenant à décrire toutes les isométries de E. On montre plus précisément qu'un endomorphisme est une isométrie si et seulement s'il existe une b.o.n. dans laquelle sa matrice est diagonale par bloc, chaque bloc étant de la forme 1, -1, $R(\theta)$.
Dans cette section, E désigne un espace euclidien.

Lemme 12.3.25 – Stabilité des orthogonaux

Soit u une isométrie, et F un sous-espace stable par u. Alors $F^\perp$ est aussi stable par u.

◁ **Éléments de preuve.**
Passer par u^*, et le fait que u^{-1} est un isomorphisme. ▷

Lemme 12.3.26 – Existence d'un sous-espace strict stable

Soit E de dimension au moins 3, et soit $u \in \mathrm{O}(E)$. Il existe F un sous-espace de E, distinct de $\{0\}$ et de E, tel que F soit stable par u.

◁ **Éléments de preuve.**
Considérer un facteur irréductible P de χ_u, de degré d, et un vecteur $x \in \mathrm{Ker}(P(u))$. Justifier que $\mathrm{Vect}(x, u(x), \dots, u^{d-1}(x))$ est stable par u. ▷

Théorème 12.3.27 – Réduction des isométries

Soit u un isométrie. Il existe une b.o.n. $\mathcal{B}$, des entiers naturels p, q et k, et des réels $\theta_1, \dots, \theta_k \in \mathbb{R}\backslash\pi\mathbb{Z}$ tels que $\mathrm{Mat}_{\mathcal{B}}(u)$ soit diagonale par blocs de la forme

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} I_p & 0 & \cdots & \cdots & 0 \\ 0 & -I_q & \ddots & & \vdots \\ \vdots & \ddots & R(\theta_1) & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & R(\theta_k) \end{pmatrix}$$

Réciproquement, tout endomorphisme se réduisant ainsi en base orthonormale est une isométrie.

◁ **Éléments de preuve.**
Par récurrence forte. Puisque le lemme donnant l'existence d'un sous-espace stable strict est valide uniquement à partir de la dimension 3, il faut initialiser pour E de dimension inférieure ou égale à 2. L'étude précédente nous donne le cas $\dim(E) = 2$. ▷

Corollaire 12.3.28 – Classification des isométries en dimension 3

Toute isométrie d'un espace euclidien de dimension 3 se réduit sous l'une des formes suivantes, en base orthonormale $\mathcal{B} = (b_1, b_2, b_3)$:

1. Isométrie directe : $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$, $\theta \in \mathbb{R}$
   (rotation autour de l'axe $\mathbb{R}b_1$, dans le plan $\mathrm{Vect}(b_2, b_3)$)
2. Isométries indirectes : $\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$
   (rotation autour de l'axe $\mathbb{R}b_1$ et réflexion par rapport à $\mathrm{Vect}(b_2, b_3)$)

13

Séries entières

On étudie dans ce chapitre un cas particulier important de séries de fonctions : les séries entières, de la forme $\sum a_n x^n$. Ces séries apparaissent notamment de façon naturel comme prolongement à l'infini de developpements limités en 0, et on peut donc légitimement se demander si une fonction f de classe $\mathcal{C}^\infty$ peut s'écrire, au moins localement, sous forme d'une série entière, qui correspondrait au prolongement infini de ses développements limités (donc de la formule de Taylor-Young), c'est-à-dire à sa série de Taylor. Nous nous intéresserons à cette question en fin de chapitre, après avoir donné les définitions et les propriétés générales liées à la convergence, ainsi qu'à la régularité de la somme.

Dans tout ce chapitre, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

I Définitions, convergence

I.1 Définition

Définition 13.1.1 – Série entière

- Une série entière est une suite de fonctions $\sum f_n$ où (f_n) est de la forme

$$f_n : x \mapsto a_n x^n, \qquad \text{avec } (a_n) \in \mathbb{K}^{\mathbb{N}}.$$

- On note $\sum a_n z^n$ la série entière associée à la suite (a_n), ou éventuellement $\sum\limits_{n \geqslant n_0} a_n z^n$ si l'indexation commence au rang n_0.
-

Convention 13.1.2 – Variable réelle ou complexe

On utilise généralement la notation $\sum a_n z^n$ lorsqu'on considère la série de la variable complexe $z \in \mathbb{C}$, et $\sum a_n x^n$ ou $\sum a_n t^n$ lorsqu'on considère la série de la variable réelle x ou t dans $\mathbb{R}$ (y compris si les coefficients a_n sont complexes).

Avertissement 13.1.3

Les coefficients a_n doivent être indépendants de la variable z (ou x ou t).

Exemples 13.1.4

Les séries de fonctions suivantes sont-elles des séries entières ?

1. $\sum z^n$
2. $\sum n z^n$

3. $\sum nz^{n-1}$
4. $\sum nz^{n-2}$
5. $\sum\limits_{n\geqslant 1} \ln(n) z^{n^2}$
6. $\sum 2^z z^n$

I.2 Domaine de convergence, rayon de convergence

Nous allons montrer dans ce paragraphe que les séries entières ont un domaine de convergence d'une forme très précise : il s'agit de disques centrés en 0, contenant ou non certains points de leurs bords. Ce disque peut éventuellement être réduit à un point, ou inclure $\mathbb{R}$ ou $\mathbb{C}$ tout entier. Le point clé pour obtenir cette description est le lemme d'Abel.
Dans toute cette section, $\sum a_n z^n$ désigne une série entière sur $\mathbb{K}$.

Lemme 13.1.5 – Abel

Soit $z_0 \in \mathbb{K}$. Si $(a_n z_0^n)_{n\in\mathbb{N}}$ est bornée, alors pour tout $z \in \mathring{B}(0, |z_0|)$, $\sum a_n z^n$ converge absolument.

◁ **Éléments de preuve.**

Écrire $|a_n z^n| = |a_n z_0^n| \cdot \left|\dfrac{z}{z_0}\right|^n$ ▷

Remarque 13.1.6

En particulier, si $\sum a_n z_0^n$ converge, alors $\sum a_n z^n$ converge absolument sur $\mathring{B}(0, |z_0|)$.

Le lemme d'Abel motive la définition suivante.

Définition 13.1.7 – Rayon de convergence

On appelle rayon de convergence de la série $\sum a_n z^n$ l'élément suivant, bien défini dans $\overline{\mathbb{R}}_+$:

$$R = \sup\{r \in \mathbb{R}_+, \quad (a_n r^n)_{n\in\mathbb{N}} \text{ bornée}\}.$$

Exemples 13.1.8

1. Rayon de convergence de $\sum n^\alpha z^n$, $\alpha \in \mathbb{R}$, à partir de la définition.
2. Rayon de convergence de $\sum \frac{z^n}{n!}$
3. Rayon de convergence de $\sum n! z^n$.

Par convention, si $R = +\infty$, $\mathring{B}(0, R)$ désigne $\mathbb{K}$ tout entier, et si $R = 0$, $\mathring{B}(0, R) = \varnothing$.
De la définition-même résulte la propriété suivante.

Proposition 13.1.9 – Rayon de $\sum \lambda a_n z^n$

Soit $\lambda \in \mathbb{K}^*$. Les séries $\sum a_n z^n$ et $\sum \lambda a_n z^n$ ont même rayon de convergence.

Théorème 13.1.10 – Domaine de convergence d'une série entière

Soit $\sum a_n z^n$ une série entière de rayon de convergence R.

1. La série $\sum a_n z^n$ converge absolument sur $\mathring{B}(0, R)$ (i.e. si $|z| < R$).
2. La série $\sum a_n z^n$ diverge grossièrement sur $\mathbb{K}\backslash\overline{B}(0, R)$ (i.e. si $|z| > R$)
3. La convergence pour $|z| = R$ peut varier suivant les séries et suivant la valeur de z.

$\mathring{B}(0, R)$ est appelé disque ouvert (ou intervalle ouvert dans $\mathbb{R}$) de convergence de la série entière

$\sum a_n z^n$, et est parfois aussi noté $D_O(0, R)$.

◁ **Éléments de preuve.**
Pour $z \in \mathring{B}(0, R)$, considérer r tel que $|z| < r < R$, et appliquer le lemme d'Abel. ▷

Remarque 13.1.11

Le théorème de convergence des séries entières donne un moyen de déterminer le rayon de convergence, par étude directe du domaine de convergence. Le rayon R délimite alors la zone de convergence et la zone de divergence. Il n'est pas nécessaire d'étudier toutes les propriétés de convergence : par exemple, trouver un z_0 tel que $\sum a_n z_0^n$ converge, mais $\sum a_n z^n$ diverge si $|z| > |z_0|$ permet de déterminer R.

Exemples 13.1.12

1. Reprendre les exemples précédents.
2. Rayon de convergence de $\sum \ln(n) z^{n^2}$
3. Étude sur le cercle des séries suivantes :
   (a) $\sum z^n$
   (b) $\sum \frac{z^n}{n^2}$
   (c) $\sum \frac{z^n}{n}$.

Remarque 13.1.13

- Si $\sum a_n R^n$ converge absolument, alors $\sum a_n z^n$ est convergente (absolument) sur $\overline{B}(0, R)$.
- Si $\sum a_n R^n$ diverge grossièrement, alors $suma_n z^n$ diverge (grossièrement) sur $\mathbb{K} \setminus \mathring{B}(0, \mathbb{R})$.

Corollaire 13.1.14 – Rayon de convergence associé à une suite de somme convergente

Soit (a_n) telle que $\sum a_n$ converge. Alors le rayon de convergence de $\sum a_n z^n$ est au moins égal à 1.

Nous recontrerons cette situation en probabilités : si X est une variable aléatoire à valeurs dans $\mathbb{N}$, cela permet de s'assurer que le rayon de convergence de $\sum P(X = n) t^n$ est au moins égal à 1. Cette série est appelée série génératrice de la variable aléatoire X.
Nous donnons maintenant quelques moyens de calcul du rayon de convergence. Nous ferons un bilan de ces méthodes en fin de paragraphe.

Proposition 13.1.15 – Règle de d'Alembert

Soit $\sum a_n z^n$ une série entière telle que $\frac{|a_n|}{|a_{n+1}|} \longrightarrow \ell \in \overline{\mathbb{R}}_+$. Alors le rayon de convergence de $\sum a_n z^n$ est $R = \frac{1}{\ell} \in \overline{\mathbb{R}}_+$

Remarque 13.1.16

Ainsi, sous réserve d'existence de cette limite, $R = \lim\limits_{n \to +\infty} \dfrac{|a_{n+1}|}{|a_n|}$.

Le programme officiel stipule que « la limite du rapport $\dfrac{a_{n+1}}{a_n}$ peut être utilisée directement.

Exemple 13.1.17

Rayon de convergence de $\sum \binom{2n}{n} z^n$.

Avertissement 13.1.18

Attention au fait qu'il est important ici que l'indice de a_n corresponde bien à l'exposant en z. Le résultat ne peut pas être mis en oeuvre pour des séries « à trous ».

Exemples 13.1.19

1. Rayon de convergence de $\sum \binom{2n}{n} z^{2n+1}$.
2. Rayon de convergence de $\sum \mathrm{e}^n z^{n^2}$.

Les propriétés suivantes permettent de comparer les rayons de convergence de deux séries entières.

Théorème 13.1.20 – Theorème de comparaison des rayons de convergence

Soit $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières, de rayons respectifs R_a et R_b, et $n_0 \in \mathbb{N}$.

1. Si pour tout $n \geqslant n_0$, $|a_n| \leqslant |b_n|$, alors $R_a \geqslant R_b$.
2. Si $a_n = O(b_n)$ (*a fortiori* si $a_n = o(b_n)$), alors $R_a \geqslant R_b$.
3. Si $a_n \underset{+\infty}{\sim} b_n$, alors $R_a = R_b$.

◁ **Éléments de preuve.**

Pour 1, on peut revenir à la définition. Pour 2, on se ramène à 1, en utilisant la comparaison des rayons de $\sum b_n z^n$ et de $\sum M b_n z^n$. Le 3 résulte directement de 2. ▷

Corollaire 13.1.21

Soit $\sum a_n z^n$ une série entière de rayon de convergence R. Alors $\sum_{n \geqslant n_0} n a_n z^n$ est de rayon de convergence R.

◁ **Éléments de preuve.**

Une inégalité provient de la comparaison $a_n = O(na_n)$.

La deuxième s'obtient par comparaison de $na_n z^n$ à $a_n r^n$, pour $|z| < r < R$. ▷

Exemples 13.1.22

1. $\sum a_n z^n$, $\sum n a_n z^n$ et $\sum\limits_{n \geqslant 1} n a_n z^{n-1}$ ont même rayon de convergence.
2. $\sum a_n z^n$ et $\sum\limits_{n \geqslant 1} \frac{a_n}{n} z^n$ ont même rayon de convergence.
3. Pour toute fraction rationnelle F non nulle, et n_0 assez grand, $\sum_{n \geqslant n_0} F(n) a_n z^n$ a même rayon de convergence que $\sum a_n z^n$.

Méthode 13.1.23 – Déterminer un rayon de convergence

- On pourra toujours commencer par simplifier l'expression en remplaçant a_n par un équivalent (sans condition de signe).
- Dans les situations concrètes, si $\left|\frac{a_{n+1}}{a_n}\right|$ admet une limite dans $\overline{\mathbb{R}}_+$, utiliser la règle de d'Alembert. Veillez à ce que la série soit bien de la forme $\sum a_n z^n$.
- Pour des séries à trous $\sum a_n z^{\varphi(n)}$, on peut aussi essayer d'utiliser la règle de d'Alembert, directement sous cette forme (en considérant $\lim \left|\frac{a_{n+1} z^{\varphi(n+1)}}{a_n z^{\varphi(n)}}\right|$.
- De façon plus générale, étudier le domaine de conevergence par toute autre technique disponible, pour trouver le rayon qui sépare zone de convergence et zone de divergence.
- On peut parfois s'en sortir avec juste quelques valeurs. Si on trouve par exemple z_1 et z_2 tels que $|z_1| = |z_2|$, et tels que $\sum a_n z_1^n$ converge, alors que $\sum a_n z_2^n$ diverge, alors $R = |z_1| = |z_2|$.
- Parfois une unique valeur peut suffire : la divergence non grossière de $\sum a_n z^n$ ou la semi-

convergence de $\sum a_n z^n$ suffisent à savoir que z est sur le cercle limite. Ainsi, la divergence non grossière de $\sum 1n$ suffit à conclure que $\sum \frac{z^n}{n}$ est de rayon de convergence 1.

- On peut aussi revenir à la définition et étudier le caractère borné de $(a_n z^n)$, notamment à l'aide des croissances comparées.
- Enfin, dans des situations un peu plus abstraites on pourra procéder par comparaisons.

I.3 Opérations sur les séries entières

Proposition 13.1.24 – Somme de deux séries entières

Soit $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières de rayons R_a et R_b.

1. Le rayon de convergence R de $\sum (a_n + b_n) z^n$ vérifie $R \geqslant \min(R_a, R_b)$.
2. Si $R_a \neq R_b$, le rayon de convergence R de $\sum (a_n + b_n) z^n$ vérifie l'égalité $R = \min(R_a, R_b)$
3. Si $R_a = R_b$, l'inégalité peut être stricte.

Évidemment, on a alors, pour tout $|z| < \min(R_a, R_b)$,

$$\sum_{n=0}^{+\infty} (a_n + b_n) z^n = \sum a_n z^n + \sum b_n z^n.$$

Proposition 13.1.25 – Produit de Cauchy de deux séries entières

Soit $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières de rayons R_a et R_b. Soit, pour tout $n \in \mathbb{N}$,

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Alors le rayon de convergence R_c de $\sum c_n z^n$ vérifie $R_c \geqslant \min(R_a, R_b)$, et pour tout z tel que $|z| < \min(R_a, R_b)$,

$$\sum_{n=0}^{+\infty} c_n z^n = \sum_{n=0}^{+\infty} a_n z^n \sum_{n=0}^{+\infty} b_n z^n.$$

◁ **Éléments de preuve.**

Les séries étant absolument convergentes, on peut faire le produit de Cauchy de $\sum a_n z^n$ et $\sum b_n z^n$ lorsque $|z| < \min(R_a, R_b)$. ▷

Avertissement 13.1.26

L'inégalité peut être stricte.

Exemple 13.1.27

1. Rayon de convergence du produit de Cauchy de $\sum (-1)^n x^n$ et de $\sum \frac{1}{4^{\lfloor \frac{n}{2} \rfloor}} x^n$.
2. Rayon de convergence du produit de Cauchy de $\sum (\delta_{n,0} + 2^n) z^n$ et $\sum (-1)^{\delta_{n,0}} z^n$.

Ce dernier exemple montre que le rayon de convergence du produit de Cauchy peut même être strictement plus grand que le plus grand des deux rayons (merci Bertrand Hauchecorne pour cet exemple, extrait de *Les contre-exemples en mathématiques*)

Remarque 13.1.28

On peut montrer que si pour tout $n \in \mathbb{N}$, $a_n \geqslant 0$ et $b_n \geqslant 0$, alors $R_c = \min(R_a, R_b)$.

II Régularité de la somme d'une série entière

II.1 Convergence uniforme

Théorème 13.2.1 – Convergence uniforme locale sur le disque ouvert de convergence

Soit $\sum a_n z^n$ une série entière de rayon de convergence R. Alors $\sum a_n z^n$ est normalement, donc aussi uniformément convergence sur tout disque $\overline{B}(0,r)$, avec $z < R$.

◁ **Éléments de preuve.**

Comparer à $\sum |a_n| r^n$ qui converge puisque $\sum a_n z^n$ est absolument convergente dans $\mathring{B}(0,R)$. ▷

Remarque 13.2.2

Si $\sum a_n R^n$ converge absolument, la preuve s'adapte facilement pour montrer que $\sum a_n z^n$ est uniformément convergente sur le disque fermé. $\overline{B}(0,R)$.

Le cas où on a un point du disque en lequel on a une semi-conevergence est plus délicat, mais on peut quand même en dire quelque chose. Quitte à faire une rotation (consistant à remplacer a_n par $a_n \mathrm{e}^{\mathrm{i}\, n\theta}$), on peut supposer qu'on a convergence au point R.

Théorème 13.2.3 – Convergence uniforme sur un rayon en un point de convergence, HP

Soit $\sum a_n z^n$ une série de rayon de convergence $R < +\infty$. On suppose que $\sum a_n R^n$ converge. Alors $\sum a_n z^n$ converge uniformément sur le rayon $[0,R]$

◁ **Éléments de preuve.**

Montrer que le reste $\sum\limits_{k=n+1}^{+\infty} a_k z^k$ converge uniformément vers 0, en faisant une transformation d'Abel, en écrivant $a_k R^k = R_k - R_{k+1}$, où

$$R_n = \sum_{k=n+1}^{+\infty} a_k R^k.$$

▷

Remarque 13.2.4

1. Comme dit plus haut, ce résultat n'a d'intérêt que si la convergence n'est pas absolue. Dans le cas d'une CVA, le résultat obtenu est plus fort
2. Attention à ne pas en conclure que $\sum a_n R^n$ converge uniformément au voisinage de R. Ceci est faux en général, et nécessiterait de maîtriser l'approche de R dans toutes les directions (en restant dans le disque de convergence). Ici, on ne s'approche de R qu'en suivant un rayon.

II.2 Continuité et limite radiale

Les études de convergence uniforme ci-dessus amènent dirextement les propriétés de continuité des séries entières.

Théorème 13.2.5 – Continuité d'une série entière

Soit $\sum a_n z^n$ une série entière de rayon de convergence R, et de somme $A(z)$, pour $z \in \mathring{B}(0,R)$. Alors la fonction A est continue sur $\mathring{B}(0,R)$.

Remarque 13.2.6

Si $\sum a_n R^n$ est absolument convergente, on a même la continuité de A sur $\overline{B}(0,R)$.

Théorème 13.2.7 – Théorème radial d'Abel

Soit $\sum a_n z^n$ une série entière de rayon de convergence R, et de somme $A(z)$, pour $z \in \mathring{B}(0,R)$. Si $\sum a_n R^n$ converge, alors

$$\lim_{\substack{x \to R^- \\ x \in \mathbb{R}}} a_n x^n = \sum a_n R^n$$

◁ **Éléments de preuve.**

La démonstration est HP, pour éviter le théorème de CVU sur un rayon, dont le théorème d'Abel radial est une conséquence immédiate par théorème de la double-limite.

Remarque : même si le théorème de CVU sur un rayon est HP, le théorème d'Abel radial est, lui, bien au programme. ▷

Autrement dit, en combinant les deux résultats précédents, si $\sum a_n z^n$ est de rayon R et $\sum a_n R^n$ converge, la fonction de la variable réelle $x \mapsto \sum a_n x^n$ est continue sur $[0,R]$.

Exemple 13.2.8

On retrouver $\ln(2) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n}$, connaissant le développement du logarithme sur $]-1,1[$. Remarquez que la preuve qu'on en avait donnée reposait sur une transformation d'Abel. Il s'agissait donc en fait de la preuve du théorème d'Abel radial, exprimée dans cette situation particulière.

Remarque 13.2.9

Le programme ne donne pas de résultat de primitivation spécifique aux séries entières, mais la convergence uniforme sur les $\overline{B}(0,r)$, $r < R$, permet d'utiliser le théorème de primitivation (et d'intégration) des suites de fonctions sur tout intervalle $[a,b] \subset]-R,R[$.

II.3 Dérivabilité et classe $\mathcal{C}^\infty$

On rappelle que $\sum a_n z^n$ et $\sum_{n\geqslant 1} n a_n z^{n-1}$ ont même rayon de convergence. On peut itérer.

Proposition 13.2.10 – Rayon de convergence des séries dérivées

Soit $\sum a_n z^n$ une série entière de rayon R. Alors, pour tout $p \in \mathbb{N}$, la série $\sum\limits_{n\geqslant p} n(n-1)\dots(n-p+1)z^{n-p} = \sum\limits_{n\geqslant 0}(n+1)\dots(n+p)z^n$ est de rayon de convergence R.

Pour étudier les propriétés de dérivabilité, on se limite à la variable réelle. On pourrait obtenir des propriétés de dérivabilité en la variable complexe aussi (la dérivée étant aussi définie par limite du taux d'accroissement), mais cela débouche sur la théorie des fonctions holomorphes, et dépasse largement le cadre du programme.

Théorème 13.2.11 – Régularité de la somme d'une série entière

Soit $\sum a_n x^n$ une série entière de la variable réelle x, de rayon de convergence $R > 0$, et de somme égale à $A(x)$, pour $x \in]-R,R[$. Alors A est de classe $\mathcal{C}^\infty$ sur $]-R,R[$, et pour tout $p \in \mathbb{N}$, et tout $x \in]-R,R[$,

$$A^{(p)}(x) = \sum_{n=0}^{+\infty} (n+1)\dots(n+p)x^n.$$

Exemple 13.2.12

On retrouve la formule du binôme négatif, en dérivant p fois la relation

$$\sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}$$

sur $]-1,1[$, cette série entière étant de rayon 1.

Corollaire 13.2.13 – Expression des coefficients d'une série entière

Soit $\sum a_n x^n$ une série entière de la variable réelle x, de rayon de convergence $R > 0$, et de somme égale à $A(x)$, pour $x \in]-R, R[$. Alors pour tout $n \in \mathbb{N}$,

$$a_n = \frac{A^{(n)}}{n!}.$$

Corollaire 13.2.14 – Unicité du développement en série entière

1. Soit $\sum a_n x^n$ et $\sum b_n x^n$ de sommes A et B, définies sur un intervalle $]0, \alpha]$ ($\alpha > 0$), et telles que $A = B$ sur $]0, \alpha]$. Alors pour tout $n \in \mathbb{N}$, $a_n = b_n$.
2. En particulier, soit f une fonction de la variable réelle, à valeurs dans $\mathbb{C}$, définie au moins sur un intervalle $]0, a]$, $\alpha > 0$. S'il existe une suite (a_n) telle que

$$\forall x \in]0, \alpha], \quad f(x) = \sum_{n \geqslant 0} a_n x^n,$$

alors cette suite est unique.

On termine par une étude de dérivabilité au bord, qui sera bien utile en probabilités, pour l'étude des séries génératrices de variables aléatoires, et le rapport avec les calculs d'espérance et de variance.

Proposition 13.2.15 – Dérivabilité au bord, HP ?

Soit $\sum a_n x^n$ une série entière de rayon R, et de somme $A(x)$, pour $x \in]-R, R[$.

1. Si $\sum a_n R^n$ et $\sum n a_n R^n$ convergent, alors A se prolonge par continuité en R en une fonction de classe $\mathcal{C}^1$ sur $]-R, R]$, et

$$A'(R) = \sum_{n=1}^{\infty} n a_n R^{n-1}.$$

2. Réciproquement, on suppose que pour tout $n \in \mathbb{N}$, $a_n \geqslant 0$, que $\sum a_n R^n$ converge, ce qui permet de définir A sur $]-R, R]$. On suppose de plus que A est dérivable à gauche en R. Alors $\sum_{n \geqslant 1} n a_n R^{n-1}$ converge, et

$$A'(R) = \sum_{n=1}^{+\infty} n a_n R^{n-1}.$$

◁ **Éléments de preuve.**

1. Dans le sens direct, c'est le théorème radial d'Abel et le théorème de la classe $\mathcal{C}^1$ par prolongement (corollaire du théorème de la limite de la dérivée).
2. La démonstration du sens réciproque n'est pas exigible. Les hypothèses de positivité permettent d'obtenir la convexité sur $]0, R[$, donc aussi sur $]0, R]$. Majorer alors $\sum_{k=1}^{n} k a_k x^n$ par $A'(R)$, puis passer à la limite quand $x \to R^-$, puis quand $n \to +\infty$.

▷

Remarque 13.2.16

Ce théorème n'est explicitement au programme que dans un contexte de probabilités, avec un rayon $R = 1$, et des coefficients tous positifs. Il doit donc être en théorie redémontré dans toute autre circonstance, notamment le sens direct, lorsqu'on n'a pas l'hypothèse de positivité.

III Développements en séries entières

III.1 Généralités sur les fonctions développables en séries entières

Définition 13.3.1 – Fonction développable en série entière

Soit f une fonction définie au voisinage de 0 (dans $\mathbb{K}$). On dit que f est développable en série entière au voisinage de 0, s'il existe $r > 0$ et une suite $(a_n)_{n\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$ tel que pour tout $z \in \mathring{B}(0, r)$,

$$f_n(z) = \sum_{n\in\mathbb{N}} a_n z^n.$$

Pour un $r > 0$ tel que cette égalité soit satisfaite, on dira que f est développable en série entière sur $\mathring{B}(0, r)$.

Dans cette définition, si R est le rayon de convergence de $\sum a_n z^n$, on a nécessairement $r \leqslant R$, mais pas nécessairement l'égalité.

Proposition 13.3.2 – CN, et unicité du developpement en série entière

Soit f une application définie au voisinage de 0 (dans $\mathbb{R}$ ou $\mathbb{C}$). Alors, pour que f soit développable en série entière, la fonction de la variable réelle $f_{|\mathbb{R}}$, définie au voisinage de 0 doit être de classe $\mathcal{C}^\infty$ et le développement s'écrit alors

$$\forall z \in \mathring{B}(0, r), \quad f(z) = \sum_{n=0}^{+\infty} a_n z^n, \qquad \text{avec } a_n = \frac{f^{(n)}(0)}{n!}.$$

En particulier, en cas d'existence, le développement en série entière est unique.

Remarque 13.3.3 – Lien avec la série de Taylor

1. Lorsque f est DSE de la variable réelle, la série entière associée est nécessairement égale à la série de Taylor $\sum \frac{f^{(n)}(0)}{n!} x^n$.
2. Réciproquement, pour f de classe $\mathcal{C}^\infty$, la série de Taylor en 0 peut converger sur un intervalle $]-r, r[$ sans être égale à f. Dans ce cas, f ne peut pas être développable en série entière.

Exemple 13.3.4

Soit f définie par $f(x) = \mathrm{e}^{-\frac{1}{x^2}}$ si $x \neq 0$, et $f(0) = 0$. Montrer que f est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$, que sa série de Taylor converge sur $\mathbb{R}$ tout entier, mais que f n'est pas développable en série entière au voisinage de 0.

Proposition 13.3.5 – Opérations sur les fonctions développables en séries entières

Soit f et g deux fonctions développables en séries entières au vosinage de 0, et $\lambda \in \mathbb{C}$, dont le DSE est donné respectivement par $\sum a_n z^n$ et $\sum b_n z^n$. Alors :

1. $f + \lambda g$ est développable en série entière, de DSE $\sum (a_n + \lambda b_n) z^n$;

2. fg est développable en série entière, de DSE égal au produit de Cauchy de $\sum a_n z^n$ et $\sum b_n z^n$.

III.2 DSE de l'exponentielle

On commence par redéfinir de façon propre la fonction exponentielle réelle

Proposition/Définition 13.3.6 – Définition de l'exponentielle réelle

Il existe une unique fonction f développable en série entière sur $\mathbb{R}$, vérifiant $f' = f$ et $f(0) = 1$. Cette fonction est appelée exponentielle, notée $\exp(x)$, et vérifie

$$\forall x \in \mathbb{R}, \quad \exp(x) = \sum_{n=0}^{+\infty} \frac{x^n}{n!}.$$

◁ **Éléments de preuve.**

Rechercher f sous la forme $\sum a_n x^n$, dériver et injecter dans l'équation différentielle. La relation sur (a_n) et la condition en 0 déterminent entièrement (a_n). ▷

Remarque 13.3.7

1. Retenir cette méthode de recherche de développement en s'aidant d'une équation différentielle. Cela permet de trouver de façon efficace le DSE de fonctions solutions d'équations différentielles. Il est assez fréquent de commencer par rechercher ainsi les solutions développables en série entière d'une ED, qui aident souvent à trouver les autres.
2. Une fois l'existence d'une telle fonction obtenue, toutes les propriétés vue en première année sont valides. En particulier, les propriétés liées à la résolution des équations différentielles d'ordre 1. En particulier, le théorème de Cauchy Lipschitz pour les équations différentielles d'ordre 1 est vérifié, ce qui assure en particulier l'unicité de la solution de $f' = f$, tel que $f(0) = 1$. On en déduit que la fonction exp est non seulement unique parmi les fonctions DSE, mais également parmi toutes les focntions dérivables.

Proposition 13.3.8 – DSE de cos et sin

La fonction $\varphi : t \mapsto \cos(t) + \mathrm{i}\sin(t)$ vérifie l'équation différentielle $\varphi' = \mathrm{i}\,\varphi$, et $\varphi(0) = 1$, donc l'unique solution développable en série entière est

$$t \mapsto \sum_{n=0}^{+\infty} \frac{(\mathrm{i}\,t)^n}{n!}.$$

Ainsi cos et sin sont DSE, et

$$\forall t \in \mathbb{R}, \quad \cos(t) = \sum_{n=0}^{+\infty} (-1)^n \frac{t^{2n}}{(2n)!} \qquad \text{et} \qquad \sin(t) = \sum_{n=0}^{+\infty} (-1)^n \frac{t^{2n+1}}{(2n+1)!}.$$

De plus, on définit, pour $t \in \mathbb{R}$,

$$\exp(\mathrm{i}\,t) = \mathrm{e}^{\mathrm{i}\,t} = \sum_{n=0}^{+\infty} \frac{(\mathrm{i}\,t)^n}{n!} = \cos(t) + \mathrm{i}\sin(t).$$

◁ **Éléments de preuve.**

Même technique que pour exp, mais maintenant on connaît par avance l'unicité de la solution de cette ED. ▷

Remarque 13.3.9

On peut aussi trouver les DSE de cos et sin en montrant que le reste de Taylor-Lagrange tend vers 0, par exemple avec l'inégalité de Taylor-Lagrange.

La définition suivante est un rappel :

Définition 13.3.10 – Exponentielle complexe

Pour $z = x + \mathrm{i}\, y$, avec $(x, y) \in \mathbb{R}^2$, on définit

$$\exp(z) = \mathrm{e}^z = \mathrm{e}^x \mathrm{e}^{\mathrm{i}\, y}.$$

Théorème 13.3.11 – DSE de l'exponentielle complexe

La fonction exp est DSE sur $\mathbb{C}$ et

$$\exp(z) = \sum_{n=0}^{+\infty} \frac{z^n}{n!}.$$

◁ **Éléments de preuve.**

Par produit de Cauchy. ▷

III.3 DSE usuels

Définition 13.3.12 – Coefficient binomial étendu

Soit $\alpha \in \mathbb{C}$ et $n \in \mathbb{N}$. On note $\displaystyle\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(a-n+1)}{n!}$.

Proposition 13.3.13 – DSE usuels

1. $\displaystyle \mathrm{e}^z = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$ pour $z \in \mathbb{C}$;
2. $\displaystyle \frac{1}{1-z} = \sum_{n=0}^{+\infty} z^n$ pour $z \in \mathbb{C}$, $|z| < 1$
3. $\displaystyle \frac{1}{(1-z)^p} = \sum_{n=0}^{+\infty} \frac{(n+1)\ldots(n+p-1)}{p!} z^n = \sum_{n=0}^{+\infty} (-1)^n \binom{-p}{n} z^n$ pour $z \in \mathbb{C}$, $|z| < 1$, $p \in \mathbb{N}^*$.
4. $\displaystyle (1+x)^\alpha = \sum_{n=0}^{+\infty} \binom{\alpha}{n} x^n$ pour $x \in]-1, 1[$, $\alpha \in \mathbb{C}$.
5. $\displaystyle \ln(1+x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} \cdot x^n$, pour $x \in]-1, 1[$,
6. $\displaystyle \cos(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!}$, pour $x \in \mathbb{R}$;
7. $\displaystyle \sin(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$, pour $x \in \mathbb{R}$;
8. $\displaystyle \mathrm{ch}(x) = \sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!}$, pour $x \in \mathbb{R}$;
9. $\displaystyle \mathrm{sh}(x) = \sum_{n=0}^{+\infty} \frac{x^{2n+1}}{(2n+1)!}$, pour $x \in \mathbb{R}$;

10. $\text{Arctan}(x) = \sum_{n=0}^{+\infty} \frac{x^{2n+1}}{2n+1}$, pour $x \in]-1,1[$.

$\triangleleft$ **Éléments de preuve.**

1. Vu au paragraphe précédent
2. Somme géométrique.
3. Produits de Cauchy
4. Résolution de l'ED $(1+x)y' = \alpha y$.
5. Par intégration (ou, ce qui revient en même, en dérivant cette série entière).
6. Vu précédemment
7. Vu précédemment
8. À partir du DSE de l'exponentielle
9. Idem
10. Par intégration.

$\triangleright$

14

Espaces probabilisés

Le cours de probabilité de première année se contruit entièrement dans le cadre d'un univers Ω fini.
Dans ce cadre, une expérience d'univers des possibles Ω est un processus aléatoire amenant un résultat (une issue) ω appartenant à Ω. Un évènement A est alors un sous-ensemble de Ω. Dire que A lors d'une expérience signifie que l'expérience débouche sur une issue ω appartenant à A.
Nous voyons dans ce chapitre comment élargir ce cadre, en acceptant notamment des univers infinis. Ce cadre formel un peu abstrait est indispensable pour pouvoir définir de façon rigoureuse une mesure de probabilité.

I Espaces probabilisés

Dans cette section, Ω désigne un ensemble quelconque, fini ou infini.

I.1 Tribus et espaces probabilisables

Souvent, la définition de la mesure de probabilité sur tous les sous-ensembles de Ω, n'est pas pertinente, ou peut poser des problèmes techniques. Pour cette raison, il peut être intéressant de pouvoir ne définir la mesure de probabilité que sur une classe particulière de sous-ensembles de Ω. L'objet de ce paragraphe et du suivant est de définir un type satisfaisant de classe de sous-ensembles à laquelle se restreindre.
Voici nos exigences pour la construction d'une telle classe :

- L'événement certain Ω et l'événement impossible ont une probabilité, donc appartiennent à la classe ;
- Pour tout événement A dont on sait calculer la probabilité, on voudrait pouvoir calculer la probabilité de l'événement contraire $\overline{A}$;
- Si on dispose d'une probabilité sur A et B, on voudrait disposer d'une probabilité de $A \cup B$.
- Plus généralement, on souhaite pouvoir calculer la probabilité d'une union infinie dénombrable d'événements.

Exemple 14.1.1

Le dernier point est motivé par l'étude d'un tirage infini à Pile ou Face : quelle est la probabilité qu'un tirage donné (par exemple n'obtenir que des Piles) ait lieu ? Pour ce calcul, on peut encore se dispenser des unions (ou intersections par passage au complémentaire) infinies, en effectuant des majorations, si on définit correctement la probabilité ; mais en revanche, si on se demande maintenant quelle est la probabilité qu'un résultat soit stationnaire, on a besoin de regrouper une infinité de cas possibles (stationnaires à partir d'un certain rang).

Cela nous amène à la définition suivante :

Définition 14.1.2 – σ-algèbre, ou tribu

Soit Ω un univers (fini ou non). Une *σ-algèbre* $\mathcal{T}$ d'événements sur Ω (ou tribu) est un sous-ensemble $\mathcal{T}$ de $\mathcal{P}(\Omega)$ telle que :

1. $\Omega \in \mathcal{T}$;
2. $\forall A \in \mathcal{P}(\Omega),\ A \in \mathcal{T} \Longrightarrow \overline{A} \in \mathcal{T}$ (stabilité par complémentation) ;
3. Pour tout famille dénombrable $(A_i)_{i\in I}$ d'éléments de $\mathcal{T}$, $\displaystyle\bigcup_{i\in I} A_i$ est dans $\mathcal{T}$ (stabilité par union dénombrable)

Nous complétons les propriétés imposées par la définition par les suivantes :

Proposition 14.1.3 – Propriétés des tribus

Soit $\mathcal{T}$ une tribu d'événements sur Ω.

1. $\varnothing \in \mathcal{T}$;
2. $\forall (A,B) \in \mathcal{P}(\Omega),\ (A,B) \in \mathcal{T}^2 \Longrightarrow A \cup B \in \mathcal{T}$ (stabilité par union finie) ;
3. Pour tout famille dénombrable $(A_i)_{i\in I}$ d'éléments de $\mathcal{T}$, $\displaystyle\bigcap_{i\in I} A_i \in \mathcal{T}$.
4. $\forall (A,B) \in \mathcal{P}(\Omega),\ (A,B) \in \mathcal{T}^2 \Longrightarrow A \cap B \in \mathcal{T}$ (stabilité par intersection finie) ;

◁ **Éléments de preuve.**
Le point 2 se montre en complétant la famille (A,B) en une famille dénombrable, en ajoutant l'ensemble vide, ce qui ne change pas l'union. Les autres s'obtiennent par passage au complémentaire.
▷

Par une récurrence immédiate, les points 2 et 4 se généralisent facilement aux unions ou intersections d'un nombre fini quelconque d'événements.

Définition 14.1.4 – Espace probabilisable et événements

- Un *espace probabilisable* est un couple $(\Omega, \mathcal{T})$, où Ω est un ensemble quelconque, appelé univers, et $\mathcal{T}$ une σ-algèbre d'événements sur Ω.
- Les éléments de $\mathcal{T}$ sont appelés *événements* (de l'espace probabilisable $(\Omega, \mathcal{T})$).

Remarque 14.1.5

1. Une tribu est aussi fréquemment notée $\mathcal{A}$ au lieu de $\mathcal{T}$ (en référence à la terminologie σ-algèbre). C'est la notation utilisée dans le programme officiel
2. « La manipulation des tribus n'est pas un objectif du programme », dixit le programme officiel. La notion est surtout donnée pour avoir le bon cadre de travail. Ce cadre sera souvent fourni dans les exercices ou problème, sans que vous ayez à vous demander précisément à quoi il correspond.
3. En particulier, toute la très riche théorie des tribus, indispensable à une théorie un peu plus poussée des probabilités, ne sera pas abordée.

Exemple 14.1.6

$\mathcal{P}(\Omega)$ est une tribu sur Ω. Dans le cas où Ω est fini ou dénombrable, on choisira en général cette tribu. Elle s'avère en général trop riche lorsque Ω n'est pas dénombrable, et on devra dans ces cas définir des tribus plus petites pour éviter certains écueils techniques.

On rappelle alors le vocabulaire probabiliste usuel, qui est le même que dans le cas fini

Terminologie 14.1.7 – Vocabulaire probabiliste

Soit $(\Omega, \mathcal{T})$ un espace probabilisé. On définit les terminologies ci-dessous.

- *Évément* : un objet $A \in \mathcal{T}$
- *Événement élémentaire*, ou *épreuve* : un événement de la forme $\{om\}$ (attention, tout singleton n'est pas nécessairement un événement)
- *Événement certain* : l'événement Ω.
- *Événement impossible* : l'événement $\varnothing$.
- *Événement contraire de l'événement A* : l'événement $\overline{A} = \complement_\Omega A$.
- *A entraîne B si $A \subset B$*
- *Événement « A et B »* : $A \cap B$ (réalisation simultanée)
- *Événement « A ou B »* : $A \cup B$
- A et B sont dits *incompatibles* si $A \cap B = \varnothing$.
- une famille $\{A_i, i \in I\}$ est dite *constituée d'événements deux à deux incompatibles* si pour tout $(i,j) \in I^2$ tel que $i \neq j$, A_i et A_j sont incompatibles.

Définition 14.1.8 – Système complet d'événements, SCE

Un système complet d'événements est une famille $\{A_i, i \in I\}$ formant une partition de Ω. Autrement dit :

- Les événements A_i sont non vides ;
- La famille est constituée d'événements deux à deux incompatibles ;
- $\biguplus\limits_{i \in I} A_i = \Omega$.

I.2 Mesures de probabilité, espaces probabilisés

Intuitivement, une probabilité mesure la possibilité qu'un événement donné se produise. Il s'agit donc d'un réel p indiquant la fréquence probable de réalisation de l'événement lorsqu'on répète un grand nombre de fois l'expérience dans des situations similaires, c'est-à-dire le rapport moyen entre le nombre de réalisations et le nombre d'expériences.

Ainsi, une probabilité est un réel positif et inférieur à 1. De plus, toutes les expériences amenant une issue de Ω, la probabilité de l'événement Ω doit être 1. Enfin, si A et B sont deux événements ne pouvant pas être réalisés simultanément, le nombre d'expériences amenant A ou B est la somme du nombre d'expériences amenant A et du nombre d'expériences amenant B (puisqu'une expérience ne peut amener à la fois A et B). De l'interprétation ci-dessus, il doit alors venir que la probabilité de l'union disjointe $A \sqcup B$ doit être la somme des probabilités de A et de B. Cela motive la définition suivante :

Définition 14.1.9 – Mesure de probabilité

Soit $(\Omega, \mathcal{T})$ un espace probabilisable. Une mesure de probabilité (ou simplement une probabilité) sur $(\Omega, \mathcal{T})$ est une application $\mathbb{P} : \mathcal{T} \to \mathbb{R}$ telle que :

1. $\forall A \in \mathcal{T}$, $0 \leqslant \mathbb{P}(A) \leqslant 1$;
2. $\mathbb{P}(\Omega) = 1$
3. (propriété de σ-additivité) Pour tout famille dénombrable $(A_i)_{i \in I}$ d'événements deux à deux incompatibles, $(\mathbb{P}(A_i))_{i \in I}$ est sommable, et on a : $\mathbb{P}\left(\biguplus\limits_{i \in I} A_i\right) = \sum\limits_{i \in I} \mathbb{P}(A_i)$.

Remarque 14.1.10

Très souvent $I = \mathbb{N}$. Dans ce contexte, la famille $(\mathbb{P}(A_n))_{n \in \mathbb{N}}$ étant positive, sa sommabilité équivaut à la convergence de la série $\sum \mathbb{P}(A_n)$.

Définition 14.1.11 – Espace probabilisé, ou modèle probabiliste de Kolmogorov

Un *espace probabilisé* est un triplet $(\Omega, \mathcal{T}, \mathbb{P})$ où $(\Omega, \mathcal{T})$ est un espace probabilisable, et $\mathbb{P}$ une mesure de probabilité sur $(\Omega, \mathcal{T})$.

Proposition 14.1.12 – Propriétés d'une mesure de probabilité

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé. Alors :

1. $\mathbb{P}(\varnothing) = 0$;
2. *(Additivité)* Pour tout couple (A, B) d'événements incompatibles, $\mathbb{P}(A \uplus B) = \mathbb{P}(A) + \mathbb{P}(B)$;
3. $\forall A \in \mathcal{T}$, $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$;
4. $\forall (A, B) \in \mathcal{T}$, $A \subset B \Longrightarrow \mathbb{P}(B \backslash A) = \mathbb{P}(B) - \mathbb{P}(A)$;
5. $\forall (A, B) \in \mathcal{T}$, $A \subset B \Longrightarrow \mathbb{P}(A) \leqslant \mathbb{P}(B)$;
6. $\forall (A, B) \in \mathcal{T}$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;

◁ **Éléments de preuve.**

$\mathbb{P}(\varnothing)$ s'obtient en étudiant la convergence de $\sum \mathbb{P}(\varnothing)$, convergence assurée par la σ-additivité. L'additivité en découle en complétant la famille finie (A, B) en une famille dénombrable par ajout de $\varnothing$. Les points 3,4,5,6 sont évident, en décomposant le plus grand ensemble en union disjointe d'ensembles plus petits. ▷

Proposition 14.1.13 – Sous-additivité finie de la mesure de probabilité

Soit $(A_i)_{i \in I}$ une famille finie d'événements. Alors

$$P\left(\bigcup_{i \in I} A_i\right) \leqslant \sum_{i \in I} P(A_i).$$

◁ **Éléments de preuve.**

Sans perte de généralité, considérer $I = [\![1, n]\!]$, et faire une récurrence sur $\mathbb{N}$ à partir de l'expression de $\mathbb{P}(A \cup B)$ qui nous donne le cas $n = 2$. ▷

On verra un peu plus loin que cette propriété se généralise au cas dénombrable.

Proposition 14.1.14 – Probabilités associées à un SCE

Soit $(A_i)_{i \in I}$ un SCE au plus dénombrable. Alors $\sum_{i \in I} \mathbb{P}(A_i) = 1$.

I.3 Espaces probabilisés discrets

Il est fréquent que, même si Ω est de grand cardinal (infini non dénombrable), il existe un sous-ensemble A de Ω, au plus dénombrable, tel que presque sûrement, toutes les issues soient dans A (autrement dit, si A est lui-même un élément de la tribu, $\mathbb{P}(A) = 1$)

Définition 14.1.15 – Espace probabilisé discret

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé tel que $\mathcal{T}$ contienne tous les singletons. On dit que $(\Omega, \mathcal{T}, \mathbb{P})$ est discret si la famille $(\mathbb{P}(\omega))_{\omega \in \Omega}$ est sommable de somme égale à 1

Remarque 14.1.16

Un résultat sur les familles sommables (démontré dans le chapitre sur la dénombrabilité) montre alors que $\{\omega \mid \mathbb{P}(\omega) \neq 0\}$ est au plus dénombrable, ce qui justifie la terminologie.

Définition 14.1.17 – Distribution discrète sur Ω

Soit Ω un ensemble. Une distribution de probabilités sur Ω est une famille $(p_\omega)_{\omega\in\Omega}$ à valeurs dans $\mathbb{R}_+$ telle que

$$\sum_{\omega\in\Omega} p_\omega = 1.$$

L'ensemble $\{\omega \in \Omega,\ p_\omega \neq 0\}$ est appelé support de la distribution de probabilités $(p_\omega)_{\omega\in\Omega}$, et est au plus dénombrable, en vertu de la remarque précédente.

Ainsi, $(\Omega, \mathcal{T}, \mathbb{P})$ est un espace probabilisé discret si et seulement si $\mathcal{T}$ contient les singletons, et si $(\mathbb{P}(\omega))_{\omega\in\Omega}$ est une distribution de probabilités discrète sur Om.

Proposition/Définition 14.1.18 – Probabilité définie par une distribution discrète

Soit Ω un ensemble quelconque, et $(p_\omega)_{\omega\in\Omega}$ une distribution de probabilités discrète. On définit, pour $A \in \mathcal{P}(\Omega)$,

$$\mathbb{P}(A) = \sum_{\omega\in A} p_\omega,$$

bien défini dans $\overline{\mathbb{R}}_+$.
L'application $\mathbb{P}$ est à valeurs dans $[0,1]$, et définit une mesure de probabilités sur l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$.

Proposition 14.1.19 – Détermination d'une probabilité sur un univers dénombrable

Soit Ω un ensemble fini ou dénombrable. Une mesure de probabilités sur l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$ est entièrement déterminée par la distribution de probabilités $(\mathbb{P}(\{\omega\}))_{\omega\in\Omega}$

◁ **Éléments de preuve.**

Unicité car les singletons engendrent $\mathcal{P}(\Omega)$. Expliciter p ainsi obtenue en fonction des probabilités des singletons, et vérifier qu'il s'agit bien d'une mesure de probabilité. ▷

Ainsi, pour définir une mesure de probabilité sur Ω fini ou dénombrable, il suffit de se donner une distribution de probabilités $(p_\omega)_{\omega\in\Omega}$.

Exemple 14.1.20 – Probabilité uniforme sur un univers fini

Si Ω est fini, on peut définir la mesure de probabilité uniforme en posant $\mathbb{P}(\{om\}) = \frac{1}{|\Omega|}$. On obtient alors, pour tout événement $A \in \mathcal{P}(\Omega)$,

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} \qquad \text{(formule de Laplace)}$$

I.4 Continuité monotone

Le théorème suivant découle directement de la σ-additivité.

Théorème 14.1.21 – Propriété de limite monotone, ou continuité monotone de $\mathbb{P}$

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé.

1. Soit $(A_n)_{n\in\mathbb{N}}$ une suite croissante d'événements (pour l'inclusion). Alors :

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$

2. Soit $(A_n)_{n\in\mathbb{N}}$ une suite décroissante d'événements. Alors :

$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$

◁ **Éléments de preuve.**
Écrire $\bigcup\limits_{n\in\mathbb{N}} A_n = \biguplus\limits_{n\in\mathbb{N}} (A_n \backslash A_{n-1})$, avec $A_{-1} = \varnothing$, puis σ additivité. Passer au complémentaire pour l'intersection. ▷

Exemple 14.1.22

Lors d'une succession infinie de tirages à Pile ou Face, calculer la probabilité de n'obtenir que des Piles.

Corollaire 14.1.23

Soit $(A_n)_{n\in\mathbb{N}}$ une famille quelconque d'événements. Alors

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{N\to+\infty} \mathbb{P}\left(\bigcup_{n=0}^{N} A_n\right) \qquad \text{et} \qquad \mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{N\to+\infty} \mathbb{P}\left(\bigcap_{n=0}^{N} A_n\right).$$

Corollaire 14.1.24 – Sous-additivité dénombrable de $\mathbb{P}$, ou inégalité de Boole

Soit $(A_i)_{i\in I}$ une famille au plus dénombrable d'événements. Alors, dans $\overline{\mathbb{R}}_+$,

$$\mathbb{P}\left(\bigcup_{i\in I} A_i\right) \leqslant \sum_{i\in I} \mathbb{P}(A_i).$$

En définissant pour une suite croissante (A_n), $\bigcup\limits_{n\in\mathbb{N}} A_n = \lim \uparrow A_n$ et de même pour une suite décroissante, $\bigcap\limits_{n\in\mathbb{N}} A_n = \lim \downarrow A_n$, la propriété ci-dessus se réécrit :

$$\mathbb{P}(\lim \uparrow A_n) = \lim \mathbb{P}(A_n) \qquad \text{et} \qquad \mathbb{P}(\lim \downarrow A_n) = \lim \mathbb{P}(A_n).$$

Il s'agit donc d'une propriété de continuité séquentielle, d'où le nom donné au théorème.

Exemple 14.1.25 – Limite supérieure et limite inférieure, HP

Soit (A_n) une suite d'événements. On définit

- la limite supérieure $\limsup A_n = \bigcap\limits_{n\in\mathbb{N}} \bigcup\limits_{k=n}^{\infty} A_k$: réalisation d'une infinité de A_k ;
- la limite inférieure $\liminf A_n = \bigcup\limits_{n\in\mathbb{N}} \bigcap\limits_{k=n}^{\infty} A_k$: réalisation de presque tous les A_k (tous sauf un nombre fini, c'est-à-dire tous à partir d'un certain rang).

Ainsi,

- La probabilité qu'un nombre infini de A_k se réalisent est

$$\mathbb{P}(\limsup A_n) = \lim \mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right).$$

- La probabilité que presque tous les A_k se réalisent est

$$\mathbb{P}(\liminf A_n) = \lim \mathbb{P}\left(\bigcap_{k=n}^{\infty} A_k\right).$$

Par exemple, calculer la probabilité que lors d'une suite infinie de tirages à P/F, la suite obtenue soit stationnaire.

Corollaire 14.1.26

Soit (A_n) une suite d'événements.

- La probabilité qu'un nombre infini de A_k se réalisent est

$$\mathbb{P}(\limsup A_n) = \lim \mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right).$$

- La probabilité que presque tous les A_k se réalisent est

$$\mathbb{P}(\liminf A_n) = \lim \mathbb{P}\left(\bigcap_{k=n}^{\infty} A_k\right).$$

Exemple 14.1.27

On tire une infinité de fois à P ou F. Probabilité que la suite de P/F soit stationnaire.

Remarque 14.1.28

Si (A_n) est décroissante, $\bigcup_{k=n}^{+\infty} A_k = A_n$ et $\limsup A_n = \lim \downarrow A_n$; de même si (A_n) est croissante, $\liminf A_n = \lim \uparrow A_n$. Le corollaire précédent englobe donc le théorème de continuité monotone.

I.5 Évenements négligeables, presque sûrs

Définition 14.1.29 – Ensembles négligeables

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé, et soit $A \in \mathcal{T}$.

1. On dit que A est *négligeable*, ou *presque-impossible*, ou *quasi-impossible* si $\mathbb{P}(A) = 0$.
2. On dit que A est *presque sûr*, ou *presque-certain* ou *quasi-certain* si $\mathbb{P}(A) = 1$.
3. Une propriété est vraie *presque sûrement* si l'événement « la propriété est satisfaite » est presque-certain.

Exemple 14.1.30

Les événements « n'obtenir que des P » lors d'une succession infinie de tirages P/F, ou « obtenir une suite stationnaire » sont quasi-impossibles, sans être impossibles.
Leur complémentaire est donc presque sûr, sans être certain.

Remarque 14.1.31

La terminologie recommandée par le programme officiel est « négligeable » et « presque sûr », mais les autres terminologies sont aussi très répandues dans les ouvrages.

Proposition 14.1.32 – Intersection et union d'événements presques impossibles/sûrs

1. Toute union d'un nombre au plus dénombrable d'événements négligeables est négligeable.
2. Toute intersection d'un nombre au plus dénombrable d'événements presque sûrs est presque sûr.

◁ **Éléments de preuve.**
Par sous-additivité dénombrable, puis complémentation. ▷

Remarques 14.1.33

1. Dans $(\Omega, \mathcal{T}, \mathbb{P})$, par convention, si I est vide,
$$\bigcup_{i\in I} A_i = \varnothing \qquad \text{et} \qquad \bigcap_{i\in I} A_i = \Omega.$$
Ainsi, les propriétés sont aussi vraies pour des familles vides.
2. À condition d'enlever le cas des familles vides, la propriété est aussi trivialement vraie pour les intersections quelconques d'ensembles négligeables, et les unions quelconques d'ensembles presque sûrs.

On définit maintenant une notion voisine des SCE, mais qui permet plus de souplesse dans les hypothèses de certains théorèmes.

Définition 14.1.34 – Système quasi-complet d'événements, SQCE

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisable, et $(A_i)_{i\in I}$ une famille au plus dénombrable d'éléments de $\mathcal{T}$. On dit que $(A_i)_{i\in I}$ est un système quasi-complet d'événements si :

(i) les A_i sont 2 à 2 incompatibles ;

(ii) $\mathbb{P}\left(\bigcup_{i\in I} A_i\right) = 1$.

Ainsi, la différence avec un système complet réside dans le fait que certaines parts peuvent être vides, et que l'union $\bigcup_{i\in I} A_i$ n'est égale à Ω qu'à un ensemble négligeable près.

Un système complet est bien entendu aussi quasi-complet, la réciproque étant fausse.

Proposition 14.1.35 – Système quasi-complet à probabilités non nulles

Soit $(A_i)_{i\in I}$ un SQCE, et $J = \{i \in I,\ P(A_i) > 0\}$. Alors $(A_j)_{j\in J}$ est encore un SQCE.

II Conditionnement et indépendance

II.1 Probabilités conditionnelles

Les définitions relatives aux probabilités conditionnelles s'étendent bien au contexte étendu qu'on vient de définir.

Définition 14.2.1 – Probabilités conditionnelles

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé, A et B deux événements de $\mathcal{T}$ tels que B ne soit pas presque-impossible. Alors, la *probabilité conditionnelle de A en B* (ou la *probabilité de A sachant B*) est :
$$\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$
On trouve également souvent la notation $\mathbb{P}(A \mid B)$.

Cette définition se comprend bien pour la mesure uniforme : il s'agit dans ce cas de la proportion des éléments de B qui satisfont A (c'est-à-dire la proportion des éléments de B qui sont dans $A \cap B$).

Proposition 14.2.2

Soit B un événement non presque-impossible. Alors $\mathbb{P}_B$ définit une mesure de probabilités sur l'espace probabilisable $(\Omega, \mathcal{T})$.

◁ **Éléments de preuve.**

Vérifications sans difficulté. ▷

Remarque 14.2.3

La définition des probabilités conditionnelles est souvent utilisée sous la forme suivante : $\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\mathbb{P}(B)$.
En effet, expérimentalement, on a plus facilement accès aux probabilités conditionnelles qu'aux probabilités d'intersections, et on s'en sert pour le calcul de probabilités d'intersections. C'est d'ailleurs le cas initial de la formule plus générale des probabilités composées (théorème 14.2.6).

II.2 Formules liées aux probabilités conditionnelles

Théorème 14.2.4 – Formule des probabilités totales

Soit $(A_i)_{i \in I}$ un système quasi-complet fini ou dénombrable, tels que les événements A_i ne sont pas presque-impossibles, et soit $B \in \mathcal{T}$. Alors la famille $(\mathbb{P}(B \mid A_i)\mathbb{P}(A_i))_{i \in I}$ est sommable, et :

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B \mid A_i)\mathbb{P}(A_i).$$

◁ **Éléments de preuve.**

Écrire $B = \left(\biguplus_{i \in I} B \cap A_i \right) \uplus (B \cap C)$, où $\mathbb{P}(C) = 0$. ▷

Remarque 14.2.5

1. La formule des probabilités totales permet de retrouver une probabilité $\mathbb{P}(B)$ connaissant les probabilités conditionnelles $\mathbb{P}(B \mid A_i)$. Ainsi, on peut calculer $\mathbb{P}(B)$ en « distinguant suivant un certain nombre de cas » (suivant que A_i est réalisé), si ces cas partagent l'ensemble (ou presque) des cas possibles. Ainsi, cette formule est adapté au cas où le résultat d'une expérience dépend de résultats antérieurs, donc si on est amené à faire une discussion pour décrire le résultat de l'expérience.
2. Si un SCE ou un SQCE contient des événements quasi-impossibles, on a vu ci-dessus qu'en enlevant ces événements, on obtient encore un SQCE avec lequel la formule des probabilités totales est valide. Ainsi, plus généralement, si $(A_i)_{i \in I}$ est un SQCE quelconque,

$$\mathbb{P}(B) = \sum_{\substack{i \in I \\ \text{tq } P(A_i) \neq 0}} \mathbb{P}(B \mid A_i)\mathbb{P}(A_i).$$

Théorème 14.2.6 – Formule des probabilités composées

Soit $n \geqslant 2$, et $(A_1, \dots, A_n) \in \mathcal{T}^n$ tel que $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) \neq 0$ (i.e. $A_1 \cap \dots \cap A_{n-1}$ n'est pas presque-impossible). Alors :

$$\begin{aligned}\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) &= \mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1)\mathbb{P}(A_3 \mid A_1 \cap A_2)\cdots\mathbb{P}(A_n \mid A_1 \cap \cdots \cap A_{n-1}) \\ &= \mathbb{P}(A_1)\prod_{i=2}^{n}\mathbb{P}\left(A_i \;\middle|\; \bigcap_{j=1}^{i-1} A_j\right).\end{aligned}$$

◁ **Éléments de preuve.**

Récurrence sur n. ▷

Remarque 14.2.7

L'intérêt de cette formule est de donner une formule pour le calcul des probabilités d'intersections d'événements, notamment dans le cas d'une succession d'expérience. Cette formule dévoile toute son utilité lorsque les événements considérés ne sont pas mutuellement indépendants. En cas d'indépendance mutuelle, elle ne dit rien de plus que la proposition 14.2.15 ci-dessous.

Théorème 14.2.8 – Formule de Bayes simple

Soit A et B deux événements tels que $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$. Alors

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B)}$$

◁ **Éléments de preuve.**

Calculer $\mathbb{P}(A \cap B)$ en conditionnant de deux manières. ▷

De façon plus générale, la connaissance des probabilités d'un système complet, et des probabilités d'un événement B conditionnées au système complet permet de retourner une à une les probabilités conditionnelles :

Théorème 14.2.9 – Formule de Bayes sur un système complet

Soit $(A_i)_{i \in I}$ un système quasi-complet fini ou dénombrable, et B un événement non quasi-impossible. On suppose que pour tout $i \in I$, $\mathbb{P}(A_i) \neq 0$. Soit $j \in I$. Alors :

$$\mathbb{P}(A_j \mid B) = \frac{\mathbb{P}(B \mid A_j)\mathbb{P}(A_j)}{\sum\limits_{i \in I}\mathbb{P}(B \mid A_i)\mathbb{P}(A_i)}.$$

◁ **Éléments de preuve.**

Exprimer $\mathbb{P}(B)$ avec la formule des probabilités totales dans la formule précédente. ▷

Remarque 14.2.10 – Intérêt de la formule de Bayes

Le système (A_i) représente souvent une liste de causes pouvant amener l'événement B lors d'une étape suivante de l'expérience. Il est alors généralement facile de déterminer la probabilité qu'une certaine conséquence B ait lieu, sachant que la cause A_i a eu lieu, c'est-à-dire la probabilité conditionnelle $\mathbb{P}(B \mid A_i)$ (on respecte dans ce cas l'ordre temporel). Ces données permettent, grâce à la formule de Bayes, de remonter le temps, en déterminant la probabilité qu'une certaine cause A_i ait eu lieu sachant la conséquence B. Pour cette raison, cette formule est aussi souvent appelée *formule de probabilité des causes*.

Exemple 14.2.11

Dans une population composée d'autant d'hommes que de femmes, 5 % des hommes et 0,25 % des femmes sont daltoniens. Quelle est la probabilité pour qu'une personne daltonienne choisie au hasard soit une femme ?

II.3 Indépendance

Les probabilités conditionnelles mesurent l'impact de la réalisation d'un événement sur un autre. Si cet impact est nul (c'est-à-dire si la connaissance de la réalisation d'un événement n'influe pas sur la probabilité de l'autre), on dira que ces deux événements sont indépendants.

Ainsi, intuitivement, deux événements A et B sont indépendants si la connaissance de l'un ne modifie pas la probabilité de l'autre, donc si $\mathbb{P}(A \mid B) = \mathbb{P}(A)$ et $\mathbb{P}(B \mid A) = \mathbb{P}(B)$. L'inconvénient de cette définition réside dans le fait qu'elle n'est valable que sous une hypothèse de non quasi-impossibilité, afin de pouvoir considérer les probabilités conditionnelles. On constate cependant que dans le cas de non quasi-impossibilité, les deux égalités sont équivalentes à l'égalité : $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Cette égalité pouvant être considérée sans hypothèse de quasi-impossibilité, c'est elle que nous prenons comme définition de l'indépendance. Elle a en outre l'avantage d'être symétrique.

Définition 14.2.12 – Événements indépendants

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé, et $(A, B) \in \mathcal{T}^2$. Les événements A et B sont *indépendants* si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

En faisant le cheminement inverse de celui précédant la définition, il vient :

Proposition 14.2.13 – Probabilités conditionnelles et indépendance

Soit A et B deux événements indépendants tel que B ne soit pas quasi-impossible. Alors

$$\mathbb{P}(A) = \mathbb{P}(A \mid B).$$

On a évidemment aussi $\mathbb{P}(B) = \mathbb{P}(B \mid A)$ si A n'est pas quasi-impossible.
On note dans ce qui suit $\mathcal{P}_f(I)$ l'ensemble des parties **finies** de I.

Définition 14.2.14 – Indépendance d'une famille

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé, et $(A_i)_{i \in I}$ une famille d'événements. On dit que les A_i, $i \in I$, sont indépendants (ou mutuellement indépendants), si et seulement si pour tout $J \in \mathcal{P}_f(I)$,

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j).$$

En particulier si I lui-même est fini en prenant $J = I$:

Proposition 14.2.15 – Probabilité d'une intersection d'événements indépendants

Soit $(A_i)_{i \in I}$ une famille finie d'événements indépendants. Alors :

$$\mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}(A_i).$$

Avertissement 14.2.16

L'égalité précédente ne suffit pas à avoir l'indépendance. Il faut bien contrôler TOUS les produits des sous-familles (finies)

On verra dans le paragraphe suivant comment calculer la probabilité d'une intersection lorsque les événements ne sont pas indépendants (formule des probabilités composées, théorème 14.2.6).

Exemple 14.2.17

On fait des tirages à pile ou face

- A est réalisé si et seulement si le premier tirage est Pile.
- B est réalisé si et seulement si lors des 3 premiers tirages, il y a au plus un Pile.
- $C = B$

On a $\mathbb{P}(A \cap B \cap C) = \frac{1}{8} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$, mais les événements ne sont clairement pas indépendants !

De la définition même découle :

Proposition 14.2.18 – Stabilité de l'indépendance par restriction

Soit $(A_i)_{i\in I}$ une famille d'événements indépendants et $J \subset I$. Alors $(A_j)_{j\in J}$ est encore composée d'événements indépendants.

En particulier, en prenant toutes les sous-familles de 2 événements, l'indépendance implique l'indépendance 2 à 2 des événements.

Avertissement 14.2.19

La réciproque est fausse : l'indépendance 2 à 2 d'une famille n'implique pas l'indépendance (mutuelle).

Exemple 14.2.20

On procède à trois tirages à Pile ou Face. Soit A l'événement consistant à obtenir exactement 1 Pile lors des tirages 2 et 3, B de même lors des tirages 1 et 3, C de même lors des tirages 1 et 2. Les événements A, B et C sont 2 à 2 indépendants, mais pas mutuellement.

Proposition 14.2.21 – Indépendance des intersections de sous-familles disjointes, HP

Soit $(A_i)_{i\in I}$ une famille d'événements indépendants, et $(I_k)_{k\in K}$ une famille de sous-ensembles finis de I, 2 à 2 disjoints. Alors la famille $\left(\bigcap_{i\in I_k} A_i\right)_{k\in K}$ est constitué d'événements indépendants.

◁ **Éléments de preuve.**

Prendre une intersection finie de tels ensembles, c'est aussi une intersection finie de A_i. Écrire par indépendance la probabilité comme produit de $\mathbb{P}(A_i)$, puis regrouper par paquets de I_k en réutilisant l'indépendance sur ces paquets. ▷

Proposition 14.2.22 – Indépendance et complémentation

Si A et B sont indépendants, alors A et $\overline{B}$ sont indépendants.

◁ **Éléments de preuve.**

Calculer $\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A \backslash (A \cap B))$. ▷

Proposition 14.2.23 – Indépendance et complémentation pour une famille finie

Si $A_1, \dots, A_n$ sont indépendants, il en est de même de $\overline{A_1}, A_2, \dots, A_n$

◁ **Éléments de preuve.**

Considérer l'intersection d'une sous-famille. Seul cas à voir : $\overline{A_1}$ est dans la sous-famille. Se ramener au cas précédent en justifiant que A_1 est indépendant de l'intersection des autres termes de la sous-famille. ▷

Corollaire 14.2.24 – Indépendance et complémentation pour une famille quelconque

Soit $(A_i)_{i\in I}$ une famille d'événements indépendants et $(B_i)_{i\in I}$ une famille telle que pour tout $i \in I$, $B_i = A_i$ ou $\overline{A_i}$. Alors $(B_i)_{i\in I}$ est une famille d'événements indépendants.

III Bilan des méthodes calculatoires en probabilités

Nous exposons ici le principe général du calcul (et de la rédaction de ce calcul) d'une probabilité. Évidemment, dans la pratique, les situations sont très variées, et on peut être amené à s'écarter légèrement de la voie tracée ci-dessous.

Méthode 14.3.1 – Comment aborder un calcul de probabilité

1. Si on est dans une situation d'équiprobabilité, on peut se diriger vers un argument combinatoire (et utiliser la formule de Laplace). Dans les autres cas, la démarche générale est indiquée ci-dessous.
2. Définir des événements simples issus de l'expérience décrite, en n'oubliant pas la numérotation de ces événements en cas d'expérience répétée. Ces événements simples auront pour vocation de décrire ceux qui nous intéressent.

   Il ne faut pas hésiter à introduire des événements par vous-même, à donner des notations. Ce travail n'est pas nécessairement fait pour vous dans le sujet.
3. Décrire à l'aide d'une **phrase** l'événement dont on recherche la probabilité, sous forme d'une condition nécessaire et suffisante de réalisation :

   « L'événement A est réalisé si et seulement si ... »

   Cette condition nécessaire et suffisante de réalisation de l'événement A doit s'exprimer en fonction de la réalisation ou non des événements simples définis dans la première étape. Il s'agit donc de « décomposer » l'événement A en événements simples.
4. Traduire cette description dans le langage ensembliste (c'est-à-dire sous forme d'union, intersection ou complémentaires d'événements élémentaires). Les deux étapes sont nécessaires : la première (description verbale) fournit une explication facilement lisible, et éclaire la formule, la seconde (description ensembliste) donne de la rigueur à l'argument, et permet ensuite de s'orienter vers la bonne méthode de calcul.
5. Calcul de la probabilité $\mathbb{P}(A)$, en se servant de la décomposition ensembliste précédente, et des règles de calcul des probabilités d'unions, d'intersections ou de complémentaires. Nous rappelons l'essentiel de ces règles ci-dessous.

Ainsi, il est important de savoir s'orienter rapidement vers la bonne technique de calcul (probabilité d'une union, d'une intersection), suivant la situation rencontrée. Voici les différents cas rencontrés les plus fréquemment :

Méthode 14.3.2 – Calcul de la probabilité d'un complémentaire

Une seule formule à connaître, issue de la définition d'une mesure de probabilité :

$$\forall A \in \mathcal{T}, \quad \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A).$$

Méthode 14.3.3 – Calcul de la probabilité d'une intersection

Suivant les cas :

- Intersection dénombrable d'événements décroissants pour l'inclusion :

  On utilise le théorème de la limite monotone.
- Intersection d'un nombre fini d'événements mutuellement indépendants :

  On utilise la relation issue de la définition de l'indépendance mutuelle :

  $$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2)\ldots\mathbb{P}(A_n).$$
- Intersection d'un nombre fini d'événements non nécessairement indépendants :

  On utilise la formule des probabilités composées.

  Cette formule est surtout utile lorsqu'il y a un enchainement temporel des événements (A_1 étant le premier à advenir, A_n le dernier), chaque résultat influant sur les suivants. C'est le cas par exemple dans le cadre de tirages successifs dans une urne évolutive, l'évolution se faisant différemment suivant la boule tirée à un tirage donné.

  Cette formule permet de formaliser le calcul de la probabilité associée à une branche de l'arbre des possibilités.

Méthode 14.3.4 – Calcul de la probabilité d'une union

- Union dénombrable d'événements croissants pour l'inclusion :

  On utilise ici aussi le théorème de la limite monotone.
- Union d'un nombre fini ou dénombrable d'événements 2 à 2 incompatibles :

  On utilise l'additivité (cas fini) ou la σ-additivité (cas dénombrable) de la mesure de probabilité.
- Union d'un nombre fini d'événements mutuellement indépendants :

  On a intérêt dans ce cas à considérer l'événement complémentaire. On est alors ramené au calcul de la probabilité d'une intersection d'événements mutuellement indépendants.
- Union d'un nombre fini d'événements lorsqu'on a des informations sur les intersections :

  On utilise la formule du crible de Poincaré. Les cas $n = 2$ et $n = 3$ sont les plus utiles :
  * $n = 2$: $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$
  * $n = 3$: $\mathbb{P}(A_1 \cup A_2 \cup A_3) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) - \mathbb{P}(A_2 \cap A_3) - \mathbb{P}(A_1 \cap A_3) - \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_1 \cap A_2 \cap A_3)$

Méthode 14.3.5 – Cas d'une expérience dont l'issue dépend de résultats antérieurs

Dans certaines situations, par exemple lorsque le mode opératoire d'une expérience (par exemple la composition d'une urne) dépend du résultat d'une expérience précédente, il peut être nécessaire de discuter suivant le résultat obtenu lors de l'expérience antérieure : autrement dit, dans ces situations, il est aisé de calculer des probabilités conditionnellement à chacun des résultats possibles de la première expérience. Il faut ensuite récupérer la probabilité globale à partir de toutes ces probabilités conditionnelle. On utilise pour cela la formule des probabilités totales.

A retenir : $\boxed{\text{discussion à faire} \longrightarrow \text{formule des probabilités totales}}$

Enfin, vu qu'elles interviennent dans de nombreuses formules, nous faisons quelques remarques concernant le calcul des probabilités conditionnelles.

Méthode 14.3.6 – Calcul de probabilités conditionnelles

Puisque la probabilité conditionnelle $\mathbb{P}_B$ sachant un événement B donné est une mesure de probabilité, tous les résultats, et toutes les règles vues ci-dessus pour le calcul des probabilités s'appliquent au calcul des probabilités conditionnelles. Lorsque ces formules font elles-même intervenir des probabilités conditionnelles, les conditions s'ajoutent à celle déjà présente (sous forme d'une intersection) :

$$(\mathbb{P}_B)_C = \mathbb{P}_{B\cap C}.$$

À titre d'exemple, voici la formule des probabilités totales pour une probabilité conditionnelle : $(A_i)_{i\in I}$ étant un système complet au plus dénombrable tel que pour tout $i \in I$ $\mathbb{P}_C(A_i) \neq 0$, on a :

$$\mathbb{P}(B \mid C) = \sum_{i\in I} \mathbb{P}(A_i \mid C)\mathbb{P}(B \mid C \cap A_i).$$

Avertissement 14.3.7

Attention, on ne peut pas définir de notion d'événement conditionnel. Ce n'est pas l'événement qui est conditionnel, mais sa réalisation. Ainsi, si vous adoptez la démarche générale de rédaction :

- ne décrivez pas l'événement conditionnel (cela n'a pas de sens), mais donnez bien une condition nécessaire et suffisante de réalisation, conditionnellement à un certain événement :
  « Sachant que B est réalisé, A est réalisé si et seulement si ... »
- Gardez-vous bien de transcrire cette phrase de manière ensembliste ; cela vous amènerait inévitablement à considérer des « événements conditionnels ». Sautez cette étape et passez directement aux probabilités. Vous pouvez transcrire votre phrase de façon ensembliste à l'intérieur des probabilités, la condition étant alors précisée non pas au niveau des ensembles, mais au niveau de la probabilité.

Avertissement 14.3.8

L'indépendance est une notion dépendant de la mesure de probabilité. Ce n'est pas parce que A et B sont indépendants (lorsqu'on dit cela, on sous-entend l'indépendance pour la mesure de probabilité totale $\mathbb{P}$) qu'ils sont indépendants pour une mesure de probabilité conditionnelle $\mathbb{P}_C$.

Exemple 14.3.9

On effectue 3 tirages à Pile ou Face indépendants. L'événement C est réalisé si et seulement si on obtient exactement 1 Pile lors des tirages 1 et 2, A de même avec les tirages 2 et 3. et B de même avec les tirages 1 et 3.
Les événements A et B sont indépendants pour $\mathbb{P}$, mais pas pour $\mathbb{P}_C$.

“

15

Variables aléatoires discrètes

Dans ce chapitre, sauf mention contraire, $(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé quelconque.

La situation à laquelle vous serez le plus fréquemment confrontés est celle d'une variable aléatoire discrète à valeurs réelles (variable aléatoire réelle), ou à valeurs dans $\mathbb{R}^n$ (on parle de vecteur aléatoire, c'est équivalent à la donnée de n variables aléatoires réelles) mais on peut définir une variable aléatoire discrète à valeurs dans n'importe quel ensemble E.
La définition d'une variable aléatoire X générale (non discrète) dépasse en revanche le cadre du programme, et nécessite un cadre un peu plus précis : l'ensemble E dans lequel X prend ses valeurs doit lui-même être muni d'une tribu, et X doit vérifier une propriété de « mesurabilité », qui est en quelque sorte un respect des tribus, comme la continuité est le respect des topologies.

I Loi d'une variable aléatoire discrète

I.1 Généralités sur les variables aléatoires discrètes

Définition 15.1.1 – Variable aléatoire discrète

Soit $(\Omega, \mathcal{T})$ un espace probabilisable. Une *variable aléatoire discrète* X à valeurs dans E est une application $X : \Omega \to E$ telle que :

(i) $X(\Omega)$ (l'ensemble des valeurs possibles de X) est au plus dénombrable ;

(ii) pour tout $xinX(\Omega)$, $X^{-1}(\{x\}) \in \mathcal{T}$.

Définition 15.1.2 – Variables et vecteurs aléatoires réels discrets

- si $E = \mathbb{R}$ et X vérifie les conditions de la définition 15.1.1, on dit que X est une variable aléatoire réelle discrète (abrégé en v.a.r.d.)
- si $E = \mathbb{R}^n$ et X vérifie les conditions de la définition 15.1.1, on dit que X est un vecteur aléatoir réel discret.

Proposition 15.1.3 – Image réciproque des parties de E

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, et $X : \Omega \to E$ une variable aléatoire discrète. Alors,

$$\forall A \in \mathcal{P}(E), \quad X^{-1}(A) \in \mathcal{T}.$$

◁ **Éléments de preuve.**
C'est la stabilité par union finie ou dénombrable, du fait que $X(\Omega) \cap A$ est fini ou dénombrable. ▷

Notation 15.1.4 – Événements liés à une variable aléatoire

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, et X une variable aléatoire à valeurs dans E.

- Soit $A \in \mathcal{P}(E)$. L'événement $X^{-1}(A) \in \mathcal{T}$ est noté $(X \in A)$ ou $\{X \in A\}$ ou encore $[X \in A]$.
- Soit $x \in A$. L'événement $X^{-1}(\{x\}) = (X \in \{x\})$ sera simplement noté $(X = x)$, ou $\{X = x\}$, ou $[X = x]$.

Remarque 15.1.5

Les notations $(X \in A)$ et $\{X \in A\}$ figurent toutes deux dans le programme, contrairement à $[X \in A]$, qui est néanmoins largement présente dans beaucoup d'ouvrages (en particulier pour les événements $[X = x]$).
Nous laisserons désormais de côté cette notation.

Notation 15.1.6 – Événements liés à une v.a.r.d.

Dans le cas d'une variable aléatoire réelle discrète,

- l'événement $\{\omega \mid X(\omega) < a\}$ est noté $(X < a)$ ou $\{X < a\}$.
- l'événement $\{\omega \mid X(\omega) \leqslant a\}$ est noté $(X \leqslant a)$ ou $\{X \leqslant a\}$;
- l'événement $\{\omega \mid a < X(\omega) < b\}$ est noté $(a < X < b)$ ou $\{a < X < b\}$;
- l'événement $\{\omega \mid a < X(\omega) \leqslant b\}$ est noté $(a < X \leqslant b)$ ou $\{a < X \leqslant b\}$
- etc.

Proposition 15.1.7 – Stabilité par produit cartésien

Soit $X : \Omega \to E$ et $Y : \Omega \to F$ deux variables aléatoires discrètes. Alors le couple (X, Y) est une variable aléatoire discrète.

Proposition 15.1.8 – Composition d'une variable aléatoire discrète

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{T})$, à valeurs dans E, et f une application de $X(\Omega)$ dans un ensemble F. Alors $f \circ X$ est une variable aléatoire discrète, notée $f(X)$.

◁ **Éléments de preuve.**

- $f(X)^{-1}(\{y\})$ est une union au plus dénombrable de $X^{-1}(\{x\})$.
- $f : X(\Omega) \to f(X)(\Omega)$ est une surjection, et $X(\Omega)$ est au plus dénombrable.

▷

On en déduit :

Proposition 15.1.9 – Algèbre des v.a.r.d.

Le sous-ensemble de $\mathbb{R}^\Omega$ constitué des variables aléatoires réelles discrètes est une sous-algèbre de $\mathbb{R}^\Omega$. Ainsi, si X et Y deux v.a.r.d. sur Ω, λ un réel quelconque :

1. 1 est une v.a.r.d.
2. $X + Y$ est une v.a.r.d.
3. λX est une v.a.r.d.
4. XY est une v.a.r.d.

De plus, si X est une v.a.r.d., il en est de même de $|X|$.

◁ **Éléments de preuve.**

2 et 4 s'obtiennent en composant (X, Y) par la somme ou le produit. ▷

I.2 Loi d'une variable aléatoire discrète

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{T}, \mathbb{P})$, à valeurs dans E. par définition, $\{X = x\}$ est un événement, donc $\mathbb{P}(X = x)$ est donc bien défini.

Proposition 15.1.10

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{T}, \mathbb{P})$ à valeurs dans E. Alors

- la famille $(\{X = x\})_{x \in X(\Omega)}$ est un SCE d'événements ;
- la famille $(\mathbb{P}(X = x))_{x \in X(\Omega)}$ est une distribution de probabilités.

◁ **Éléments de preuve.**

Le premier point est de la vérification facile (c'est en fait la partition associée à une surjection

Le deuxième point et la σ-additivité utilisée sur le SCE précédent. ▷

Définition 15.1.11 – Loi d'une variable aléatoire discrète

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé, et $X : \Omega \to E$ une variable aléatoire.
La loi de X est la mesure de probabilité définie sur $(X(\Omega), \mathcal{P}(X(\Omega)))$ par la distribution de probabilités discrète $(P(X = x))_{x \in X(\Omega)}$.

Remarque 15.1.12

Au besoin, il est possible de définit $\mathbb{P}$ sur l'espace probabilisable $(F, \mathcal{P}(F))$, pour tout ensemble F tel que $X(\Omega) \subset F$, en prolongeant la mesure de probabilité $(P(X = x))_{x \in X(\Omega)}$, en posant

$$p_x = \begin{cases} \mathbb{P}(X = x) & \text{si } x \in X(\Omega) \\ 0 & \text{sinon.} \end{cases}$$

Cela permet par exemple de définir $\mathbb{P}_X$ sur $(E, \mathcal{P}(E))$ tout entier.

Proposition 15.1.13 – Expression de la loi d'une v.a. discrète

Soit $A \in \mathcal{P}(X(\Omega))$,

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A) = \sum_{x \in A} \mathbb{P}(X = x),$$

cette somme étant bien définie.

◁ **Éléments de preuve.**

Écrire A comme union de singletons. ▷

Remarque 15.1.14

En définissant $\mathbb{P}_X$ sur $(E, \mathcal{P}(E))$, on peut plus généralement écrire

$$\forall A \in \mathcal{P}(E), \quad \mathbb{P}_X(A) = \mathbb{P}(X \in A) = \sum_{x \in A} \mathbb{P}(X = x) = \sum_{x \in A \cap X(\Omega)} \mathbb{P}(X = x),$$

Notation 15.1.15 – Variables de même loi

On note $X \sim Y$ pour désigner le fait que X et Y suivent la même loi.
Plu précisément, étant données deux variables aléatoires $X : \Omega_1 \to E$ et $Y : \Omega_2 \to E$, définies sur les espaces probabilisés $(\Omega_1, \mathcal{T}_1, \mathbb{P}_1)$ et $(\Omega_2, \mathcal{T}_2, \mathbb{P}_2)$, $X \sim Y$ si et seulement si les mesures $\mathbb{P}_X$ et $\mathbb{P}_Y$ définies sur E vérifient

$$\mathbb{P}_X = \mathbb{P}_Y$$

Ainsi, $X \sim Y$ si et seulement si les supports S_1 et S_2 des distributions de probabilités $(\mathbb{P}_1(X = x))_{x \in X(\Omega_1)}$ et $(\mathbb{P}_2(Y = y))_{y \in Y(\Omega_2)}$ sont les mêmes, et si

$$\forall x \in S_1 = S_2, \quad \mathbb{P}_1(X = x) = \mathbb{P}_2(Y = x).$$

Exemple 15.1.16

Si $X = Y$ presque sûrement (i.e. il existe $\Omega' \subset \Omega$ tel que $\mathbb{P}(\Omega') = 1$ et pour tout $\omega \in \Omega'$, $X(\omega) = Y(\omega)$), alors $X \sim Y$.

Remarque 15.1.17

1. Contrairement à l'exemple précédent, la relation $X \sim Y$ ne nécessite pas que X et Y soient définis sur le même espace probabilisé.
2. La relation $\sim$ est clairement réflexive, symétrique et transitive. Mais dans le cadre du programme, il est un peu abusif d'en déduire que c'est une relation d'équivalence, car son « domaine » de définition n'est pas très bien défini. Il n'est pas clair que l' « ensemble » des variables aléatoires discrètes soit bien un ensemble, car, même en fixant E, la relation nécessite de faire varier Ω. Or, la collection des ensembles n'est pas un ensemble.

On peut bien évidemment aussi définir des lois conditionnelles. En effet, la définition d'une variable aléatoire $X : \Omega \to E$ n'implique que l'espace probabilisable $(\Omega, \mathcal{T})$, et non l'espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$. Ainsi, étant donné un événement $B \in \mathcal{T}$, on peut munir $(\Omega, \mathcal{T})$ de la mesure de probabilité conditionnelle $\mathbb{P}_B$ et définir la loi de la variable X en référence à l'espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P}_B)$. Concrètement, la loi conditionnelle de X est donc définie par la distribution de probabilités

$$\forall x \in X(\Omega), \quad \mathbb{P}_{B,X}(x) = \mathbb{P}_B(X = x).$$

I.3 Loi de $f(X)$

Proposition 15.1.18 – Loi de $f(X)$

1. Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{T}, \mathbb{P})$, à valeurs dans E, et $f : E \to F$ La loi de $f(X)$, définie sur $(F, \mathcal{P}(F))$ (ou seulement sur $(f(X(\Omega)), \mathcal{P}(f(X(\Omega))))$, est donnée par

   $$\mathbb{P}_{f(X)} = \mathbb{P}_X \circ \widehat{f^{-1}}.$$

   Autrement dit, pour tout $A \in \mathcal{P}(F)$,

   $$\mathbb{P}_{f(X)}(A) = \mathbb{P}_X(f^{-1}(A)) = \mathbb{P}(X \in f^{-1}(A)).$$

2. En particulier,

   $$\forall x \in f(X(\Omega)), \quad \mathbb{P}(f(X) = x) = \sum_{\substack{y \mid f(y) = x \\ y \in X(\Omega)}} \mathbb{P}(X = y).$$

◁ **Éléments de preuve.**

1. $\{X \in f^{-1}(A)\} = \{f(X) \in A\}$.
2. $\{f(X) = x\} = \{X \in f^{-1}(x)\}$.

▷

Exemples 15.1.19 – Loi d'une somme, loi du minimum

Voici deux exemples particulièrement important, à bien connaître, dans le cas où X et Y sont des variables aléatoires réelles discrètes.

1. $\forall z \in \mathbb{R}, \quad \mathbb{P}(X+Y=z) = \displaystyle\sum_{\substack{(x,y)\in X(\Omega)\times Y(\Omega) \\ x+y=z}} \mathbb{P}(X=x, Y=y).$
2. $\forall z \in X(\Omega) \cup Y(\Omega), \quad \mathbb{P}(\min(X,Y)=z) = \displaystyle\sum_{\substack{y\in Y(\Omega) \\ y\geqslant z}} \mathbb{P}(X=z, Y=y) + \sum_{\substack{x\in X(\Omega) \\ x>z}} \mathbb{P}(X=x, Y=z).$

   Exemple : tirage de deux dés.

Corollaire 15.1.20 – Stabilité de $\sim$ par image directe

Soit $X : \Omega_1 \to E$ et $Y : \Omega_2 \to E$ telles que $X \sim Y$, et $f : E \to F$. Alors $f(X) \sim f(Y)$.

I.4 Loi d'un vecteur aléatoire réel discret

Proposition 15.1.21 – Caractérisation d'un vecteur aléatoire par ses coordonnées

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé, et $X : \Omega \to \mathbb{R}^n$ une application. On note $X = (X_1, \dots, X_n)$ les coordonnées de X. Les propositions suivantes sont équivalentes:

(i) X est une variable aléatoire discrète (donc un vecteur aléatoire discret)

(ii) pour tout $i \in [\![1, n]\!]$, X_i est une variable aléatoire réelle discrète.

Notation 15.1.22 – $\mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n)$

Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire discret, à valeurs dans $\mathbb{R}^n$. Pour alléger les notations, on note $\mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n)$ au lieu de $\mathbb{P}([\{X_1 \in A_1\} \cap \cdots \cap \{X_n \in A\})$. En particulier, on utilisera la notation $\mathbb{P}(X_1 = x_1, \dots, X_n = x_n)$.

Pour éviter l'abus de notations techniques stériles, nous donnons les définitions dans le cas de couples de vriables aléatoires, elles s'étendent facilement au cas de vecteurs aléatoires à davantage de coordonnées. Pour commencer, on peut remarquer que si (X,Y) est une couple de variables aléatoires réelles discrètes, et $V = (X,Y)$, alors

$$V(\Omega) \subset X(\Omega) \times Y(\Omega).$$

Définition 15.1.23 – Loi conjointe, lois marginales

Soit (X,Y) un couple de variables aléatoires réelles discrètes.

- La loi conjointe du couple (X,Y) est la loi du vecteur aléatoire $V = (X,Y)$. Il s'agit donc d'une mesure de probabilité $\mathbb{P}_{(X,Y)}$ définie par la distribution de probabilités

$$(\mathbb{P}(X=x, Y=y))_{(x,y)\in X(\Omega)\times Y(\Omega)}.$$

- La première loi marginale du couple (X,Y) est la loi de X
- La seconde loi marginale du couple (X,Y) est la loi de Y

De la même façon, on parle de la loi conjointe, et de la k-ième loi marginale d'un n-uplet $(X_1, \dots, X_n)$. De façon évidente :

Proposition 15.1.24 – Expression d'une loi marginale

Soit (X,Y) un couple de variables aléatoires réelles discrètes. La loi conjointe détermine les lois marginales par les formules :

$$\begin{cases} \forall x \in X(\Omega), \quad \mathbb{P}(X = x) = \displaystyle\sum_{y \in Y(\Omega)} \mathbb{P}(X = x, Y = y). \\ \forall y \in X(\Omega), \quad \mathbb{P}(Y = y) = \displaystyle\sum_{x \in X(\Omega)} \mathbb{P}(X = x, Y = y). \end{cases}$$

De façon similaire, la k-ième loi marginale d'un n-uplet $(X_1, \dots, X_n)$ s'obtient en sommant sur toutes les valeurs possibles prises par les $n-1$ autres variables aléatoires.

Avertissement 15.1.25

Les lois marginales sont déterminées par la loi conjointe, mais la réciproque est fausse : les lois marginales ne déterminent pas la loi conjointe !

Exemple 15.1.26

Soit une urne contenant 1 boule blanche, 1 boule noire. On effectue un tirage, avec équiprobabilité. On note X_1, Y_1 et X_2 les trois variables (égales l'une à l'autre) égales à 1 si on tire la boule noire, et 0 sinon. On note Y_2 la variable égale à 1 si on tire la boule blanche, et 0 sinon. Alors, les lois marginales de (X_1, Y_1) et de (X_2, Y_2) sont les mêmes, pourtant les lois conjointes sont distinctes.

I.5 Variables indépendantes

Proposition/Définition 15.1.27 – Variables aléatoires discrètes indépendantes

Soit X et Y deux variables aléatoires discrètes définies sur $(\Omega, \mathcal{T}, \mathbb{P})$, à valeurs dans E_1 et E_2 respectivement. Les propositions suivantes sont équivalentes:

(i) Pour tout $(A, B) \in \mathcal{P}(E_1) \times \mathcal{P}(E_2)$, $\mathbb{P}_{(X,Y)}(A \times B) = \mathbb{P}_X(A)\mathbb{P}_Y(B)$;

(ii) Pour tout $(A, B) \in \mathcal{P}(E_1) \times \mathcal{P}(E_2)$, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants, *i.e.*

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B);$$

(iii) Pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, les événements $\{X = x\}$ et $\{Y = y\}$ sont indépendants, i.e.

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y);$$

Dans ce cas, on dit que X et Y sont indépendances, et on note $X \perp\!\!\!\perp Y$.

Remarque 15.1.28

Conformément à la remarque faite lors de la définition de la loi d'une variable aléatoire, on peut se contenter de considérer $(A, B) \in \mathcal{P}(X(\Omega)) \times \mathcal{P}(Y(\Omega))$.s

Proposition/Définition 15.1.29 – Famille de variables mutuellement indépendantes

Soit $(X_k)_{k \in K}$ une famille quelconque de variables aléatoires discrètes sur $(\Omega, \mathcal{T}, \mathbb{P})$, X_k étant à valeurs dans E_k. Les propositions suivantes sont équivalentes:

(i) pour toute famille $(A_k)_{k \in K}$ telle que pour tout $k \in K$, $A_k \subset E_k$, les événements $[X_k \in A_k]$, $k \in K$, sont indépendants ;

(ii) pour tout $J \in \mathcal{P}_f(K)$ et toute famille $(A_j)_{j \in J}$ telle que pour tout $j \in J$, $A_j \subset E_j$,

$$\mathbb{P}\left(\bigcap_{j \in J} \{X_j \in A_j\}\right) = \prod_{j \in J} \mathbb{P}(X_j \in A_j);$$

(iii) pour toute famille $(x_k)_{k\in K}$ telle que pour tout $k \in K$, $x_k \in X_k(\Omega)$, les événements $\{X_k = x_k\}$, $k \in K$, sont indépendants ;

(iv) pour tout $J \in \mathcal{P}_f(K)$ et toute famille $(x_j)_{j\in J}$ telle que pour tout $j \in J$, $x_j \in X_j(\Omega)$,

$$\mathbb{P}\left(\bigcap_{j\in J}\{X_j = x_j\}\right) = \prod_{j\in J}\mathbb{P}(\{X_j = x_j\}).$$

Dans ce cas, on dit que les variables aléatoires X_k, $k \in \mathbb{K}$, sont indépendantes (ou mutuellement indépendantes).

Proposition 15.1.30 – Sous-famille d'une famille de variables indépendantes

Si la famille $(X_k)_{k\in K}$ est une famille de variables aléatoires discrètes indépendantes, alors, pour tout $J \subset K$, $(X_j)_{j\in K}$ est une famille de variables aléatoires discrètes indépendantes.

Corollaire 15.1.31 – Indépendance 2 à 2

Si $(X_k)_{k\in K}$ est une famille de variables alétoires discrètes indépendantes, alors les X_k sont 2 à 2 indépendants. La réciproque est fausse.

En revanche, on peut caractériser l'indépendance par l'indépendance des sous-familles finies.

Proposition 15.1.32 – Caractérisation de l'indépendance par les familles finies

Soit $(X_k)_{k\in K}$ une famille de variables aléatoires discrètes indépendantes. Les propositions suivantes sont équivalentes:

(i) $(X_k)_{k\in K}$ est une famille de variables indépendantes ;

(ii) pour tout $J \in \mathcal{P}_f(K)$, $(X_j)_{j\in J}$ est une famille de variables indépendantes ;

Proposition 15.1.33 – Fonctions de variables indépendantes

Soit $X \perp\!\!\!\perp Y$ deux v.a.r.d. independantes, et f et g deux fonctions définies respectivement sur $X(\Omega)$ et $Y(\Omega)$. Alors $f(X) \perp\!\!\!\perp g(Y)$.

Ceci est un cas particulier du résultat plus général ci-dessous :

Théorème 15.1.34 – Lemme des coalitions, pour 2 coalitions

Soit $1 \leqslant m < n$, et $(X_1, \dots, X_m, \dots X_n)$ une famille de variables aléatoires discrètes indépendantes. Soit f et g deux fonctions définies sur $X_1(\Omega) \times \dots X_m(\Omega)$ et $X_{m+1}(\Omega) \times \cdots \times X_n(\Omega)$ respectivement. Alors $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ sont indépendantes.

◁ **Éléments de preuve.**

Démonstration un peu technique, en exprimant tout en fonction des répartitions de probabilités des X_i. ▷

Par récurrence, on en déduit la version générale, qu'on exprime volontairement sous une forme un peu lâche.

Théorème 15.1.35 – Lemme des coalitions

Soit $(X_k)_{k\in K}$ une famille de variables aléatoires discrètes indépendantes, et $(K_i)_{i\in I}$ une famille de sous-ensembles finis deux à deux disjoints de K. Soit $(Y_i)_{i\in I}$ une famille de variables aléatoires, telle que pour tout $i \in I$, Y_i s'écrive comme fonction des Y_k, $k \in K_i$ (et pas des autres). Alors $(Y_i)_{i\in I}$ est une famille de variables aléatoires discrètes indépendantes.

◁ **Éléments de preuve.**

Le cas K infini et I fini s'obtient par récurrence à partir du lemme des 2 coalitions (seul un nombre fini de X_k intervenant, puisque les parts K_i sont finies)

Le cas I infini provient alors du fait que l'indépendance d'une famille équivaut à l'aindépendance de toutes ses sous-familles finies. ▷

On admet le résultat suivant, qui permet d'assurer l'existence d'espaces probabilisés modélisant des situations un peu complexes, notamment infinies. Par exemple, la modélisation précise d'une succession infinie de tirages à Pile ou Face indépendants n'est pas chose facile. L'univers Ω est naturellement $\{0,1\}^{\mathbb{N}}$, mais cet ensemble est en bijection avec $\mathbb{R}$, et le munir de la tribu $\mathcal{P}(\Omega)$ ne permet pas ici de modéliser correctement le tirage à pile ou face.
En effet, on peut montrer que toute mesure de probabilité sur $(\mathbb{R}, \mathcal{P}(\mathbb{R}), \mathbb{P})$ est déterminée par une répartition de probabilités discrète. Or, pour tout $\omega \in \Omega$, on obtient alors dans cet espace probabilisé $\mathbb{P}(\omega) = 0$ (par le même calcul que la probabilité de n'obtenir que des Piles). Ceci contredit le fait que $\mathbb{P}$ soit déterminée par une répartition de probabilités discrète.
La construction d'un modèle convenable est difficile, et n'est pas un objectif du programme. Nous utiliserons donc le théorème suivant, admis, pour justifier l'existence d'un modèle convenable.

Théorème 15.1.36 – Espace probabilisé modélisant une suite de lois, admis

Soit E un ensemble et $(P_n)_{n\in\mathbb{N}}$ une suite de lois de probabilités discrètes sur E (ce qui équivaut à la donnée d'une suite de distributions de probabilités discrètes).
Alors il existe un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et une suite $(X_n)_{n\in\mathbb{N}}$ de variables indépendantes, définies sur $(\Omega, \mathcal{T}, \mathbb{P})$, à valeurs dans E, et telle que

$$\forall n \in \mathbb{N}, \quad \mathbb{P}_{X_n} = P_n.$$

◁ **Éléments de preuve.**

Démonstration hors-programme. ▷

Il est fréquent, pour étudier les caractéristiques d'une expérience, de la répéter une infinité de fois dans les mêmes conditions, de façon indépendante, et d'observer le résultat sous forme de valeur prise par une variable aléatoire.

Définition 15.1.37 – variables i.i.d.

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires. On dit que la suite (X_n) est indépendante identiquement distribuée (en abrégé officiel i.i.d.), si les X_n sont indépendants, et suivent la même loi (i.e. pour tout $(i,j) \in \mathbb{N}^2$, $X_i \sim X_j$).

Si les X_i suivent une loi classique, on dira, par exemple, que $(X_n)_{n\in\mathbb{N}}$ est une suite i.i.d. de variables suivant une loi binomiale $\mathcal{B}(n,p)$.

Corollaire 15.1.38 – Modélisation d'une répétition infinie d'expériences indépendantes

Soit P une loi de probabilités discrète sur E. Alors il existe un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et $(X_n)_{n\in\mathbb{N}}$ une suite i.i.d. de variables de loi P.

En appliquant cela à une loi de Bernoulli $\mathcal{B}(p)$, on obtient l'existence d'un espace probabilisé modélisant une succession infinie de tirages (éventuellement déséquilibrés) à Pile ou Face

Corollaire 15.1.39 – Espace probabilisé modélisant des tirages à P/F

Il existe un espace probabilisé modélisant une succession infinie de tirages à Pile ou Face indépendants, avec une probabilité p d'obtenir Pile à chaque tirage.

II Espérance et variance

On étend les définitions vues en première année au cas des variables discrètes.
Dans toute cette section,

- $(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé qu'on ne réintroduira pas à chaque énoncé ;
- sauf mention explicite du contraire, les variables sont définies sur $(\Omega, \mathcal{T}, \mathbb{P})$, et à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$.

II.1 Espérance d'une variable positive

Définition 15.2.1 – Espérance d'une variable positive

Soit X une variable aléatoire à valeurs dans $\overline{\mathbb{R}}_+$. L'espérance de X est définie par

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x),$$

où, par convention, $+\infty \times 0 = 0$. L'espérance est toujours bien définie dans $\overline{\mathbb{R}}_+$

Remarque 15.2.2

Si $+\infty \in X(\Omega)$:

- si $\mathbb{P}(X = +\infty) > 0$, alors $\mathbb{E}(X) = 0$;
- si $\mathbb{P}(X = +\infty) = 0$, alors X a même espérance que sa restriction à $\Omega' = \Omega \backslash \{X = +\infty\}$

Plus généralement,

Proposition 15.2.3 – Espérance de variables de même loi

Si X et Y sont des variables à valeurs dans $\overline{\mathbb{R}}_+$, et $X \sim Y$, alors $\mathbb{E}(X) = \mathbb{E}(Y)$.

Remarque 15.2.4

Une variable aléatoire X positive peut admettre une espérance infinie même si X ne prend que des valeurs finies.

Exemple 15.2.5

$X(\Omega) = \mathbb{N}^*$ et $\forall n \in \mathbb{N}^*$, $\mathbb{P}(X = n) = \dfrac{6}{\pi^2 n^2}$

Proposition 15.2.6 – Positivité et stricte positivité

Soit X une variable aléatoire à valeurs dans $\overline{\mathbb{R}}_+$.

1. $\mathbb{E}(X) \geqslant 0$;
2. $\mathbb{E}(X) = 0 \iff X \ \sim 0$ (i.e. X est presque sûrement nulle).

Dans le cas de variables à valeurs dans $\mathbb{N}$, on dispose d'une autre description possible de $\mathbb{E}(X)$.

Proposition 15.2.7 – Réxpression de l'espérance lorsque $X(\Omega) \subset \mathbb{N} \cup \{+\infty\}$

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N} \cup \{+\infty\}$. Alors

$$\mathbb{E}(X) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geqslant n).$$

II.2 Espérance d'une variable complexe

Définition 15.2.8 – Espérance

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{C}$. On dit que X admet une espérance finie si la famille $(x\mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable, et dans ce cas :

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x).$$

La notation $X \in L^1$ signifie que X admet une espérance finie.

Remarque 15.2.9

Il arrive souvent que l'ensemble dénombrable $X(\Omega)$ soit énuméré, c'est-à-dire décrit sous forme d'une suite $(x_n)_{n \in \mathbb{N}}$. La condition d'existence d'une espérance finie équivaut dans ce cas à la convergence absolue de la série $\sum x_n \mathbb{P}(X = x_n)$.
On prendra garde au fait que la semi-convergence n'est pas suffisante pour assurer la sommabilité. En effet, comme on l'a vu, la valeur de la somme est dans ce cas fortement dépendante de l'ordre de sommation. Or l'ordre de sommation défini par l'énumération $(x_n)_{n \in \mathbb{N}}$ est purement arbitraire, et n'a pas de raison d'être privilégié par rapport à un autre.

Exemple 15.2.10

$X(\Omega) = \{(-1)^n n,\ n \in \mathbb{N}^*\}$, et : $\forall n \in \mathbb{N}^*,\quad \mathbb{P}(X = (-1)^n n) = \dfrac{6}{\pi^2 n^2}$.

Remarque 15.2.11

Lorsque $X \geqslant 0$, $X \in L^1 \iff \mathbb{E}(X) < +\infty$.

En revanche, lorsque X n'est pas positive, cela n'a pas de sens de donner une valeur infinie à $\mathbb{E}(X)$. Comme dans le cas positif, on peut comparer l'espérance de deux variables de même loi.

Proposition 15.2.12 – Espérance de variables de même loi

Soit $X \sim Y$ deux variables discrètes à valeurs complexes de même loi. Alors :

- X admet une espérance finie ssi Y admet une espérance finie ;
- dans ce cas, $\mathbb{E}(X) = \mathbb{E}(Y)$.

II.3 Formule de transfert et conséquences

Théorème 15.2.13 – Formule de transfert, version positive

Soit X une variable aléatoire à valeurs dans E et $f : X(\Omega) \to \overline{\mathbb{R}}_+$. Alors, dans $\overline{\mathbb{R}}_+$, l'égalité suivante est toujours valide :

$$\mathbb{E}(f(X)) = \sum_{x \in X(\Omega)} f(x)\mathbb{P}(X = x)$$

◅ **Éléments de preuve.**
Partir de l'espérance de $f(X)$, et, par propriété d'associativité positive, regrouper les valeurs de x selon les parts $\{f(X) = y\}$. ▻

Théorème 15.2.14 – Formule de transfert, version complexe

Soit X une variable aléatoire à valeurs dans E et $f : X(\Omega) \to \overline{\mathbb{R}}_+$. Les propositions suivantes sont équivalentes:

(i) $f(X) \in L^1$;

(ii) $(f(x)\mathbb{P}(X = x))_{x \in X(\Omega)}$.

Dans ce cas,

$$\mathbb{E}(f(X)) = \sum_{x \in X(\Omega)} f(x)\mathbb{P}(X = x).$$

Exemple 15.2.15

Si X est une variable discrète réelle. En cas de sommabilité,

$$\mathbb{E}(X^k) = \sum_{x \in X(\Omega)} x^k \mathbb{P}(X = k).$$

On dit dans ce cas que $X \in L^k$, et $\mathbb{E}(X^k)$ est appelé moment d'ordre k de X.

Remarque 15.2.16

Si (X, Y) est un couple de variables aléatoires définies sur E et F respectivement, et $f : E \times F \to \mathbb{C}$, on peut considérer (X, Y) comme une variable aléatoire à valeurs dans $E \times F$, ce qui permet d'utiliser la formule de transfert : $f(X, Y)$ admet une espérance finie ssi $(f(x, y)\mathbb{P}(X = x, Y - y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est sommable. Dans ce cas,

$$\mathbb{E}(f(X, Y)) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} f(x, y)\mathbb{P}(X = x, Y = y).$$

La formule de transfert permet d'obtenir des propriétés importantes de l'espérance.

Proposition 15.2.17 – Inégalité triangulaire

Soit X une variable aléatoire à valeurs complexes. Alors

$$X \in L^1 \iff \mathbb{E}(|X|) < +\infty,$$

et dans ce cas,

$$|\mathbb{E}(X)| \leqslant \mathbb{E}(|X|).$$

Proposition 15.2.18 – Linéarité de $\mathbb{E}$, cas positif

Soit X et Y deux variables aléatoires discrètes à valeurs dans $\overline{\mathbb{R}}_+$, et $\lambda \in \mathbb{R}_+^*$. Alors, dans $\overline{\mathbb{R}}_+$,

$$\mathbb{E}(\lambda X + Y) = \lambda \mathbb{E}(X) + \mathbb{E}(Y)$$

Proposition 15.2.19 – Linéarité de $\mathbb{E}$, cas complexe

Soit X et Y deux variables aléatoires discrètes à valeurs dans $\mathbb{C}$, et $\lambda \in \mathbb{C}$. Si $X \in L^1$ et $Y \in L^1$, alors $\lambda X + Y \in L^1$ et dans ce cas,

$$\mathbb{E}(\lambda X + Y) = \lambda \mathbb{E}(X) + \mathbb{E}(Y).$$

◁ **Éléments de preuve.**
Théorème de transfert, appliqué à $f : (x, y) \mapsto \lambda x + y$, puis manipulations simples de somme et expression des lois marginales en fonction de la loi conjointe. ▷

La propriété de positivité de la proposition suivante a déjà été démontré.

Proposition 15.2.20 – Positivité et croissance de l'espérance

1. Soit (X, Y) deux variables à valeurs dans $\overline{\mathbb{R}}_+$. Alors, dans $\overline{\mathbb{R}}_+$,

$$X \leqslant Y \quad \Longrightarrow \quad \mathbb{E}(X) \leqslant \mathbb{E}(Y).$$

2. Soit $X, Y \in L^1$ deux variables aléatoires réelles discrètes admettant une espérance
   - Si $X \geqslant 0$ ps, alors $\mathbb{E}(X) \geqslant 0$, avec égalité ssi $X = 0$ ps.
   - Si $X \geqslant Y$ ps, alors $\mathbb{E}(X) \geqslant \mathbb{E}(Y)$.

◁ **Éléments de preuve.**
Dire que $X \geqslant Y$ ps signifie qu'il existe $\Omega' \subset \Omega$ tel que $\mathbb{P}(\Omega') = 1$ et $\forall \omega \in \Omega',\ X(\omega) \geqslant Y(\omega)$. ▷

Corollaire 15.2.21 – L^1 par majoration

Soit X et Y deux variables à valeurs dans $\mathbb{C}$. Si $|X| \leqslant |Y|$ et $Y \in L^1$, alors $X \in L^1$.

Proposition 15.2.22 – Espérance d'un produit de variables indépendantes

Soit X et Y deux variables à valeurs dans $\mathbb{C}$. Si X et Y sont L^1, alors $XY \in L^1$ et

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Par récurrence immédiate, si $X_1, \dots, X_n$ sont des variables indépendantes et L^1, alors $X_1 \cdots X_n$ est aussi L^1, et

$$\mathbb{E}(X_1 \dots X_n) = \prod_{k=1}^{n} \mathbb{E}(X_k).$$

II.4 Moment d'ordre 2 et variance d'une variable réelle

Dans ce paragraphe et le suivant, les variables aléatoires sont discrètes et réelles.
Comme dans le cas des variables finies, la valeur moyenne n'est pas une donnée suffisante pour avoir une bonne compréhension d'une variable aléatoire. En effet, l'espérance ne mesure pas la dispersion de la variable aléatoire : les valeurs prises peuvent tout aussi bien rester toujours très proches de $\mathbb{E}(X)$, ou alors s'en éloigner beaucoup, de part et d'autre, de façon à ce qu'en moyenne cela se compense.
Pour des raisons techniques, il est plus facile de travailler avec la moyenne quadratique $\mathbb{E}((X - \mathbb{E}(X))^2)$ de la distance à $\mathbb{E}(X)$, plutôt que la moyenne arithmétiques $\mathbb{E}(|X - \mathbb{E}(X)|)$. En effet, ce choix nous ramène à des méthodes bilinéaires.
On a déjà défini les moments d'ordre k. Les seuls au programme sont les moments d'ordre 1 (l'espérance) et 2.

Définition 15.2.23 – Moment d'ordre 2

Soit X une varable aléatoire réelle discrète.

- Le moment d'ordre 2 de X est l'espérance de X^2. Ce moment d'ordre 2 est toujours bine défini dans $\overline{\mathbb{R}}_+$.

- Ainsi, d'après la formule de transfert,
$$\mathbb{E}(X^2) = \sum_{x \in X(\Omega)} x^2 \mathbb{P}(X = x).$$
- On dit que X admet un moment d'ordre 2 fini (et on note $X \in L^2$) si $\mathbb{E}(X^2) < +\infty$.

Proposition 15.2.24 – Espace des variables L^2 sur Ω

1. Soit X et Y deux variables aléatoires réelles discrètes sur Ω. Si X et Y sont dans L^2, alors $XY \in L^1$.
2. L'ensemble $L^2(\Omega, \mathbb{R})$ des variables aléatoires définies sur Ω et admettant un moment d'ordre 2 fini est un espace vectoriel sur $\mathbb{R}$.

Définition 15.2.25

$XY \in L^1$ s'obtient en majorant $|XY|$. Ensuite, développer $\mathbb{E}((\lambda X + Y)^2)$.

Corollaire 15.2.26 – Espérance d'une variable L^2

Soit X une variable aléatoire réelle discrète. Si $X \in L^2$, alors $X \in L^1$.

◁ **Éléments de preuve.**
Prendre $Y = 1$. ▷

Théorème 15.2.27 – Inégalité de Cauchy-Schwarz pour l'espérance

- Pour toutes variables aléatoires réelles discrètes X et Y dans L^2,
$$\mathbb{E}(XY)^2 \leqslant \mathbb{E}(X^2)\mathbb{E}(Y^2).$$
- Dans l'inégalité précédente, le cas d'égalité est obtenu si et seulement si X et Y sont colinéaires presque sûrement, donc s'il existe des constantes λ et μ réelles telles que $\lambda X + \mu Y \sim 0$.

◁ **Éléments de preuve.**
C'est l'inégalité de Cauchy Schwarz pour la forme bilinéaire symétrique positive $(X, Y) \mapsto \mathbb{E}(XY)$ définie sur $L^2(\Omega, \mathbb{R})$.
On rappelle que l'inégalité de Cauchy-Schwarz pour une forme bilinéaire symétrique positive φ peut être obtenue sans l'hypothèse de séparation, et que le cas d'égalité est alors obtenu pour le couple (x, y) si et seulement s'il existe λ et μ tels que $\varphi(\lambda x + \mu y, \lambda x + \mu y) = 0$. ▷

Afin d'alléger les énoncés, nous ne rappelons plus par la suite à chaque fois que les vaiables consédérées sont réelles discrètes (ce que nous sous-entendrons par l'écriture $X \in L^2$ que nous n'avons définie que dans ce cas).

Définition 15.2.28 – Variance et écart-type d'une variable aléatoire

Soit $X \in L^2$. Alors, en vertu de ce qui précède, $X - \mathbb{E}(X)$ est bien défini et L^2. On définit alors :
- la variance de X :
$$\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \sum_{x \in X(\Omega)} (x - \mathbb{E}(X))^2 \mathbb{P}(X = x).$$
- l'écart type de X :
$$\sigma(X) = \sqrt{\mathbb{V}(X)} = \sqrt{\mathbb{E}((X - \mathbb{E}(X))^2)}.$$

Proposition 15.2.29 – Positivité et stricte positivité de $\mathbb{V}$

Soit $X \in L^2$. Alors $\mathbb{V}(X) \geqslant 0$, l'égalité étant obtenue si et seulement si X est constante presque sûrement.

Dans la pratique, on calcule souvent la variance avec le résultat suivant.

Théorème 15.2.30 – Formule de König-Huygens

Soit $X \in L^2$. Alors

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

Proposition 15.2.31 – Quadraticité et invariance par translation de la variance

Soit a et b deux réels et $X \in L^2$. Alors $aX + b \in L^2$, et

$$\mathbb{V}(aX + b) = a^2\mathbb{V}(X).$$

Il arrive souvent, de normaliser les variables aléatoires, afin de pouvoir les comparer. Il faut dans ce cas à la fois normaliser la valeur moyenne et la dispersion.

Définition 15.2.32 – variable centrée, réduite

Une variable X est dite :
- centrée si $X \in L^1$, et $\mathbb{E}(X) = 0$
- réduite si $X \in L^2$, et $\mathbb{V}(X) = 1$
- centrée réduite si elle est à la fois centrée et réduite.

Proposition 15.2.33 – Variable centrée réduite associée à X

Soit $X \in L^2$ telle que $\sigma(X) > 0$. Il existe d'uniques réels $a > 0$ et b tels que la variable X^* définie par $X^* = aX + b$ soit centrée réduite. Plus précisément,

$$X^* = \frac{X - \mathbb{E}(X)}{\sigma(X)}.$$

On dit que X^* est la variable aléatoire centrée réduite associée à X.

◁ **Éléments de preuve.**
Analyse synthèse par linéarité de $\mathbb{E}$ et quadraticité de $\mathbb{V}$. ▷

II.5 Covariance

Pour l'instant, contrairement au cas de l'espérance, nous n'avons pas donné de moyen de calculer la variance d'une somme. Nous avons en effet besoin pour cela d'un outil supplémentaire, qui nous permet de mesurer la dépendance entre deux variables.
En effet on comprend bien que la variance de $X + Y$ va fortement dépendre de la corrélation entre X et Y. Pour prendre deux cas extrêmes, supposons que X et Y suivent une même loi symétrique par rapport à 0 :
- si $X = Y$, on aura alors $\mathbb{V}(X + Y) = \mathbb{V}(2X) = 4\mathbb{V}(X)$;
- si $X = -Y$, on aura alors $\mathbb{V}(X + Y) = 0$.

Ainsi, la seule donnée des lois de X et Y ne peut pas suffire à déterminer $\mathbb{V}(X + Y)$. La covariance $\mathrm{cov}(X, Y)$ va nous permettre de mesurer la propension de deux variables à évoluer de façon parallèle ou opposée.

Définition 15.2.34 – Covariance

Soit X et Y deux variables L^2. Alors $(X-\mathbb{E}(X))(Y-\mathbb{E}(Y)) \in L^1$ et on définit la *covariance* de X et Y par :

$$\operatorname{cov}(X,Y) = \mathbb{E}\big((X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\big).$$

Remarque 15.2.35 – Interprétation de la covariance

Si $\operatorname{cov}(X,Y) > 0$, $X-\mathbb{E}(X)$ et $Y-\mathbb{E}(Y)$ ont tendance à être de même signe, donc X et Y ont tendance à se situer du même côté de leur espérance. Ainsi :

- $\operatorname{cov}(X,Y) > 0$ signifie que X et Y ont tendance à évoluer parallèlement
- $\operatorname{cov}(X,Y) < 0$ signifie que X et Y ont tendance à avoir une évolution opposée.

Définition 15.2.36 – Variables décorrélées

Les variables X et Y sont dites décorrélées lorsque $\operatorname{cov}(X,Y) = 0$.

Proposition 15.2.37 – Propriétés de la covariance

Soit $X, Y \in L^2$.

1. $\operatorname{cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$;
2. $\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X)$ (symétrie)
3. cov est bilinéaire sur l'espace des *v.a.r.d.* sur $(\Omega, \mathcal{T}, \mathbb{P})$ admettant un moment d'ordre 2.
4. $\operatorname{cov}(X,X) = \mathbb{V}(X)$.
5. $\operatorname{cov}(1,X) = 0$.
6. Si X et Y sont indépendantes, $\operatorname{cov}(X,Y) = 0$.

◁ **Éléments de preuve.**

Assez immédiat par linéarité de l'espérance. ▷

Avertissement 15.2.38

La réciproque de la dernière propriété est fausse : il existe des variables décorrélées, mais non indépendantes. Voir les exercices pour des exemples.

Les propriétés 2, 3 et 4 permettent d'affirmer que cov est une forme bilinéaire symétrique positive, de forme quadratique associée égale à la variance. En particulier, on dispose d'une deuxième inégalité de Cauchy-Schwarz :

Théorème 15.2.39 – Inégalité de Cauchy-Schwarz pour la covariance, HP

Soit $X, Y \in L^2$. Alors

$$|\operatorname{cov}(X,Y)| \leqslant \sigma(X)\sigma(Y),$$

avec égalité si et seulement s'il existe une relation affine presque sûrement entre X et Y (c'est-à-dire une relation non triviale $aX + bY + c = 0$)

Remarque 15.2.40

Sans être explicitement au programme, cette inégalité peut être considérée comme étant au programme, puisqu'il ne s'agit de rien d'autre que l'inégaliét de Cauchy-Schwarz pour l'espérance appliquée aux variables centrées $X-\mathbb{E}(X)$ et $Y-\mathbb{E}(Y)$.

II.6 Variance d'une somme

Une autre formule issue directement de l'algèbre bilinéaire (et du lien entre une forme bilinéaire symétrique et la forme quadratique associée) est la formule de polarisation :

Proposition 15.2.41 – Formule de polarisation

Soit $X, Y \in L^2$. Alors :

$$\mathrm{cov}(X,Y) = \frac{1}{2}(\mathbb{V}(X+Y) - \mathbb{V}(X) - \mathbb{V}(Y)).$$

Cette formule est dans notre contexte surtout utilisée pour calculer la variance d'une somme à l'aide des covariances. Nous la reexprimons donc de la façon suivante :

Proposition 15.2.42 – Variance d'une somme

1. Si $X, Y \in L^2$, alors $X+Y \in L^2$, et :

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2 \cdot \mathrm{cov}(X,Y).$$

2. Si de plus $X \perp\!\!\!\perp Y$, alors $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

En utilisant la propriété de bilinéarité généralisée, on obtient de même l'expression de la variance d'une somme de n terme :

Proposition 15.2.43 – Variance de $X_1 + \cdots + X_n$

1. Soit $X_1, \dots, X_n \in L^2$; alors $X_1 + \cdots + X_n \in L^2$ et :

$$\mathbb{V}(X_1 + \cdots + X_n) = \mathbb{V}(X_1) + \cdots + \mathbb{V}(X_n) + 2 \cdot \sum_{1 \leqslant i < j \leqslant n} \mathrm{cov}(X_i, X_j).$$

2. Si de plus les X_i sont 2 à 2 indépendantes, ou plus généralement 2 à 2 décorrélées,

$$\mathbb{V}(X_1 + \cdots + X_n) = \mathbb{V}(X_1) + \cdots + \mathbb{V}(X_n).$$

On peut réexprimer cette formule matriciellement.

Proposition 15.2.44 – Reexpression matricielle de la variance d'une somme

Soit $C = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$. Alors :

$$\mathbb{V}(X_1 + \cdots + X_n) = C^\top \cdot \mathbb{V}(X_1, \dots, X_n) \cdot C,$$

où $\mathbb{V}(X_1 + \cdots + X_n)$ est la matrice ci-dessous, appelée matrice des variances-covariances :

$$V(X_1, \dots, X_n) = (\mathrm{cov}(X_i, X_j))_{(i,j) \in [\![1,n]\!]^2}.$$

Ainsi, la variance est la somme des coefficients de la matrice des variances-covariances. Ce point de vue simplifie souvent les manipulation sommatoires, notamment lorsque l'expression des covariances suit un motif en diagonales (constante sur chaque diagonale par exemple). Les considérations combinatoires déviennent plus évidentes sur la représentation matricielle.

III Variables aléatoires à valeurs dans $\mathbb{N}$, lois usuelles

Nous voyons maintenant les lois classiques au programme. Avant de faire leur étude, on remarque que toutes les lois classiques sont à valeurs dans $\mathbb{N}$. Dans ce contexte on dispose d'un outil calculatoire efficace, apprécié également dans un contexte combinatoire.

III.1 Fonction génératrice d'une variable à valeurs dans $\mathbb{N}$

Définition 15.3.1 – Série et fonction génératrice d'une variable à valeurs dans $\mathbb{N}$

- Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. La série génératrice de X est la série entière $\sum_{n=0}^{+\infty} \mathbb{P}(X = n)t^n$.
- La fonction génératrice de X est la somme de la série génératrice et est notée G_X :

$$\Gamma_X(t) = \sum_{n=0}^{+\infty} \mathbb{P}(X = n)t^n,$$

pour tout t dans le domaine de convergence de la série génératrice.

Proposition 15.3.2 – Rayon de convergence de la série génératrice

- La série génératrice $\sum \mathbb{P}(X = n)t^n$ converge normalement sur $\overline{B}(0, 1) \subset \mathbb{C}$.
- En particulier, son rayon R vérifie $R \geqslant 1$.

En utilisant la formule de transfert, on obtient alors :

Corollaire 15.3.3 – Réexpression de G_X comme une espérance

Pour tout $t \in \mathring{B}(0, R) \cup \overline{B}(0, 1)$,

$$\Gamma_X(t) = \mathbb{E}(t^X).$$

Corollaire 15.3.4 – Régularité de G_X

1. La fonction génératrice G_X est de classe $\mathcal{C}^\infty$ sur $\mathring{B}(0, R)$ (donc en particulier sur $\mathring{B}(0, 1)$).
2. La fonction génératrice est continue sur $\overline{B}(0, 1)$.

Remarque 15.3.5

Si $R > 1$, le point 2 est redondant. En revanche, il apporte une information nouvelle si $R = 1$.

Les fonctions génératrices donnent un outil calculatoire efficace d'étude des variables aléatoires. En particulier, elle permet de reconnaître et comparer des variables aléatoires. En effet :

Théorème 15.3.6 – Détermination de la loi de X par sa fonction génératrice

1. Soit X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$. Si $G_X = G_Y$, alors $X \sim Y$.
2. Plus précisément,

$$\forall n \in \mathbb{N}, \quad \mathbb{P}(X = n) = \frac{G_X^{(n)}(0)}{n!}.$$

La fonction génératrice donne également un accès direct à l'espérance, et presque directe à la variance.

Théorème 15.3.7 – Expression de l'espérance à l'aide de la fonction génératrice

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$, et G_X sa fonction génératrice, considérée comme fonction d'une variable réelle. Les propositions suivantes sont équivalentes:

(i) $X \in L^1$

(ii) G_X est dérivable en 1.

Dans ce cas, on a $\mathbb{E}(X) = G'_X(1)$.

◁ **Éléments de preuve.**

- Si le rayon R de la série génératrice vérifie $R > 1$, R est aussi le rayon de la série dérivée, et cela résulte des propriétés de dérivation des séries entières.
- Sinon, le résultat est un corollaire du théorème 13.2.15 (la dérivabilité correspond dans ce cas à la dérivabilité à gauche)

  Comme le théorème 13.2.15 n'est pas explicitement au programme, on rappelle que le sens direct $(i) \implies (ii)$ est une application simple du théorème d'Abel radial. La démonstration du sens réciproque $(ii) \implies (i)$ n'est pas exigible.

▷

Théorème 15.3.8 – Expression de l'espérance à l'aide de la fonction génératrice

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. Les propositions suivantes sont équivalentes:

(i) $X \in L^2$

(ii) G_X est deux fois dérivable en 1.

Dans ce cas, on a $\mathbb{V}(X) = G''_X(1) + G'_X(1) - G'_X(1)$.

◁ **Éléments de preuve.**

$X \in L^2$ ssi $X(X-1) \in L^1$. Cela correspond, par formule de transfert, à la dérivée seconde de la série génératrice.

Ensuite utiliser la formule de König-Huygens. ▷

Une dernière propriété importante est l'accès facile à la loi d'une somme de variables indépendantes.

Théorème 15.3.9 – Fonction génératrice d'une somme de variables indépendantes

1. Soit X et Y deux variables aléatoires indépendantes, à valeurs dans $\mathbb{N}$, et R_X et R_Y les rayons de convergence de leurs séries génératrices. Alors

$$\forall t \in \mathring{B}(0, \min(R_X, R_Y) \cup \overline{B}(0,1), \quad G_{X+Y}(t) = G_X(t)G_Y(t).$$

2. Soit $X_1, \dots, X_n$ des variables indépendantes à valeurs dans $\mathbb{N}$, et R_i le rayon de la série génératrice de X_i. Alors

$$\forall t \in \mathring{B}(0, \min_{i \in [\![1,n]\!]}(R_i) \cup \overline{B}(0,1), \quad G_{X+Y}(t) = G_X(t)G_Y(t).$$

◁ **Éléments de preuve.**

Le point 1 par produit de Cauchy. Le point 2 par récurrence. ▷

III.2 Rappels sur les lois usuelles finies

On rappelle rapidement les lois usuelles finies classiques, en complétant leur étude par la description de leur fonction génératrice. On présente une alternative au calcul direct de leur espérance et de leur variance, en se servant des séries génératrices, et on se sert des fonctions génératrices également pour justifier une propriété de stabilité.

Définition 15.3.10 – Loi uniforme

Soit $n \in \mathbb{N}^*$. Une v.a.r. X suit la loi uniforme de paramètre n si $X(\Omega) = [\![1, n]\!]$ et si :

$$\forall k \in [\![1, n]\!], \quad \mathbb{P}(X = k) = \frac{1}{n}.$$

On note $X \sim \mathcal{U}(n)$.

Proposition 15.3.11 – Fonction génératrice, espérance et variance d'une loi uniforme

Soit $X \sim \mathcal{U}(n)$. Alors

$$\forall t \in \mathbb{R}, \quad \Gamma_X(t) = \begin{cases} \frac{t}{n} \cdot \frac{1-t^n}{1-t} & \text{si } t \neq 1 \\ 1 & \text{si } t = 1 \end{cases}$$

et :

$$\mathbb{E}(X) = \frac{n+1}{2} \qquad \text{et} \qquad \mathbb{V}(X) = \frac{n^2-1}{12}.$$

◁ **Éléments de preuve.**
C'est juste pour refaire le calcul avec la fonction génératrice, qui donne une version plutôt plus compliquée ici que la méthode directe. ▷

Pour tout ensemble fini E, on note plus généralement $X \sim \mathcal{U}(E)$ pour une v.a.r. suivant la loi uniforme sur X, c'est-à-dire une v.a.r. telle que :

$$\forall x \in E, \quad \mathbb{P}(X = x) = \frac{1}{|E|}.$$

Proposition 15.3.12 – Expérience-type associée à une loi uniforme

Soit U une urne contenant n boules numérotées de 1 à n. On effectue un tirage avec équiprobabilité, et on note X le numéro de la boule obtenu. Alors $X \sim \mathcal{U}(n)$.

Plus généralement, si X est une variable suivant une loi uniforme sur l'ensemble $[\![a, b]\!]$, elle vérifie :

$$\mathbb{E}(X) = \frac{a+b}{2} \qquad \text{et} \qquad \mathbb{V}(X) = \frac{(b-a+1)^2 - 1}{12}.$$

Définition 15.3.13 – Loi de Bernoulli

Soit X une v.a.r. et $p \in]0, 1[$. On dit que X suit une loi de Bernoulli de paramètre p (et on note $X \sim \mathcal{B}(1, p)$ ou $X \sim \mathcal{B}(p)$) si :

$$X(\Omega) = \{0, 1\} \qquad \text{et} \qquad \begin{cases} \mathbb{P}(X = 0) = 1 - p \\ \mathbb{P}(X = 1) = p \end{cases}$$

Désormais, on se donne un réel $p \in]0, 1[$, et on note $q = 1 - p$.

Proposition 15.3.14 – Espérance et variance d'une loi de Bernoulli

Soit $X \sim \mathcal{B}(p)$. Alors

$$\forall t \in \mathbb{R}, \quad G_X(t) = (1-p) + tp, \qquad \mathbb{E}(X) = p \qquad \text{et} \qquad \mathbb{V}(X) = pq.$$

◁ **Éléments de preuve.**
Un peu trivial... Mais le vérifier aussi avec la série génératrice. ▷

Exemple 15.3.15 – Fonction indicatrice d'un événement

Étude de $X(\omega) = \mathbb{1}_A(\omega)$, $A \in \mathcal{P}(\Omega)$.

La motivation même de la définition de la loi de Bernoulli amène :

Proposition 15.3.16 – Expérience-type associée à une loi de Bernoulli

Soit une expérience à 2 issue (par exemple P/F, ou tirage dans une urne contenant 2 types de boules), l'une représentant le succès et l'autre l'échec, la probabilité d'obtenir un succès étant p. Soit X prenant la valeur 1 en cas de succès et la valeur 0 en cas d'échec. Alors $X \sim \mathcal{B}(p)$.

Une expérience telle que décrite dans la proposition précédente est usuellement appelée « expérience de Bernoulli ».

Remarque 15.3.17

Toute variable aléatoire prenant ses valeurs dans $\{0,1\}$ est une variable de Bernoulli, à condition d'élargir la définition en acceptant de considérer les cas dégénérés $p = 0$ et $p = 1$.
De façon équivalente, si une variable aléatoire à valeurs dans $\mathbb{N}$ a une fonction génératrice affine, alors il s'agit d'une variable de Bernoulli (avec cette dafinition étendue).

Définition 15.3.18 – Loi binomiale

Soit X une v.a.r.. On dit que X suit la loi binomiale de paramètres (n,p) (et on note $X \sim \mathcal{B}(n,p)$) si :

$$X(\Omega) = [\![0,n]\!], \qquad \text{et} \qquad \forall k \in [\![0,n]\!],\ \mathbb{P}(X=k) = \binom{n}{k} p^k q^{n-k}, \quad \text{où } q = 1-p.$$

Ceci définit bien une loi de probabilités. En effet : $\displaystyle\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1$.

Un calcul élémentaire (fait en MPSI) amène :

Proposition 15.3.19 – Expérience-type associée à la loi binomiale

Soit X le nombre de succès obtenus lors d'une répétition de n épreuves de Bernoulli indépendantes de même paramètre p. Alors $X \sim \mathcal{B}(n,p)$.

Proposition 15.3.20 – Espérance et variance d'une loi binomiale

Soit $X \sim \mathcal{B}(n,p)$. Alors

$$\forall t \in [-1,1], \quad G_X(t) = ((1-p)+tp)^n, \varepsilon\mathbb{E}(X) = np, \qquad \text{et} \qquad \mathbb{V}(X) = npq.$$

◁ **Éléments de preuve.**
La série génératrice se calcule par la formule du binôme. Dériver pour obtenir l'espérance et la variance. ▷

III.3 Loi géométrique

Nous ajoutons à ce catalogue deux nouvelles lois, discrètes, mais non finies.
La première correspond à la loi du temps d'attente dans une succession infinie de tirages à Pile ou Face. On sait, d'après le théorème 15.1.36, définir un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ modélisant cette expérience. On se place dans cet espace probabilisé.

Proposition 15.3.21 – Temps d'attente du premier succès

Soit T la variable égale au rang du premier succès lors d'une succession infinie d'expériences de Bernoulli indépendantes de paramètre p, évenutellement égal à $+\infty$ si le succès n'est jamais obtenu. Alors $X(\Omega) = \mathbb{N}^* \cup \{+\infty\}$, et

$$\forall n \in \mathbb{N},\ \mathbb{P}(X=n) = pq^{n-1}, \qquad \text{et} \qquad P(X=+\infty) = 0.$$

Ainsi, la valeur $+\infty$ n'est pas pertinente dans la description de la loi, et on peut définir la loi générique de X sans la valeur infinie.

Définition 15.3.22 – Loi géométrique

On dit que X suit une loi géométrique de paramètre $p \in]0,1[$, si $(X \in \mathbb{N}^*)$ est presque sûr, et

$$\forall n \in \mathbb{N}^*, \quad \mathbb{P}(X = n) = pq^{n-1}, \qquad \text{avec } q = 1 - p.$$

On note $X \sim \mathcal{G}(p)$.

Et par notre motivation-même :

Proposition 15.3.23 – Expérience-type associée à une loi géométrique

Soit X le temps d'attente du premier succès lors d'une répétition infinie d'expériences de Bernoulli indépendantes de même paramètre p. Alors $X \sim \mathcal{G}(p)$.

Proposition 15.3.24 – Fonction génératrice, espérance et variance d'une loi géométrique

Soit $X \sim \mathcal{G}(p)$. Alors

$$\forall t \in \left]-\frac{1}{q}, \frac{1}{q}\right[, \quad \Gamma_X(t) = \frac{pt}{1-qt}, \qquad \mathbb{E}(X) = \frac{1}{p} \qquad \text{et} \qquad \mathbb{V}(X) = \frac{q}{p^2}.$$

◁ **Éléments de preuve.**

On présente le calcul de l'espérance de X par fonction génératrice et par calcul direct. ▷

III.4 Loi de Poisson

Définition 15.3.25 – Loi de Poisson

Soit $\lambda \in \mathbb{R}_+$ et X une v.a.r.. On dit que X suit la loi de Poisson de paramètre λ, et on note $X \sim \mathcal{P}(\lambda)$, si :

$$X(\Omega) = \mathbb{N} \qquad \text{et} \qquad \forall n \in \mathbb{N}, \quad \mathbb{P}(X = n) = \mathrm{e}^{-\lambda} \cdot \frac{\lambda^n}{n!}.$$

D'après les résultats sur les séries exponentielles, cela définit bien une loi de probabilité. De plus :

Proposition 15.3.26 – Espérance et variance d'une loi de Poisson

Soit $X \sim \mathcal{P}(\lambda)$. Alors

$$\forall t \in \mathbb{R}, \quad G_X(t) = \mathrm{e}^{\lambda(t-1)}, \qquad \mathbb{E}(X) = \lambda \qquad \text{et} \qquad \mathbb{V}(X) = \lambda.$$

Proposition 15.3.27 – Limite en loi de variables binomiales

Soit (X_n) une suite de variables binomiales de paramtères (n, p_n), telles que $p_n \sim_{+\infty} \frac{\lambda}{n}$. Alors

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(X_n = k) \longrightarrow \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}.$$

On dit que la suite (X_n) converge en loi vers une loi de Poisson de paramètre λ.

Remarque 15.3.28

Pour cette raison, la loi de Poisson modélise le nombre de succès lorsque le succès est un événement rare (i.e. de petite probabilité) et que la population est très grande.
Une situation typique est l'affluence (à un magasin, à un parc d'attraction etc), lors d'un laps de

temps donné (par exemple une journée, dans une population très grande. Chaque individu de la population a une probabilité p très faible d'aller à ce lieu lors d'une journée donnée, et la population N est très grande. Ainsi, la loi $\mathcal{B}(N,p)$ représentant l'affluence au lieu donné pendant la période donnée est assez voisine de la loi $\mathcal{P}(Np)$. Ainsi, en posant $\lambda = Np$, la loi de Poisson $\mathcal{P}(\lambda)$ est une bonne modélisation de l'affluence.

III.5 Stabilités (HP classique)

Par stabilité d'une loi classique, on entend la chose suivante : si X et Y suit un certain type de loi et sont indépendante, alors $X+Y$ suit une loi de même type (avec des paramètres éventuellement différents). Nous énonçons ici deux propriétés de stabilité.

Proposition 15.3.29 – Stabilité des lois binomiales

Soit $(m,n) \in (\mathbb{N}^*)^2$ et $p \in]0,1[$. Si $X \sim \mathcal{B}(n,p)$ et $Y \sim \mathcal{B}(m,p)$ sont indépendantes, alors $X+Y \sim \mathcal{B}(n+m,p)$.

◁ **Éléments de preuve.**
Se comprend bien par l'expérience. Assez immédiat par fonctions génératrices. ▷

Corollaire 15.3.30 – Somme de variables de Bernoulli indépendantes

Soit $X_1,\dots,X_n$ des variables aléatoires mutuellement indépendantes suivant la loi $\mathcal{B}(p)$. Alors $X_1+\cdots+X_n \sim \mathcal{B}(n,p)$.

Proposition 15.3.31 – Stabilité des lois de Poisson

Soit $(\lambda,\mu) \in (\mathbb{R}_+)^2$. Si $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$ sont indépendantes, alors $X+Y \sim \mathcal{P}(\lambda+\mu)$.

◁ **Éléments de preuve.**
Même principe de calcul. ▷

IV Inégalités probabilistes

Pour terminer, nous généralisons au cas de variables discrètes les inégalités de Markov et de Bienaymé-Tchebychev vues en MPSI pour les varables finies.

IV.1 Inégalités de Markov et de Bienaymé-Tchebychev

Théorème 15.4.1 – Inégalité de Markov

Soit X une variable aléatoire réelle discrète dans L^1. Alors, pour tout $a>0$,

$$\mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}(X)}{\varepsilon}.$$

◁ **Éléments de preuve.**
Par croissance de $\mathbb{E}$, en remarquant que $X \geqslant a\mathbb{1}_{\{X\geqslant a\}}$. ▷

Corollaire 15.4.2 – Réexpression quadratique de l'inégalité de Markov dans le cas L^2

Soit X une variable aléatoire réelle discrète dans L^2, et $a \in \mathbb{R}_+^*$. Alors

$$\mathbb{P}(|X| \geqslant \varepsilon) \leqslant \frac{\mathbb{E}(X^2)}{\varepsilon^2}.$$

◁ **Éléments de preuve.**
Si $\mathbb{E}(X^2) \neq 0$, appliquer ce qui précède à X^2. Et sinon ? ▷

Théorème 15.4.3 – Inégalité de Bienaymé-Tchebychev

Soit Y une variable aléatoire réelle discrète dans L^2, d'espérance m et de variance σ^2. Alors :

$$\forall \varepsilon > 0,\ \mathbb{P}\big(|Y - m| \geqslant \varepsilon\big) \leqslant \frac{\sigma^2}{\varepsilon^2}.$$

◁ **Éléments de preuve.**
Inégalité de Markov quadratique appliquée à... ▷

IV.2 Loi faible des grands nombres

On a déjà vu la notion de convergence en loi de X_n vers X. Cette notion de convergence ne dit rien en terme de réalisation des variables, mais juste sur les probabilités.
Une notion plus probabiliste de convergence mesure à quel point deux variables ont tendance à prendre des valeurs proches l'une de l'autre.

Définition 15.4.4 – Convergence en probabilités

Soit (X_n) une suite de variables aléatoires L^1 et X une variable L^1. On dit que (X_n) converge en probabilités vers X si

$$\forall \varepsilon > 0, \quad \lim_{n \to +\infty} \mathbb{P}(|X_n - X| \geqslant \varepsilon) = 0.$$

Un des exemples les plus importants de convergence en probabilité est celle qui résulte du calcul de la moyenne obtenue lors d'une répétition de variables aléatoires.

Théorème 15.4.5 – Loi faible des grands nombres, HP

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles discrètes L^2, i.i.d, d'espérance m et de variance σ^2. Soit $(Z_n)_{n \in \mathbb{N}^*}$ définie par :

$$\forall n \in \mathbb{N}, \quad Z_n = \frac{X_1 + \cdots + X_n}{n}.$$

Alors $(Z_n)_{n \in \mathbb{N}^*}$ converge en probabilité vers la variable certaine égale à m. Plus précisément :

$$\forall \varepsilon > 0,\ \forall n \in \mathbb{N}^*, \quad \mathbb{P}\big(|Z_n - m| \geqslant \varepsilon\big) \leqslant \frac{\sigma^2}{n\varepsilon^2}.$$

◁ **Éléments de preuve.**
Inégalité de B-T appliquée à Z_n dont on sait facilement déterminer l'espérance et la variance. ▷

Exemple 15.4.6 – Théorème d'or de Bernoulli

1. La loi faible des grands nombres utilisée avec une suite de variables de Bernoulli i.i.d. de loi $\mathcal{B}(p)$, avec p non connu, on obtient

$$\forall \varepsilon > 0,\ \forall n \in \mathbb{N}^*, \quad \mathbb{P}\big(|Z_n - p| \geqslant \varepsilon\big) \leqslant \frac{1}{4n\varepsilon^2}.$$

Le fait qu'on puisse faire cette majoration indépendamment de p permet d'estimer la valeur de p en répétant un grand nombre de fois une expérience, et en donnant un intervalle de confiance.

2. On tire 1000 fois à pile ou face avec une pièce déséquilibrée dont la probabilité d'obtention de Pile est p. On obtient 570 fois Pile. Donner un intervalle I tel que la probabilité que $p \in I$ soit supérieure à 0.9.

16

Calcul différentiel et optimisation

L'objectif de ce chapitre est l'étude de propriétés généralisant la dérivation, dans le cadre de fonctions définies sur un e.v.n. de dimension finie E.
En MPSI, vous avez par exemple étudié le cas de fonctions de 2 variables, c'est-à-dire de fonctions $f : \Omega \to \mathbb{K}$ où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et Ω est un ouvert de $\mathbb{R}^2$. Tout élément de Ω peut être décrit dans la base canonique sous forme d'un couple $(x, y) \in \Omega$, ce qui permet de considérer une telle fonction comme une fonction de 2 variables réelles x et y. On a notamment vu dans ce cadre comment définir des dérivées partielles, et plus généralement des dérivées selon un vecteur.
On comprend facilement comment généraliser pour des fonctions de n variables, c'est à dire définies sur un ouvert Ω de $\mathbb{R}^n$ et à valeurs dans $\mathbb{K}$, ou plus généralement, à valeurs dans un $\mathbb{K}$-e.v.n. F.
De façon encore plus générale, on peut se demander si on peut s'affranchir de ce cadre très rigide de $\mathbb{R}^n$ pour la variable de départ, c'est-à-dire si on peut directement considérer des fonctions de Ω dans F, où Ω est un ouvert d'un $\mathbb{R}$-e.v.n. E, et F est un $\mathbb{K}$-e.v.n..
La notion de dérivée selon un vecteur se définit alors de même, et, si on est en dimension finie, est indépendante du choix des normes. Les dérivées partielles correspondent à des dérivées selon certains vecteurs particuliers formant une base de E. Les dérivées partielles peuvent donc se définir de façon plus générale en référence à une base de E.
Le point de vue des dérivées partielles est pratique d'un point de vue calculatoire, mais donne souvent un point de vue un peu trop partiel et morcelé sur la fonction f étudiée. Il peut être intéressant de regrouper les informations. Dans ce but, nous définirons la notion de différentielle en un point a, qui est un objet qui regroupe toutes les dérivées directionnelles en un point. Plus précisément, la différentielle de f au point a est (sous réserve d'existence et de linéarité) l'application qui à tout vecteur $\vec{u}$ de E associe le vecteur dérivé en a selon la direction de $\vec{u}$.
Cette notion de différentielle permet d'avoir des formules beaucoup plus compactes et naturelles que celles provenant de l'étude des dérivées partielles, et surtout, elle permet de s'affranchir de la donnée d'une base.

Dans tout le chapitre, $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$, E est un $\mathbb{R}$-e.v.n. (nécessairement réel) de dimension finie, F un $\mathbb{K}$-e.v.n. de dimension finie, et Ω un ouvert de E. Dans certains contextes liés à l'optimisation (recherche d'extrema), il pourra cependant être intéressant de considérer des fonctions définies sur un fermé, ou même sur un compact, afin d'assurer l'existence d'extrema globaux.

I Compléments sur les fonctions vectorielles

Cette section a pour but de généraliser rapidement certaines notions vues dans le cadre de fonctions de la variable réelle et à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ au cas de fonctions de la variable réelle aussi, mais à valeurs dans un e.v.n. F de dimension finie. Nous ne nous attardons pas sur ces différentes propriétés, qui sont assez limpides une fois qu'on a compris que, du fait de l'équivalence des normes, il suffit de se ramener systématiquement aux coordonnées. Ces propriétés seront complétées dans les sections suivantes par des propriétés un peu plus spécifiques (dérivation de formes bilinéaires ou multilinéaires composées par des fonctions dérivables) qui seront vues comme cas particuliers de propriétés similaires pour les fonctions d'une variable vectorielle.

I.1 Extention des notations o et O

Dans cette sous-section, E est un $\mathbb{R}$-e.v.n. de dimension finie, F est un $\mathbb{K}$-e.v.n. de dimension finie, et $D \subset E$ est une partie quelconque de E.

Définition 16.1.1 – Extension des notations o et O

Soit f est une application de D dans F et $g : D \to \mathbb{R}$. Soit $x_0 \in \overline{D}$. On dira que

- $f(x) \underset{x_0}{=} o(g(x))$ s'il existe $\alpha : D \to F$ tel que

$$\exists V \in \mathcal{V}(x_0),\quad \forall x \in V \cap D,\quad f(x) = g(x)\alpha(x),\qquad \text{et} \qquad \lim_{x\to x_0} \alpha(x) = 0;$$

- $f(x) \underset{x_0}{=} O(g(x))$ s'il existe $\alpha : D \to F$ tel que

$$\exists V \in \mathcal{V}(x_0),\quad \forall x \in V \cap D,\quad f(x) = g(x)\alpha(x),\qquad \text{et} \qquad \alpha \text{ bornée sur } V \cap D;$$

Remarque 16.1.2

Le terme d'erreur sera toujours contrôlé par une fonction à valeur dans $\mathbb{R}$, et non une fonction vectorielle.

Proposition 16.1.3 – Caractérisation de o et O par les coordonnées

Soit $f : I \to F$ et $g : I \to \mathbb{R}$, et $\mathcal{B} = (b_1, \dots, b_n)$ une base de F, et pour tout $x \in I$

$$f(x) = \sum_{i=1}^{n} f_i(x) b_i.$$

Alors

- $f(x) \underset{x_0}{=} o(g(x))$ si et seulement si pour tout $i \in [\![1, n]\!]$, $f_i(x) \underset{x_0}{=} o(g_i(x))$;
- $f(x) \underset{x_0}{=} O(g(x))$ si et seulement si pour tout $i \in [\![1, n]\!]$, $f_i(x) \underset{x_0}{=} O(g_i(x))$;

◁ **Éléments de preuve.**
Écrire α avec ses coordonnées, et utiliser la norme $\|\cdot\|_{\mathcal{B},\infty}$. ▷

Exemples 16.1.4

1. $\begin{pmatrix} t \\ t^2 \\ \ln(t) \end{pmatrix} \underset{+\infty}{=} o(t^3)$
2. $\begin{pmatrix} e^t & \sin(t) \\ \frac{1}{1+t} & \operatorname{Arctan}(t)) \end{pmatrix} \underset{0}{=} \begin{pmatrix} 1+t & t \\ 1-t & t \end{pmatrix} + O(t^2).$
3. Pour $A \in \mathcal{M}_n(\mathbb{C})$, $e^A \underset{A\to 0}{=} \sum_{k=0}^{n} \frac{A^k}{k!} + o(|\!|\!|A|\!|\!|^n).$

4. On en déduit, pour $t \in \mathbb{R}$: $e^{At} \underset{t\to 0}{=} I_n + tA + o(t)$.

La situation la plus fréquente sera le cas d'un terme d'erreur en $\|x\|^\alpha$, x étant une variable de l'e.v.n. E de dimension finie.

Lemme 16.1.5 – Invariance vis-à-vis de la norme en dimension finie

Soit E de dimension finie et $\|\cdot\|_1$ et $\|\cdot\|_2$ deux normes quelconques sur E. Soit $D \subset E$ et $f : E \to F$, et $x_0 \in \overline{F}$. Soit $\alpha \in \mathbb{R}$. Alors

$$f(x) \underset{x\to x_0}{=} o(\|x\|_1^\alpha) \iff f(x) \underset{x\to x_0}{=} o(\|x\|_2^\alpha)$$

et

$$f(x) \underset{x\to x_0}{=} O(\|x\|_1^\alpha) \iff f(x) \underset{x\to x_0}{=} O(\|x\|_2^\alpha).$$

Notation 16.1.6 – Notation $o(h)$, $O(h)$

Pour $h \in E$, on trouve parfois par abus la notation $o(h)$ à la place de $o(\|h\|)$, et de même pour $O(h)$, dans le but d'alléger un peu les notations.

Lorsque $F = \mathbb{R}$ ou $\mathbb{C}$, on peut former des quotients, et on retrouve la caractérisation usuelle, généralisée au cas d'une variable vectorielle.

Proposition 16.1.7 – Caractérisation de o par le quotient

Si $g : D \to \mathbb{R}$ ne s'annule pas au voisinage de x_0, les propositions suivantes sont équivalentes:

(i) $f(x) \underset{x_0}{=} o(g(x))$;

(ii) $\dfrac{f(x)}{g(x)} \underset{x\to 0}{\longrightarrow} 0$.

Proposition 16.1.8 – Caractérisation de O par le quotient

Si $g : D \to \mathbb{R}$ ne s'annule pas au voisinage épointé de x_0, les propositions suivantes sont équivalentes:

(i) $f(x) \underset{x_0}{=} O(g(x))$;

(ii) $\dfrac{f(x)}{g(x)}$ est bornée au voisinage de x_0.

Les propriétés opératoires sur les o et O sont les mêmes que dans le cas réel (sommation, transitivité etc).

I.2 Dérivation de fonctions vectorielles d'une variable réelle

Dans cette sous-section, on suppose que $E = \mathbb{R}$ et $I \subset E$ est un intervalle ouvert. On considère donc $f : I \to F$ une fonction d'une variable réelle à valeurs dans E.

Définition 16.1.9 – Fonction dérivable en a

Soit $f : I \to F$ et $a \in I$. On dit que f est dérivable en a si le taux d'accroissement

$$\tau_{x_0} f(x) = \frac{f(x) - f(a)}{x - a}$$

admet une limite finie $\ell \in F$ lorsque x tend vers a. Le vecteur dérivé en a est alors défini par :

$$f'(a) = \lim_{x\to x_0} \frac{f(x) - f(a)}{x - a} = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}.$$

La deuxième limite est prise en considérant $h \in I - x_0 = \{x - x_0, x \in I\}$. Ainsi, si x_0 est la borne inférieure de I, il s'agit d'une limite lorsque h tend vers 0^+.

Définition 16.1.10 – Fonction dérivable sur I, fonction dérivée

Soit $f : I \to F$. On dit que f est dérivable sur I si f est dérivable en tout $x_0 \in I$. On définit alors la fonction dérivée $f' : I \to F$ par :

$$\forall x \in I, \quad f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x}.$$

Exemples 16.1.11

1. Dérivée de $f : t \mapsto tX_0$, $X_0 \in \mathbb{K}$.
2. Plus généralement, dérivée de $f : t \mapsto g(t)X_0$, $g : I \to \mathbb{R}$ étant une fonction dérivable en a , et X_0 un vecteur de F.
3. Dérivée de $t \mapsto \mathrm{e}^{\mathrm{i}\,t}$.

Proposition 16.1.12 – Caractérisation de la dérivabilité par les coordonnées

Soit $f : I \to F$, $x_0 \in I$, et $\mathcal{B}$ une base de F. On note, pour tout $x \in I$,

$$f(x) = \sum_{i=1}^{n} f_i(x) b_i,$$

la décomposition de f suivant ses composantes sur la base $\mathcal{B}$. Les propositions suivantes sont équivalentes:

(i) f est dérivable en x_0 ;

(ii) pour tout $i \in [\![1, n]\!]$, f_i est dérivable en x_0.

Dans ce cas, on a

$$f'(x_0) = \sum_{i=1}^{n} f_i'(x_0) b_i.$$

◁ **Éléments de preuve.**
Former le taux d'accroissement coordonnée par coordonnée, et utiliser la caractéristique similaire pour les limites. ▷

La caractérisation par les coordonnées permet de montrer qu'en définissant la dérivabilité à gauche et à droite comme dans le cas d'une fonction réelle, et en utilisant la caractérisation de la dérivée par les dérivées à gauche et à droite sur chacune de coordonnée, cela permet d'obtenie une caractérisation similaire dans le cas vectoriel.

Corollaire 16.1.13 – Opérations sur les dérivées

Soit I et J des intervalles de $\mathbb{R}$, et F un e.v.n. de dimension finie.

1. La dérivation est linéaire : si $f, g : I \to F$ et $\lambda \in \mathbb{K}$, alors si f et g sont dérivables en $a \in I$, $\lambda f + g$ aussi, de dérivée
$$(\lambda f + g)'(a) = \lambda f'(a) + g'(a).$$
2. Si $\varphi : J \to I$ est une fonction réelle dérivable, et $f : I \to F$ une fonction vectorielle dérivable, alors
$$\forall t \in J, \quad (f \circ \varphi)'(t) = f' \circ \varphi(t) \cdot \varphi'(t).$$

◁ **Éléments de preuve.**
Voir coordonnées par coordonnées. ▷

Corollaire 16.1.14 – Dérivabilité à gauche, à droite

Soit $x_0 \in \mathring{I}$ et $f : I \to F$. La fonction f est dérivable en x_0 si et seulement si elle est dérivable à gauche et à droite en x_0 et $f'_g(x_0) = f'_d(x_0)$.

Corollaire 16.1.15 – Dérivable implique continu

Soit $f : I \to \mathbb{R}$ et $x_0 \in I$. Si f est dérivable en x_0, alors f est continue en x_0.

◁ **Éléments de preuve.**

La caractérisation précédente nous ramène au cas des fonctions d'une variable. On peut aussi utiliser directement la propriété suivante (DL à l'ordre 1, qu'on peut ensuite restreindre à l'ordre 0). ▷

Proposition 16.1.16 – Caractérisation de la dérivabilité par DL à l'ordre 1

Soit $f : I \to F$ et $x_0 \in I$. Les propositions suivantes sont équivalentes:

(i) f est dérivable en x_0 ;

(ii) il existe A et B dans F tels que $f(x_0 + h) \underset{h\to 0}{=} A + hB + o(h)$.

◁ **Éléments de preuve.**

Caractérisation de o par quotient, puisque h ne s'annulle pas au voisinage épointé de 0. Ou encore, se ramener la la dérivabilité coordonnée par coordonnée. ▷

On donne une première conséquence, qui est plutôt un exemple, suffisamment important pour qu'il soit un résultat du cours (il servira pour l'étude des équations différntielles).

Corollaire 16.1.17 – Dérivée de $t \mapsto \exp(At)$

Soit $A \in \mathcal{M}_n(\mathbb{C})$. La fonction $f : t \mapsto \exp(At)$ est dérivable sur $\mathbb{R}$, et

$$\forall t \in \mathbb{R}, \quad f'(t) = A\exp(At).$$

Une autre conséquence de la caractérisation par les DL est l'interprétation du vecteur dérivé en terme d'approximation de la trajectoire.

Corollaire 16.1.18 – Tangente à la trajectoire

Soit $f : I \to F$ et $x_0 \in I$. On suppose que f est dérivable en x_0 et que $f'(x_0) \neq 0$. Alors la droite affine $f(x_0) + \mathbb{R}f'(x_0)$ est tangente à la trajectoire (dans le sens où c'est la droite qui approche au mieux la trajectoire au voisinage de $f(x_0)$).

Remarque 16.1.19 – Interprétation cinétique de la dérivée

Si f représente la trajectoire d'un point dans F, et si f est dérivable en x_0 :

- $\|f'(x_0)\|_F$ est la vitesse instantanée du point (dépendant de la norme choisie sur F)
- le vecteur $f'(x_0)$ est tangent à la trajectoire, et indique le sens de déplacement.

Exemple 16.1.20

Décrire la trajectoire, et la vitesse instantanée d'un point, dont la position est définie pour $t \in \mathbb{R}_+$ par

$$f(t) = (\cos(t), \sin(t), t).$$

Théorème 16.1.21 – Caractérisation des fonctions constantes

Soit $f : I \to F$, dérivable sur l'intervalle I, et à valeurs dans l'e.v.n. de dimension finie F. Les propositions suivantes sont équivalentes:

(i) f est constante sur I

(ii) $f' = 0$ sur I.

◁ **Éléments de preuve.**

Voir dans une base. ▷

I.3 Fonctions vectorielles de classe $\mathcal{C}^k$

On considère toujours dans cette sous-section le cas d'une fonction $f : I \to F$ définie sur un intervalle I ouvert, et à valeurs dans un e.v.n. F de dimension finie. Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de F, et,pour $x \in I$, $[f(x)]_{\mathcal{B}} = \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix}$.

Si f est dérivable sur I, alors elle définit une fonction dérivée

$$\forall x \in I, \quad f'(x) = \sum_{k=1}^{n} f_i'(x) b_i, \qquad \text{i.e.} \qquad [f'(x)]_{\mathcal{B}} = \begin{pmatrix} f_1'(x) \\ \vdots \\ f_n'(x) \end{pmatrix}.$$

On a encore $f' : I \to F$, et on peut donc étudier la dérivabilité de f', et ainsi de suite, comme dans le cas réel.

Si f est dérivable p fois, on note $f^{(k)}$ sa dérivée d'ordre k.

Définition 16.1.22 – Fonction de classe $\mathcal{C}^k$

La fonction $f : I \to \mathbb{R}$ est de classe $\mathcal{C}^k$ si elle est k fois dérivable et si $f^{(k)} : I \to F$ est continue.

Proposition 16.1.23 – caractérisation de la classe $\mathcal{C}^k$ et expression de la dérivée

Avec les notations introduites ci-dessus,

1. f est k fois dérivable si et seulement si chaque f_i est k fois dérivable, et dans ce cas,

$$\forall x \in I, \quad f^{(k)}(x) = \sum_{i=1}^{n} f_i^{(k)}(x) b_i \qquad \text{i.e.} \qquad [f^{(k)}(x)]_{\mathcal{B}} = \begin{pmatrix} f_1^{(k)}(x) \\ \vdots \\ f_n^{(k)}(x) \end{pmatrix};$$

2. f est de classe $\mathcal{C}^k$ si et seulement si chaque f_i l'est.

Proposition 16.1.24 – Règles opératoires sur les fonctions de classe $\mathcal{C}^k$

Soit I et J deux intervalles de $\mathbb{R}$ et F un $\mathbb{K}$-e.v.n. de dimension finie. Soit $f, g : I \to \mathbb{R}$, $\varphi : J \to I$ et $\lambda \in \mathbb{K}$.

1. Si f et g sont de classe $\mathcal{C}^k$, alors $f = \lambda g$ est de classe $\mathcal{C}^k$, et

$$(f + \lambda g)^{(k)} = f^{(k)} + \lambda g^{(k)}.$$

2. Si f et φ sont de classe $\mathcal{C}^k$, $f \circ \varphi$ est de classe $\mathcal{C}^k$. Il n'y a pas de formule simple en général

3. Si f est de classe $\mathcal{C}^k$ et $\varphi : t \mapsto at + b$, alors $(f \circ \varphi)^{(k)} = a^k (f^{(k)} \circ \varphi)$.

◁ **Éléments de preuve.**
Ces propriétés sont bien connues pour les fonctions à valeurs réelles. S'y ramener en travaillant sur les coordonnées. ▷

II Dérivées partielles et différentielles

On s'intéresse maintenant à la situation plus générale d'une fonction $f : \Omega \to F$, où Ω est un ouvert d'un $\mathbb{R}$-e.v.n. E de dimension finie, et F est un $\mathbb{K}$-e.v.n. de dimension finie.

II.1 Dérivées selon un vecteur

Définition 16.2.1 – Fonctions partielles

Soit $f : \Omega \to F$, $a \in \Omega$, et $\vec{u} \in E \backslash \{0\}$. La fonction partielle $f_{a,\vec{u}}$ au point a et de direction $\vec{u}$, est définie, en tout $t \in \mathbb{R}$ tel que $a + t\vec{u} \in \Omega$, par

$$f_{a,\vec{u}}(t) = f(a + t\vec{u}).$$

Remarque 16.2.2

1. La fonction partielle $f_{a,\vec{u}}$ est bien définie au voisinage de 0, car Ω est ouvert.
2. La flèche sur $\vec{u}$ est présente ici uniquement pour faire remarquer que c'est la structure affine naturelle de E qui est en jeu ici. Les éléments de E jouent donc un rôle double de points et de vecteurs. Les éléments en lesquels f est évalué correspondent plutôt à des points, alors que $\vec{u}$ donne une direction de translation, et doit plutôt être considéré comme un vecteur. La fonction $f_{a,\vec{u}}$ correspond plus ou moins, après reparamétrage, à la restriction de f à la droite affine passant par a et dirigée par le vecteur $\vec{u}$.

Définition 16.2.3 – Dérivée selon un vecteur

Soit $f : \Omega \to F$, $a \in \Omega$, et $\vec{u} \in E \backslash \{0\}$.

1. La fonction f est dérivable selon le vecteur $\vec{u}$ si la fonction partielle $f_{a,\vec{u}}$ est dérivable en 0.
2. La dérivée de f en a selon $\vec{u}$ est alors défini par
$$D_{\vec{u}} f(a) = f'_{a,\vec{u}}(0) = \lim_{t \to 0} \frac{f(a + t\vec{u}) - f(a)}{t} \in F.$$
3. Si f est dérivable selon $\vec{u}$ en tout point a de Ω, on peut définir la fonction $D_{\vec{u}} f : \Omega \to F$.

Remarque 16.2.4

1. Dans la suite, afin d'alléger les notations, on omettra la flèche sur u, en notant simplement $D_u f(a)$.
2. L'existence d'une dérivée selon u en a n'implique pas l'existence de dérivées selon d'autres vecteurs non colinéaires ; en revanche, si $v \neq 0$ est colinéaire à u, l'existence de $D_u f(a)$ équivaut à celle de $D_v f(a)$, et on peut relier les deux quantités.

Exemple 16.2.5

Étudier l'existence des dérivées en 0 selon $u \in \mathbb{R}^2$ de $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x, y) = \mathbb{1}_{\mathbb{R}\times\{0\}}(x, y)$$

Proposition 16.2.6 – Linéarité de D par rapport à u sur chaque droite

Soit $u \in E\backslash\{0\}$ et $f : \Omega \to F$ dérivable en $a \in \Omega$ selon u. Soit $\lambda \in \mathbb{R}^*$. Alors f est dérivable en a selon λu et

$$D_{\lambda u} f(a) = \lambda D_u f(a).$$

◁ **Éléments de preuve.**

Revenir à la définition par taux d'accroissement. ▷

Avertissement 16.2.7

L'existence de toutes les dérivées directionnelles en a n'est pas suffisante pour assurer la continuité en a

Exemple 16.2.8 – (figure 16.1)

Étudier les dérivées directionnelles en $0_{\mathbb{R}^2}$ et la continuité de $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x, y) = \begin{cases} \frac{x^2}{y} & \text{si } y \neq 0 \\ 0 & \text{si } y = 0. \end{cases}$$

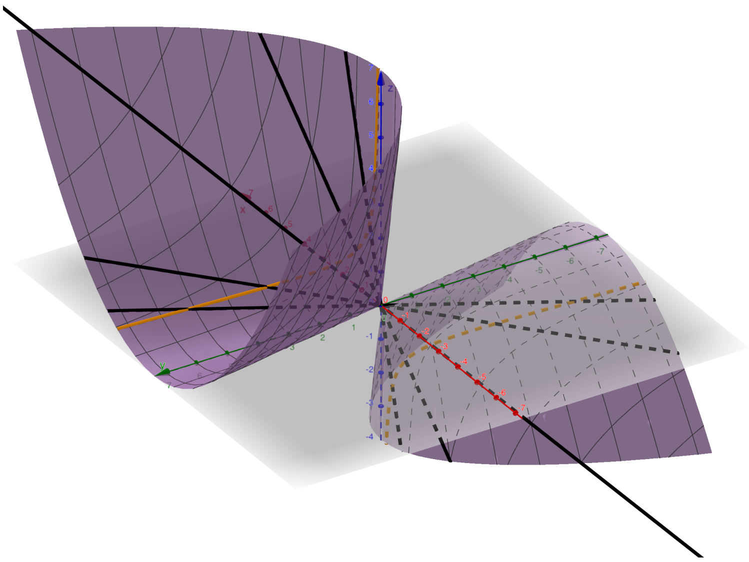

FIGURE 16.1 – Une fonction non continue en $(0, 0)$ mais y admettant des dérivées selon toute direction

Définition 16.2.9 – Dérivées partielles dans une base

Soit $f : \Omega \to F$, $a \in \Omega$ et $\mathcal{B} = (b_1, \dots, b_n)$ une base de E.

1. Soit $k \in [\![1, n]\!]$. La k-ième dérivée partielle de f dans la base $\mathcal{B}$ est la dérivée de f selon la direction b_k en a, et est notée $\partial_{\mathcal{B},k} f(a)$, ou, s'il n'y a pas d'ambiguïté sur la base utilisée,

simplement $\partial_k f(a)$. Ainsi

$$\partial_k f(a) = \partial_{\mathcal{B},k} f(a) = D_{b_k} f(a) = \lim_{t \to 0} \frac{f(a+tb_k) - f(a)}{t}.$$

2. Si f admet une k-ième dérivée partielle dans la base $\mathcal{B}$ en tout point de Ω, cela définit une application $\partial_k f : \Omega \to F$.

Remarque 16.2.10 – Identification à une fonction de plusieurs variables

Soit $f : \Omega \to F$ et $\mathcal{B} = (b_1, \dots, b_n)$ une base de E. Pour $x \in \Omega$, on note $[x]_{\mathcal{B}} = (x_1, \dots, x_n)$ ses coordonnées relativement à la base $\mathcal{B}$. On note

$$\tilde{\Omega} = \left\{ (x_1, \dots, x_n) \mid \sum_{k=1}^{n} x_k b_k \in \Omega \right\}.$$

Alors $\tilde{\Omega}$ est un ouvert de $\mathbb{R}^n$, et on peut identifier la fonction f à la fonction $\tilde{f} : \tilde{\Omega} \to F$ définie par

$$\tilde{f}(x_1, \dots, x_n) = f\left(\sum_{k=1}^{n} x_k b_k \right).$$

On obtient facilement, du fait de la possibilité de choisir la norme qu'on veut (en dimension finie) et donc notamment par exemple la norme $\| \cdot \|_{\mathcal{B},\infty}$ que :

- f est continue en x ssi $\tilde{f}$ est continue en $[x]_{\mathcal{B}}$;
- f admet en x une dérivée partielle selon b_k ssi $\tilde{f}$ admet en $[x]_{\mathcal{B}}$ une dérivée partielle selon e_k (vecteur de la base canonique), et alors

$$D_{b_k}(f)(a) = D_{e_k}(\tilde{f})(a).$$

Notation 16.2.11

Lorsque Ω est un ouvert de $\mathbb{R}^n$, on note aussi $\dfrac{\partial f}{\partial x_k}(a)$ la k-ième dérivée partielle par rapport à la base canonique. Ainsi, pour $a = (a_1, \dots, a_n)$, sous réserve d'existence,

$$\frac{\partial f}{\partial x_k}(a) = \lim_{t \to 0} \frac{f(a_1, \dots, a_k + t, \dots, a_n) - f(a_1, \dots, a_n)}{t}.$$

On l'appelle aussi dérivée partielle par rapport à la k-ième variable.

Proposition 16.2.12 – identification des dérivées partielles

Avec les notations de la remarque 16.2.10,

$$\frac{\partial \tilde{f}}{\partial x_k}(x_1, \dots, x_n) = D_{e_k}(\tilde{f})(x_1, \dots, x_n) = \partial_{\mathcal{B},k} f(x).$$

Notation 16.2.13

Après avoir fixé la base $\mathcal{B}$ et identifié f à $\tilde{f}$, on pourra donc aussi identifier la k-ième dérivée partielles par rapport à la base $\mathcal{B}$ à la dérivée partielle par rapport à la k-ième variable de $\tilde{f}$, et on s'autorisera à écrire, s'il n'y a pas d'ambiguïté sur la base :

$$\partial_k f(a) = \frac{\partial f}{\partial x_k}(a).$$

Par ailleurs, il arrive souvent, dans le cadre d'une application d'un ouvert de $\mathbb{R}^n$ dans $\mathbb{R}^p$ (situation à laquelle on se ramène par choix de bases), de travailler coordonnée par coordonnée. Dans ce cadre, on est ramené au calcul de dérivées partielles d'une fonction à valeurs dans $\mathbb{R}$:

Proposition 16.2.14 – Expression de la dérivée partielle d'une fonction à valeurs dans $\mathbb{R}^p$

Soit Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}^p$, qu'on écrit $f = (f_1, \dots, f_p)$. Ainsi, pour tout $x \in \Omega$, $f(x) = (f_1(x), \dots, f_p(x)) \in \mathbb{R}^p$, et pour tout $k \in [\![1, p]\!]$, $f_k : \Omega \to \mathbb{R}$.
Alors f admet une dérivée partielle par rapport à la k-ième variable en a ssi c'est le cas de chacun des f_i, et dans ce cas,

$$\frac{\partial f}{\partial x_k}(a) = \left(\frac{\partial f_1}{\partial x_k}(a), \dots, \frac{\partial f_p}{\partial x_k}(a)\right).$$

Par ailleurs, on rappelle comment, concrètement, calculer les dérivée partielles d'une fonction dont on connaît l'expression :

Méthode 16.2.15 – Calcul d'une dérivée partielle

Soit f une application de $\Omega \subset \mathbb{R}^n$ dans $\mathbb{R}$ (ce à quoi on peut se ramener d'après toutes les propriétés précédentes d'identification ou de travail par coordonnées). Pour calculer la dérivée partielle par rapport à x_k, on utilise les règles usuelles de calcul de la dérivée en utilisant l'expression de f, dans laquelle on dérive par rapport à la variable x_k, en considérant que toutes les autres variables sont des constantes.

Exemple 16.2.16

Calculer la dérivée partielle par rapport à x de $f : (x, y, z) \mapsto (xy \ln(x + \mathrm{e}^z), \sin(x\mathrm{e}^z)\cos(y))$.

Avertissement 16.2.17

L'existence des dérivées partielles par rapport à une base n'implique pas l'existence de dérivées selon toutes les directions en un point a.

Exemple 16.2.18 – figure 16.2

Étudier les dérivées directionnelles en $(0, 0)$ de $f : (x, y) \mapsto \sqrt[4]{|xy|}$

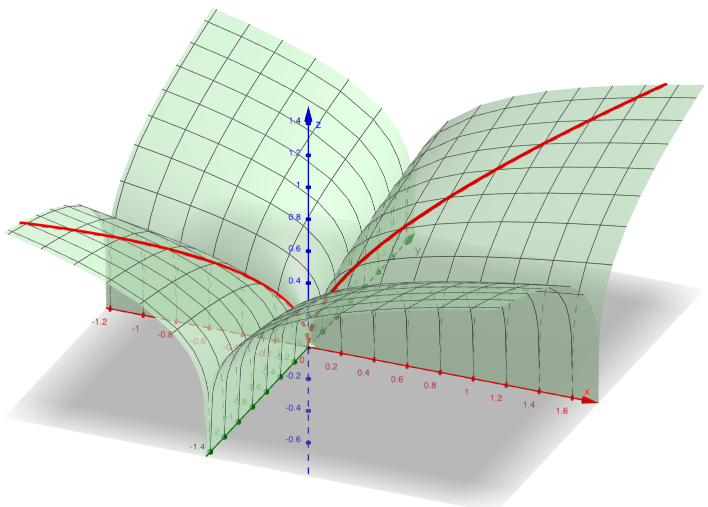

FIGURE 16.2 – Une fonction admettant des dérivées partielles en $(0, 0)$, mais pas de dérivée selon $(1, 1)$

Remarque 16.2.19

Cependant, sous des hypothèses supplémentaires de régularité, les dérivées partielles suffisent à retrouver les dérivées directionnelles. Par exemple, dans le cas des fonctions de 2 variables, vous avez montré l'année dernière que si f est de classe $\mathcal{C}^1$ au voisinage de a (dans le sens où ses deux dérivées partielles existent et sont continues au voisinage de a), alors, en notant $\nabla f(X) = \begin{pmatrix} \frac{\partial f}{\partial x}(X) \\ \frac{\partial f}{\partial x}(X) \end{pmatrix}$:

- pour tout $u = (u_x, u_y)$, $D_u f(a) = \langle \nabla f(a), u \rangle = u_x \frac{\partial f}{\partial x}(a) + u_y \frac{\partial f}{\partial y}(a)$,
- au voisinage de $a = (a_x, a_y)$, en notant $X = (x, y)$,

$$f(X) = f(a) + (x - a_x)\frac{\partial f}{\partial x}(a) + (y - a_y)\frac{\partial f}{\partial y}(a) + o(\|X - a\|) = f(a) + \langle X - a, \nabla f(a) \rangle + o(\|X - a\|)$$

(formule de Taylor-Young)

II.2 Différentielles

En particulier, d'après la remarque clôturant le paragraphe précédent, les fonctions de 2 variables suffisamment régulières se comportent à première approximation comme la fonction affine

$$X \mapsto f(a) + \langle X - a, \nabla f(a) \rangle,$$

donc la partie linéaire est l'application linéaire est

$$Y \mapsto \langle Y, \nabla f(a) \rangle = D_Y f(a).$$

Ainsi, sous certaines conditions de régularité, la dérivée directionnelle en a est une application linéaire en la direction, et permet de définir un développement limité à l'ordre 1.
C'est l'existence de cette approximation linéaire qui définit la notion de différentiabilité d'une application (tout comme la dérivabilité d'une fonction d'une variable équivaut à l'existence d'un DL_1).

Définition 16.2.20 – Différentiabilité et différentielle de f

Soit $f : \Omega \to F$, et $a \in \Omega$.

1. On dit que f est différentiable en a s'il existe une application linéaire $u \in \mathcal{L}(E, F)$ telle que au voisinage de a,
$$f(x) = f(a) + u(x - a) + o(\|x - a\|).$$
2. Dans ce cas, l'application linéaire u est notée $\mathrm{d}f(a)$, et est appelée différentielle de f au point a, ou application linéaire tangente à f en a.

Ainsi, la différentielle $\mathrm{d}f(a)$ est caractérisée par :

(i) sa linéarité ($\mathrm{d}f(a) \in \mathcal{L}(E, F)$)

(ii) le DL $f(a + h) \underset{0}{=} f(a) + [\mathrm{d}f(a)](h) + o(\|h\|)$

Pour alléger les notations, on écrira souvent $o(h)$ au lieu de $o(\|h\|)$ (et de même aux ordres supérieurs), et on omettra les crochets et parenthèses en écrivant

$$[\mathrm{d}f(a)](h) = \mathrm{d}f(a) \cdot h.$$

Remarque 16.2.21

Attention à bien comprendre les significations des différentes parenthèses. Si f est différentiable sur tout Ω par exemple,

$$\mathrm{d}f \; : \; \Omega \to \mathcal{L}(E, F)$$

est une fonction qui à tout point de Ω associe une application liénaire, qui est la partie linéaire de l'approximation affine de f en a. Pour tout a,

$$\mathrm{d}f(a)\ : E \to F$$

est donc une application linéaire, qu'on peut ensuite évaluer en $h \in E$. Nous verrons plus loin que, comme dans le cas d'une fonction de deux variables, $\mathrm{d}f(a) \cdot h$ est la dérivée directionnelle en a selon u.

Cette définition est à rapprocher de la caractérisation de la dérivabilité des fonctions d'une variable réelle par le DL_1. D'ailleurs, cette caractérisation permet d'en déduire le rapport entre dérivée et diférentielle dans le cas d'une fonction d'un intervalle I de $\mathbb{R}$ dans $\mathbb{R}$.

Proposition 16.2.22 – Différentielle d'une application d'une variable réelle

Soit Ω un ouvert de $\mathbb{R}$, $f : \Omega \to \mathbb{R}$ et $a \in \mathbb{R}$. Alors f est différentiable en a si et seulement si f est dérivable en a, et dans ce cas :

$$df(a) : \left|\begin{array}{ccc} \mathbb{R} & \longrightarrow & F \\ h & \longmapsto & f'(a) \cdot h \end{array}\right.$$

Cette proposition permet de voir la plupart des résultats suivants comme généralisation de la situation réelle bien connue.

Méthode 16.2.23 – Montrer que f est différentiable, méthode 1

Pour montrer que f est différentiable en a et calculer sa différentielle, on peut effectuer un DL_1 en a.

Exemples 16.2.24

1. Différentiabilité et différentielle en $(0, 2, 1)$ de la fonction définie sur $\mathbb{R}^3$ par $f(x, y, z) = \frac{\sin(x)-1}{1+x-y-z}$.
2. Différentiabilité en M et différentielle de $M \mapsto M^k$, définie sur $\mathcal{M}_n(\mathbb{R})$ à valeurs dans $\mathcal{M}_n(\mathbb{R})$
3. Différentiabilité et différentielle en 0 de $\exp : \mathcal{M}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R})$.

Proposition 16.2.25 – Différentiable implique continu

Soit $f : \Omega \to F$ une application différentiable en a. Alors f est continue en a.

◁ **Éléments de preuve.**
Par DL ! Remarquer que $\mathrm{d}f(a)$ est continue en 0 puisque linéaire en dimension finie. ▷

Proposition 16.2.26 – Différentielle et dérivées directionnelles

Soit $f : \Omega \to F$ une application différentiable en a. Alors f admet en a des dérivées selon toute direction $u \in E$, et

$$D_u f(a) = \mathrm{d}f(a) \cdot u.$$

En particulier, si $\mathcal{B} = (b_1, \dots, b_n)$ est une base de E, et si f est identifiée à une fonction des variables $x_1, \dots, x_n$ coordonnées dans la base $\mathcal{B}$,

$$\forall k \in [\![1, n]\!], \quad \frac{\partial f}{\partial x_k}(a) = \mathrm{d}(a) \cdot b_k.$$

◁ **Éléments de preuve.**
Encore par DL. ▷

Corollaire 16.2.27 – Matrice de la différentielle

Soit $f : \Omega \to F$ une application différentiable en a, et $\mathcal{B}$ et $\mathcal{C}$ deux bases de E et F respectivement. Soit $(f_1, \dots, f_p)$ les fonctions coordonnées de f, définies sur E à valeurs dans $\mathbb{K}$. Alors

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\mathrm{d}f(a)) = \left(\frac{\partial f_i}{\partial x_j}\right)_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant n}} \in \mathcal{M}_{p,n}(\mathbb{K}).$$

Corollaire 16.2.28 – Détermination de la différentielle par les dérivées partielles

Si f est différentiable en a, ses dérivées partielles selon une base $\mathcal{B}$ déterminent la différentielle. Plus précisément, pour tout $h \in E$ tel que $[h]_{\mathcal{B}} = (h_1, \dots, h_n)$,

$$\mathrm{d}f(a) \cdot h = \sum_{k=1}^{n} h_k \partial_k f = \sum_{k=1}^{n} h_k \frac{\partial f}{\partial x_k},$$

les dérivées partielles et les coordonnées étant prises relativement à la base $\mathcal{B}$.

Avertissement 16.2.29 – Dérivées directionnelles versus différentielle

L'existence des dérivées directionnelles ne suffit pas à obtenir la différentiabilité. L'hypothèse de différentiabilité de l'énoncé précédent est donc important.

En effet, l'exemple 16.1 définit une fonction admettant des dérivées directionnelles mais non continue en 0, donc non différentiable.

Définition 16.2.30 – Jacobienne d'une application de $\Omega \subset \mathbb{R}^n$ dans $\mathbb{R}^p$

Si $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$, la matrice de $\mathrm{d}f(a)$ relativement aux bases canoniques est appelée matrice jacobienne de f en a, et notée $J_f(a)$. Ainsi :

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1} & \cdots & \frac{\partial f_p}{\partial x_n} \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{K}),$$

où $f = (f_1, \dots, f_p)$.

Méthode 16.2.31 – Montrer que f est différentiable, méthode 2

- Dans le cas où $E = \mathbb{R}^n$, $F = \mathbb{R}^p$, étudier l'existence des dérivées partielles en a, et considérer l'application linéaire v canoniquement associée à la matrice jacobienne. Si

$$f(a+h) - f(a) - v(h) = o(h),$$

  f est différentiable en a de différentielle $\mathrm{d}f(a) = v$.
- Adaptation immédiate dans les autres situations après avoir fixé des bases.

Exemple 16.2.32

Étudier la différentiabilité en $(0,0)$ de $f : (x,y) \mapsto \frac{x^3}{x^2+y^4}$, prolongée par 0 en 0.
NB : en exercice, assurez-vous que trouver un DL_1 de façon directe n'est pas aisé du tout !

Proposition 16.2.33 – Différentielle d'une application constante

Soit $f : \Omega \to \mathbb{R}$ une application constante. Alors f est différentiable sur Ω, de différentielle nulle.

Autrement dit, pour tout $a \in \Omega$, $\mathrm{d}f(a)$ est l'application linéaire nulle de E dans F.

Proposition 16.2.34 – Différentielle d'une AL

Soit $f \in \mathcal{L}(E, F)$. Alors f est différentiable en tout point $a \in E$, et pour tout $a \in E$,

$$\mathrm{d}f(a) = f.$$

Ainsi, $\mathrm{d}f$ est constante de valeur f.

Les notations $\mathrm{d}x$, $\mathrm{d}y$, etc, très utilisée par les physiciens notamment, peuvent être définies rigoureusement comme des cas particuliers de la proposition précédente.

Définition 16.2.35 – Notation $\mathrm{d}x_i$

Étant donné un espace E de dimension n, muni d'une base $\mathcal{B} = (b_1, \dots, b_n)$, on note $\mathrm{d}x_i$ la différentielle de la projection

$$p_i : \sum_{j=1}^{n} x_j b_j \mapsto x_i.$$

Ainsi, en tout $a \in E$,

$$\mathrm{d}x_i(a) = p_i \qquad \text{i.e.} \qquad \forall h \in E,\ \mathrm{d}x_i(a) \cdot h = h_i,$$

où h_i est la coordonnée selon b_i de h sur la base $\mathcal{B}$.

Remarque 16.2.36 – Expression de la différentielle avec $\mathrm{d}x_i$

L'expression de $\mathrm{d}f(a)$ en fonction des dérivées partielles se réécrit donc :

$$\forall h \in E,\quad \mathrm{d}f(a) \cdot h = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \mathrm{d}x_i(a) \cdot h, \qquad \text{soit:} \qquad \mathrm{d}f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \mathrm{d}x_i(a),$$

ou encore, sur un domaine de différentiabilité de f :

$$\mathrm{d}f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \mathrm{d}x_i,$$

relation prisée des physiciens.

La notation $\mathrm{d}x_i$, bien qu'adhérente au programme et utile pour justifier les notations des physiciens, n'est pas explicitement au programme.

II.3 Gradient

On généralise dans cette section la notion de gradient introduite l'année dernière dans le cadre de fonctions de fonctions de deux variables réelles à valeurs dans $\mathbb{R}$. Dans ce contexte, le gradient a été défini en un point a comme le vecteur de $\mathbb{R}^2$ formé des deux dérivées partielles en a, autrement dit des deux dérivées selon les vecteurs d'une base.

On pourrait définir le gradient de la même manière, par rapport à une base, mais c'est plus commode de le définir sans référence à une base particulière. Pour cela, on se place dans le contexte où E est un espace euclidien et $F = \mathbb{R}$, et on utilise le lemme suivant, qu'on a déjà utilisé pour justifier l'existence de l'adjoint d'un endomorphisme.

Lemme 16.2.37 – Théorème de représentation de Riesz

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien, et $\varphi \in E^*$ une forme linéaire. Alors il existe un unique vecteur $a \in E$ tel que

$$\forall x \in E, \quad \varphi(x) = \langle a, x \rangle .$$

Proposition/Définition 16.2.38 – Gradient d'une application numérique

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien, et Ω un ouvert de E. Soit $f : \Omega \to \mathbb{R}$, et $a \in \Omega$ tel que f doit différentiable en a. Alors $\mathrm{d}f(a) \in E^*$. On définit alors le gradient de f en a comme étant l'unique vecteur $\nabla f(a) \in E$ tel que

$$\forall h \in E, \quad \mathrm{d}f(a) \cdot h = \langle \nabla f(a), h \rangle .$$

Théorème 16.2.39 – Expression du gradient en b.o.n.

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien, et Ω un ouvert de E. Soit $f : \Omega \to \mathbb{R}$, et $a \in \Omega$ tel que f doit différentiable en a. Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base orthonormale de E. Alors

$$\nabla f(a) = \sum_{k=1}^{n} \partial_k f(a) b_k,$$

les dérivées partielles étant prises dans la base $\mathcal{B}$. Ainsi :

$$[\nabla f(a)]_{\mathcal{B}} = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix}.$$

Corollaire 16.2.40 – Gradient d'une fonction de n variables réelles

On munit $\mathbb{R}^n$ (identifié à $\mathcal{M}_{n,1}(\mathbb{R})$) de sa structure euclidienne canonique. Soit Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}$ une application différentiable en $a \in \Omega$. Alors

$$\nabla f(a) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(a) \end{pmatrix}.$$

Remarque 16.2.41

Si $\Omega \subset \mathbb{R}^n$ et $f : \Omega \to \mathbb{R}$ est différentiable en a, alors la jacobienne en a est

$$J_f(a) = \mathrm{Mat}_{bc}(\mathrm{d}f(a)) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right) \in \mathcal{M}_{1,n}(\mathbb{R}).$$

Ainsi, $J_f(a) = \nabla f(a)^\top$.

Proposition 16.2.42 – Première interprétation géométrique du gradient

Soit $\Omega \subset \mathbb{R}^n$, et $f : \Omega \to \mathbb{R}$ différentiable en $a \in \Omega$. Le gradient $\nabla f(a)$ est positivement colinéaire au vecteur unitaire u tel que $D_u f(a)$ soit maximal :

$$\max_{\|u\|=1} D_u f(a) \iff u = \frac{\nabla f(a)}{\|\nabla f(a)\|}$$

◁ **Éléments de preuve.**

C'est l'inégalité de Cauchy-Schwarz et son cas d'égalité, le choix du signe étant dicté par le fait que $D_{-u}f(a) = -D_u f(a)$, par linéarité de $\mathrm{d}f(a)$. ▷

Ainsi, en tout point de différentiabilité, le gradient indique la direction de plus forte pente.

II.4 Opérations sur les applications différentiables

On se place à nouveau dans la situation générale : E est un e.v.n. de dimension finie sur $\mathbb{R}$, F est un e.v.n. de dimension finie de $\mathbb{K}$ et Ω est un ouvert de E.

Proposition 16.2.43 – Linéarité

Soit f et $g : \Omega \to F$ différentiables en a, et $\lambda, \mu \in \mathbb{K}$. Alors $f + \lambda g$ est différentiable en a et

$$\mathrm{d}(f + \lambda g)(a) = \mathrm{d}f(a) + \lambda \mathrm{d}g(a).$$

Définition 16.2.44

Revenir à la définition par DL.

Proposition 16.2.45 – Différentielle d'une forme multilinéaire

Soit Ω un ouvert de E, $F_1, \dots, F_n$ des $\mathbb{K}$-e.v.n. de dimension finie, et $f_i : \Omega \to F_i$ différentiables en $a \in \Omega$. Soit $M : F_1 \times \cdots \times F_n \to F$ une application multilinéaire. Alors

$$M(f_1, \dots, f_n) : x \mapsto M(f_1(x), \dots, f_n(x))$$

est différentiable en a, et

$$\forall h \in E, \quad \mathrm{d}M(f_1, \dots, f_n)(a) \cdot h = \sum_{k=1}^{n} M(f_1(a), \dots, \mathrm{d}f_k(a) \cdot h, \dots, f_n(a)).$$

◁ **Éléments de preuve.**

Par DL et majoration de la forme multilinéaire (par continuité) ▷

Exemple 16.2.46

1. Soit $f, g : \Omega \to \mathbb{K}$, $\Omega \subset E$. Exprimer la différentielle de fg.
2. Si $f : \Omega \to F$ avec F euclidien, et si f est différentiable en a, montrer que $\|f\|^2$ est différentiable en tout point et exprimer sa différentielle.
3. En particulier, si E est un espace euclidien, exprimer la différentielle en tout point de $X \mapsto \|X\|^2$.

Corollaire 16.2.47 – Dérivée de $M(f_1, \dots, f_p)$

Soit I un intervalle ouvert de $\mathbb{R}$, $F_1, \dots, F_n, F$ des $\mathbb{K}$-e.v.n. de dimension finie, et $f_i : I \to F_i$. Soit M une forme multilinéaire de $F_1 \times \cdots \times F_n$ dans F. On pose :

$$g : t \mapsto M(f_1(t), \dots, f_n(t)).$$

Soit $a \in I$. Si pour tout $i \in [\![1, n]\!]$, f_i est dérivable en a, alors g aussi, et

$$g'(a) = \sum_{k=1}^{n} M(f_1(a), \dots, f_i'(a), \dots, f_n(a)).$$

Remarque 16.2.48

C'est une généralisation de la dérivée d'un produit (le produit est bilinéaire !)

Corollaire 16.2.49

Voici 3 cas particuliers classiques, a

1. Si L est une application linéaire de F dans G, et $f : I \to F$ dérivable en a, alors $(L \circ f)$ aussi, et
$$(L \circ f)'(a) = L(f'(a)).$$
2. Soit F un espace euclidien, et $f, g : I \to E$ deux applications dérivables sur un intervalle I. Alors $\langle f, g\rangle$ aussi, et
$$\frac{\mathrm{d}}{\mathrm{d}t}\left\langle f(t), g(t)\right\rangle = \left\langle f'(t), g(t)\right\rangle + \left\langle f(t), g'(t)\right\rangle.$$
3. Soit E un espace vectoriel de dimension n, et $\mathcal{B}$ une base de E. Alors, si $f_1, \dots, f_n$ sont à valeurs dans E et dérivables,
$$\frac{\mathrm{d}}{\mathrm{d}t}\det_{\mathcal{B}}(f_1, \dots, f_n) = \sum_{k=1}^{n} \det_{\mathcal{B}}(f_1, \dots, f_k', \dots, f_n).$$

Exemples 16.2.50

1. Exprimer la dérivée de $t \mapsto A(t)B(t)$, où
$$A : t \mapsto \mathcal{M}_{n,p}(\mathbb{C}) \qquad \text{et} \qquad B : t \mapsto \mathcal{M}_{p,q}(\mathbb{C})$$
sont deux applications dérivables sur I.
Par exemple, exprimer la dérivée de $\begin{pmatrix} 2 & 4 & 1+t \\ 3 & 2+2t & 2 \\ 1 & t & t \end{pmatrix}\begin{pmatrix} 2t & 4 & -t \\ 0 & 1 & 1-t \\ 2 & 1 & t \end{pmatrix}$.
2. Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien, et $X : I \to E\backslash\{0\}$ dérivable sur I. Montrer que $t \mapsto \|X(t)\|$ est dérivable sur I, et exprimer sa dérivée.
3. Dans le même contexte, montrer que $t \mapsto \frac{X(t)}{\|X(t)\|}$ est dérivable, et exprimer sa dérivée.
4. Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$, $X : I \mapsto \mathcal{M}_{n,1}(\mathbb{K})$ et $Y : I \mapsto \mathcal{M}_{p,1}(\mathbb{K})$. On suppose X et Y dérivables. Exprimer la dérivée de $X(t)^\top AY(t)$.
5. Dérivée de $t \mapsto X(t)^n$, lorsque $X : I \to \mathcal{M}_n(\mathbb{K})$.
6. Exprimer la dérivée de l'application qui à $t \in \mathbb{R}$ associe $\begin{vmatrix} t-2 & 3 & 6 \\ 0 & t+1 & 2 \\ 2 & 5 & t-3 \end{vmatrix}$

Proposition 16.2.51 – Différentielle d'une composée

Soit E, et F deux e.v.n. de dimension finie sur $\mathbb{R}$, et G un e.v.n. de dimension finie sur $\mathbb{K}$. Soit Ω un ouvert de E et Ω' un ouvert de f, et $f : \Omega \to F$ et $g : \Omega' \to G$ deux applications. Soit $a \in \Omega$. On suppose que

(i) f est différentiable en a

(ii) $f(a) \in \Omega'$ et g est différentiable en $f(a)$.

Alors $g \circ f$ est différentiable en a, et

$$\mathrm{d}(g \circ f)(a) = \mathrm{d}g(f(a)) \circ \mathrm{d}f(a).$$

◁ **Éléments de preuve.**

Composer les DL. ▻

Corollaire 16.2.52 – Composition par une fonction réelle

Soit Ω un ouvert d'un e.v.n. E, et I un intervalle ouvert de $\mathbb{R}$. Soit $f : \Omega \to \mathbb{R}$ et $g : I \to \mathbb{R}$. Si f est différentiable en $a \in \Omega$, si $f(a) \in I$ et si g dérivable en $f(a)$, alors $g \circ f$ est différentiable en a et

$$\mathrm{d}(g \circ f)(a) = g'(f(a))\mathrm{d}f(a).$$

Exemples 16.2.53

1. Différentiabilité de $x \mapsto \|x\|$ en tout $x \neq 0$ dans un espace euclidien.
2. Différentiabilité de $\|f\|$ lorsque f l'est.

Corollaire 16.2.54 – Dérivée le long d'un arc

Soit I un intervalle ouvert de $\mathbb{R}$, Ω un ouvert de E. Soit $\gamma : I \to \Omega$ et $f : \Omega \to F$. Si γ est dérivable en $a \in I$, et f différentiable en $\gamma(a)$, alors $f \circ \gamma$ est dérivable en a et

$$(f \circ \gamma)'(a) = \mathrm{d}f(a) \cdot \gamma'(a).$$

Si E est euclidien, cela se réécrit aussi

$$(f \circ \gamma)'(a) = \left\langle \nabla f(a), \gamma'(a) \right\rangle.$$

Exemples 16.2.55

Dérivée de $t \mapsto f(a + th)$. Cohérence en 0.

Corollaire 16.2.56 – Règle de la chaîne

Avec les notations et hypothèses de la proposition 16.2.51, si $\mathcal{B} = (b_1, \dots, b_n)$ est une base de E, et $\mathcal{C} = (c_1, \dots, c_p)$ une base de F,

$$\forall k \in [\![1, n]\!], \partial_{\mathcal{B},k}(g \circ f)(a) = \sum_{i=1}^{p} \partial_{\mathcal{C},i}\, g(f(a))\partial_{\mathcal{B},k} f(a).$$

Avertissement 16.2.57

Attention à ne pas confondre $\partial_i g(f(a))$ et $\partial_i(g \circ f)(a)$.. Quelle est la différence entre les deux ?

Corollaire 16.2.58 – Reexpression de la règle de la chaîne dans $\mathbb{R}^n$

Si $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$, et si les bases considérées sont les bases canoniques, on obtient, avec les mêmes hypothèses :

$$\forall k \in [\![1, n]\!], \frac{\partial(g \circ f)}{\partial x_k}(a) = \sum_{i=1}^{p} \frac{\partial g}{\partial y_i}(f(a))\frac{\partial f}{\partial x_k}(a),$$

où $(x_1, \dots, x_n)$ représentent les coordonnées génériques dans $E = \mathbb{R}^n$ et $(y_1, \dots, y_p)$ les coordonnées génériques dans $F = \mathbb{R}^p$.

De façon équivalente, en exprimant coordonnée par coordonnée, $J_{g\circ f}(a) = J_g(f(a)) \times J_f(a)$.

Exemple 16.2.59

Soit f une application de $\mathbb{R}^3$ dans $\mathbb{R}$ admettant des dérivées partielles en tout (x, y, z). Exprimer les dérivées partielles de

$$g(u, v, w) = f(u^2 + v, 2v, 3w - e^v).$$

Proposition 16.2.60 – Différentielle de $(f_1, \dots, f_p)$

Soit Ω un ouvert de E, et $F_1, \dots, F_q$ des $\mathbb{K}$-e.v.n. de dimension finie, et soit $f_i : \Omega \to F_i$, $i \in [\![1, q]\!]$. Les propositions suivantes sont équivalentes:

(i) Pour tout $i \in [\![1, q]\!]$, f_i est différentiable en $a \in \Omega$

(ii) $f = (f_1, \dots, f_q) : E \to F_1 \times \cdots \times F_q$ est différentiable en a.

Si c'est le cas,

$$\mathrm{d}f(a) = (\mathrm{d}f_1(a), \dots, \mathrm{d}(f_q)(a)) \in \mathcal{L}(E, F_1 \times \cdots \times F_q),$$

c'est-à-dire :

$$\forall h \in E, \quad \mathrm{d}f(a) = (\mathrm{d}f_1(a) \cdot h, \dots, \mathrm{d}f_q(a) \cdot h).$$

Remarque 16.2.61

Retrouver de la sorte l'expression de la jacobienne.

III Applications de classe $\mathcal{C}^n$

III.1 Applications de classe $\mathcal{C}^1$

Définition 16.3.1 – Application de classe $\mathcal{C}^1$

Soit $f : \Omega \to F$. On dit que f est de classe $\mathcal{C}^1$ sur Ω si :

(i) f est différentiable en tout point de Ω

(ii) $\mathrm{d} : \Omega \to \mathcal{L}(E, F)$ est continue

Exemples 16.3.2

1. Soit $f : E \to F$ une application linéaire, alors f est de classe $\mathcal{C}^1$.
2. Soit f une application constante, alors f est de classe $\mathcal{C}^1$.

Proposition 16.3.3 – Opérations sur les applications de classe $\mathcal{C}^1$

Soit E, E' des $\mathbb{R}$-e.v.n. de dimension finie, et F, $F_1, \dots, F_q$ des $\mathbb{K}$-e.v.n. de dimension finie. Soit Ω, Ω' des ouverts de E et E' respectivement.

1. Soit $f, g : \Omega \to F$ de classe $\mathcal{C}^1$ et $\lambda \in \mathbb{K}$. Alors $f + \lambda g$ est de classe $\mathcal{C}^1$.
2. Soit pour $i \in [\![1, q]\!]$, $f_i : \Omega \to F_i$, et $M : F_1 \times \cdots \times F_q \to F$ une forme multilinéaire. Si les f_i sont de classe $\mathcal{C}^1$, alors $M(f_1, \dots, f_q)$ est de classe $\mathcal{C}^1$.
3. Soit $f : \Omega' \to \Omega$ et $g : \Omega \to F$. Si f et g sont de classe $\mathcal{C}^1$ sur Ω' et Ω respectivement, alors $g \circ f$ est de classe $\mathcal{C}^1$ sur Ω.
4. Si $f : \Omega \to \mathbb{K}$ ne s'annule pas et est de classe $\mathcal{C}^1$, alors $\frac{1}{f}$ aussi.

5. Soit pour $i \in [\![1, q]\!]$, $f_i : \Omega \to F_i$, et $f : \Omega \to F_1 \times \cdots \times F_q$ définie par

$$\forall x \in \Omega, \quad f(x) = (f_1(x), \dots, f_q(x)).$$

Alors f est de classe $\mathcal{C}^1$ si et seulement si tout f_i est de classe $\mathcal{C}^1$.

◁ **Éléments de preuve.**

Attention à bien comprendre sur quelle variable porte la continuité (le point $x \in \Omega$ en lequel on considère la différentielle). Cela conditionne la validité de l'argument donné. Par exemple, pour la composée, la différentielle $\mathrm{d}(g \circ f)$ s'exprime comme composée de deux différentielles, mais la composée porte sur la variable h, et non sur a qui est à voir plutôt comme un paramètre. Donc on ne peut pas obtenir la continuité de cette différentielle par propriété de continuité d'une composée.

En revanche, puisque la composition se fait entre applications linéaires d'espaces de dimension finie, l'application $(f, g) \mapsto g \circ f$ est une application bilinéaire entre espaces de dimension finie, donc elle-même continue. C'est ce point qui est en jeu ici.

Pour 2, on peut remarquer que si $M : E^n \to F$ est une application multilinéaire, alors $\widetilde{M} : E^{k-1} \times \mathcal{L}(E, F) \times E^{n-k} \to \mathcal{L}(E, F)$ définie par

$$\widetilde{M}(x_1, \dots, x_{k-1}, \alpha, x_{k+1}, \dots, x_n)(x) = M(x_1, \dots x_{k-1}, \alpha(x), x_{k+1}, \dots, x_n)$$

est également multilinéaire, donc continue, tous les espaces en jeu étant de dimension finie. ▷

Vous avez défini l'année dernière le caratère $\mathcal{C}^1$ des fonctions de 2 variables à l'aide des dérivées partielles. On retrouve ce point de vue via la caractérisation 16.3.5. On démontre d'abord le résultat suivant, qui n'apparaît qu'implicitement au programme (cela peut être vu comme un sous-produit de la caractérisation 16.3.5 et de la définition de la différentielle)

Théorème 16.3.4 – Formule de Taylor-Young à l'ordre 1 pour les dérivées partielles

Soit $f : \Omega \to \mathbb{K}$ une application à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. Soit $\mathcal{B}$ une base de E, et $a \in \Omega$. On suppose que pour tout f, $\partial_k f$ est définie et continue au voisinage de a. Alors

$$f(a + h) \underset{h \to 0}{=} f(a) + \sum_{k=0}^{n} h_k \partial_k f(a) + o(h).$$

où $(h_1, \dots, h_n) = [h]_{\mathcal{B}}$

◁ **Éléments de preuve.**

Utiliser la formule de Taylor pour les fonctions d'une variable, successivement pour chacune des n variables. ▷

Théorème 16.3.5 – Caractérisation de la classe $\mathcal{C}^1$ par les dérivées partielles

Soit $f : \Omega \to F$ et $\mathcal{B} = (b_1, \dots, b_n)$ une base de E. Les propositions suivantes sont équivalentes:

(i) f est de classe $\mathcal{C}^1$ sur Ω ;

(ii) f admet des dérivées partielles en tout point selon la base $\mathcal{B}$, et pour tout $k \in [\![1, n]\!]$, $\partial_k f : \Omega \to F$ est continue.

◁ **Éléments de preuve.**

Démonstration non exigible.

- Se ramener à $f_i : \Omega \to \mathbb{K}$, en considérant les coordonnées dans une base $\mathcal{C}$ de F.
- Utiliser la décomposition $\mathrm{d}f_i(a) = \sum_{k=1}^{n} \partial_k f_i$.

Cela revient en gros à utiliser le fait que $\mathrm{d}f$ est continue si et seulement si $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\mathrm{d}f)$ est continue, ce qui est la matrice des dérivées partielles des coordonnées dans $\mathcal{C}$ (l'analogue de la jacobienne)

Dans le sens $(ii) \Longrightarrow (i)$, ne pas oublier de montrer d'abord la différentiabilité de f, en utilisant la formule de Taylor-Young. ▻

Remarque 16.3.6

La formule de Taylor 16.3.4 se réécrit alors simplement

$$f(a+h) = f(a) + \langle \nabla f(a), h\rangle + o(h),$$

qui n'est rien de plus que la combinaison de la définition de la différentielle et du gradient ∇f.

Corollaire 16.3.7 – Continuité via les dérivées partielles

Si $E = \mathbb{R}^n$ et si f admet des dérivées partielles par rapport à tout x_i en a, et que ces dérivées partielles sont continues au voisinage de a, alors f est continue au voisinage de a.

Avertissement 16.3.8

Comme on l'a déjà vu, l'existence seule des dérivées partielles (et même plus généralement des dérivées directionnelles) ne suffit pas à établir la continuité.

Théorème 16.3.9 – Intégrale de la différentielle le long d'un arc

Soit $f : \Omega \to F$ une application de classe $\mathcal{C}^1$, et $\gamma : [0,1] \to \Omega$ un chemin de classe $\mathcal{C}^1$ tel que $\gamma(0) = a$ et $\gamma(1) = b$. Alors

$$f(b) - f(a) = \int_0^1 \mathrm{d}f(\gamma(t))\gamma'(t)\ \mathrm{d}t.$$

◅ **Éléments de preuve.**

L'intégrande est la dérivée de $f \circ \gamma : [0,1] \to F$. ▻

Remarque 16.3.10

Ce théorème peut être vu comme une généralisation du théorème fondamentale de l'analyse, puisqu'en prenant $E = \mathbb{R}$ et $\gamma : t \mapsto a(1-t) + bt$, puis en faisant un changement de variable affine, on retrouve

$$\int_a^b f(t)\ \mathrm{d}t = f(b) - f(a)$$

lorsque $f : [a,b] \to F$ est de classe $\mathcal{C}^1$.

Exemple 16.3.11

1. Si f est de classe $\mathcal{C}^1$ sur l'ouvert Ω et si $[a,b] \subset \Omega$, alors

$$f(b) - f(a) = \int_0^1 \mathrm{d}f(a(1-t) + bt) \cdot (b-a)\ \mathrm{d}t$$

2. De façon équivalente, si f est de classe $\mathcal{C}^1$ sur l'ouvert Ω et si $[a, a+u] \subset \Omega$, alors

$$f(a+u) - f(a) = \int_0^1 \mathrm{d}f(a+tu) \cdot u\ \mathrm{d}t.$$

3. Si f est de classe $\mathcal{C}^1$ sur l'ouvert Ω et si $\gamma : [0,1] \to \Omega$ est un lacet de classe $\mathcal{C}^1$ (*i.e.*

$\gamma(a) = \gamma(b)$), alors
$$\int_0^1 \mathrm{d}f(\gamma(t)) \cdot \gamma'(t)\ \mathrm{d}t = 0.$$

Théorème 16.3.12 – Caractérisation des fonctions constantes (1)

Soit Ω un ouvert convexe de E, et $f : \Omega \to F$. Les propositions suivantes sont équivalentes:

(i) f est constante sur Ω

(ii) f est différentiable et $\mathrm{d}f$ est nulle sur Ω.

◁ **Éléments de preuve.**

- Le sens direct est déjà vu.
- Pour le sens réciproque, utiliser le théorème 16.3.9 sur une paramétrisation du segment $[a, b]$ pour montrer que pour tout $a, b \in \Omega$, $f(a) = f(b)$.

▷

Ce théorème admet une généralisation qui est aussi au programme, mais sans sa démonstration (que nous esquissons quand même).

Théorème 16.3.13 – Caractérisation des fonctions constantes (2)

Soit Ω un ouvert connexe par arcs de E, et $f : \Omega \to F$. Les propositions suivantes sont équivalentes:

(i) f est constante sur Ω

(ii) f est différentiable et $\mathrm{d}f$ est nulle sur Ω.

◁ **Éléments de preuve.**

Démonstration hors-programme.

C'est toujours la réciproque qui reste à voir.

On pourrait essayer d'adapter la preuve précédente, en montrant que si Ω est ouvert et connexe par arcs, on peut toujours joindre deux points par un arc de classe $\mathcal{C}^1$. Mais cette construction n'est pas si évidente que cela à faire.

On procède donc autrement, en « grignotant » petit à petit.

Soit $a \in \Omega$, et $\Omega' \subset \Omega$ l'ensemble des points $x \in \Omega$ tels que $f(a) = f(x)$. Si $\Omega' \neq \Omega$, considérer $b \in \Omega \setminus \Omega'$ et joindre a à b. Montrer l'existence d'un point c sur cet arc, étant sur la frontière de Ω'. Trouver une contradiction en utilisant la caractérisation (1) et la convexité d'une petite boule centrée en c.
▷

III.2 Plan tangent et noyau de la différentielle

Définition 16.3.14 – Vecteur tangent à une partie, figure 16.3

Soit $X \subset E$ une partie de E, et $x \in X$.

1. Un vecteur $v \in E$ est tangent à X en x s'il existe $\varepsilon > 0$ et un arc $\gamma :]-\varepsilon, \varepsilon[\to X$ dérivable en 0 tel que $\gamma'(0) = v$.
2. On note $T_x X$ l'ensemble des vecteurs tangents à X en x.

Exemples 16.3.15

1. Quels sont les vecteurs tangents à $\varnothing$?
2. 0 est tangent à tout X non vide.

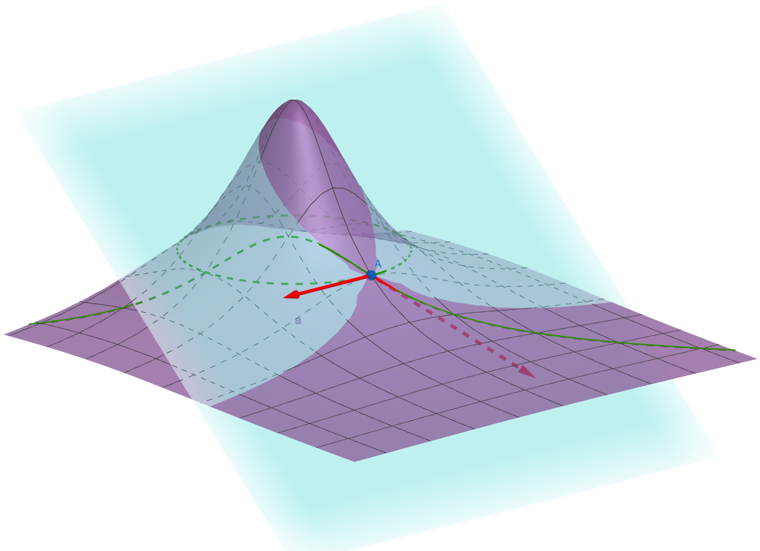

Figure 16.3 – Un sous-ensemble de $\mathbb{R}^3$ et deux vecteurs tangents en A.

3. Décrire $T_x X$ lorsque $x \in \mathring{X}$.
4. Décrire les vecteurs tangents à $\mathbb{R} \times \mathbb{R}_+$ en $(0, 0)$.

D'autres exemples importants sont mis en résultats du cours :

Proposition 16.3.16 – Vecteurs tangents à un sous-espace affine

Soit X un sous-espace affine de E de direction V. Alors, pour tout $B \in X$, $T_B(X) = V$.

◁ **Éléments de preuve.**

L'inclusion directe est évidente en paramétrant des droites incluses dans X passant par B.

Réciproquement la dérivée de γ est limite d'un taux d'accroissement qui est dans V. Comme on est en dimension finie, V est fermé. ▷

Proposition 16.3.17 – Vecteurs tangents à une sphère, figure 16.4

Soit E un espace euclidien, $c \in E$ et $r > 0$. Soit $X = S(c, r)$ et $x \in S(c, r)$. Alors $T_x X = \overrightarrow{cx}^{\perp}$.
En d'autres termes, l'ensemble des vecteurs tangents à une sphère en un point x est l'orthogonal du rayon défini par x.

◁ **Éléments de preuve.**

Considérer les grands cercles passant par x.

Pour l'autre inclusion, vérifier que si γ est à valeurs dans S, sa dérivée est orthogonale au rayon (dériver un produit scalaire). ▷

Exemple 16.3.18

Décrire sous forme d'une équation le plan vectoriel $T_A(S)$ où $S = S(0, 1) \subset \mathbb{R}^3$ et $A = (a, b, c) \in S$.

Exemple 16.3.19 – Vecteur tangent au graphe d'une fonction

Soit $f : \Omega \to \mathbb{R}$ une application numérique. Son graphe $\mathcal{G}$ est un sous-ensemble de l'e.v.n. $\mathrm{E} \times \mathbb{R}$. Soit $x_0 \in \Omega$, on peut donc définir l'ensemble des vecteurs tangents au graphe de f au point $(x_0, f(x_0)) \in \mathcal{G}$. La figure 16.3 est en fait (à condition de rajouter les axes au bon endroit) le graphe de la fonction

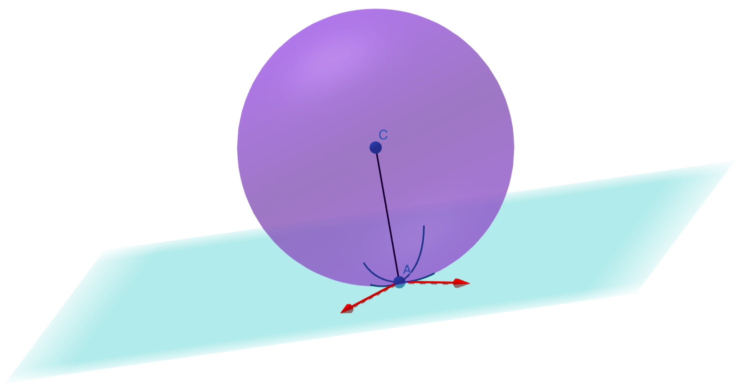

FIGURE 16.4 – Plan tangent à une sphère

$(x, y) \mapsto \dfrac{2}{1 + 2x^2 + 2y^2}$, le point de tangence étudié étant en $(1, 0)$.

Théorème 16.3.20 – Description d'un espace tangent à un graphe

Soit Ω un ouvert de $\mathbb{R}^2$, $f : \Omega \to \mathbb{R}$, et $X_0 \in \Omega$. On suppose f différentiable en X_0. L'ensemble $T_{X_0}f$ des vecteurs tangents à la courbe de f en $(X_0, f(X_0))$ est le plan d'équation

$$h\frac{\partial f}{\partial x}(X_0) + k\frac{\partial f}{\partial y}(X_0) = \ell,$$

c'est-à-dire l'ensemble des vecteurs $H = (h, k, \ell) \in \mathbb{R}^3$ vérifiant l'équation ci-dessus.

◁ **Éléments de preuve.**
Comme dans les autres exemples, étudier les propriétés des arcs du graphe, et réciproquement, trouver des arcs adéquats. ▷

Dans le cas général ($\Omega \subset E$, E de dimension finie), on obtient la description suivante :

Théorème 16.3.21 – Espace tangent à un ensemble défini par une équation

Soit $f : \Omega \to \mathbb{R}$ une fonction numérique de classe $\mathcal{C}^1$, $X = f^{-1}(\{0\})$. Ainsi X est l'ensemble défini par l'équation $f(x) = 0$.
Soit $a \in X$. Si $\mathrm{d}f(a) \neq 0$, l'ensemble des vecteurs tangents à X en a est égal à

$$T_aX = \mathrm{Ker}(\mathrm{d}f(a)).$$

◁ **Éléments de preuve.**
Hors-programme. ▷

Exemple 16.3.22

1. Retrouver à l'aide de ce théorème la description des vecteurs tangents à un espace affine de $\mathbb{R}^n$, décrit par une équation
$$a_1x_1 + \cdots + a_nx_n = b.$$
2. De même, retrouver la description des vecteurs tangents de la sphère unité de l'espace euclidien $\mathbb{R}^n$.
3. Retrouver la description de l'espace des vecteurs tangents au graphe d'une fonction $f : \Omega \to \mathbb{R}$ où $\Omega \subset \mathbb{R}^2$, lorsque f est de classe $\mathcal{C}^1$ au voisinage de X_0. Généraliser pour $\Omega \subset \mathbb{R}^n$.

Proposition 16.3.23 – Deuxième interprétation géométrique du gradient

Avec les hypothèse du théorème 16.3.21, si on suppose de plus que E est un espace euclidien, $T_aX = \nabla f(a)^{\perp}$.

◁ **Éléments de preuve.**

Dans le contexte euclidien, cette formulation est équivalente à cette du théorème, par définition du gradient.

On peut le démontrer indépendamment du théorème, en considérant un arc à valeurs dans X, et en dérivant $f \circ \gamma$. ▷

Exemple 16.3.24

Retrouver l'équation des hyperplans des vecteurs tangents à la sphère unité de $\mathbb{R}^n$.

Remarque 16.3.25

1. Si $E = \mathbb{R}^2$, X est la ligne de niveau de valeur 0. Plus généralement, en considérant $f - b$ au lieu de f, on peut décrire de la sorte toutes les lignes de niveau. La proposition affirme que le gradient en a est orthogonal à la droite tangente à la ligne de niveau en a, ce qu'on peut réexprimer de façon plus abrégée en disant que le gradient est orthogonal aux lignes de niveau.
2. Ainsi, si f est de classe $\mathcal{C}^1$ sur un ouvert de $\mathbb{R}^2$, en combinant les deux interprétations, la direction de plus forte pente au point a est toujours orthogonale à la ligne de niveau en a.

Exemple 16.3.26 – Figure 16.5

Soit $f : \mathbb{R}^2 \backslash \{(0,0)\} \to \mathbb{R}$ définie par

$$f(x,y) = \frac{x^2+y^3}{x^4+y^2}.$$

En exprimant le gradient, déterminer la tangente à la courbe de niveau en $(1,1)$. La figure 16.5 illustre cet exemple. Le gradient en $(1,1)$ est le vecteur $\overrightarrow{AB}$, reporté sur le plan $z = 1$, dans lequel est tracée en rouge la courbe de niveau au point $(1,1)$.

III.3 Applications de classe $\mathcal{C}^k$

Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de E, $(i,j) \in [\![1,n]\!]$, et $f : \Omega \to F$ admettant une dérivée partielle $\partial_i f$ sur tout Ω. Si la fonction $\partial_i f : \Omega \to F$ admet elle-même une j-ième dérivée partielle en $a \in \Omega$, on peut donc définir $\partial_j \partial_i f(a)$, la j-ième dérivée partielle de $\partial_i f$. De façon plus générale, on définit par récurrence sur k les dérivées d'ordre k :

Définition 16.3.27 – Dérivées partielles d'ordre k

On définit, si c'est possible, les dérivées d'ordre k de f par récurrence sur k :

- Si $k = 1$, il s'agit des dérivées partielles $\partial_i f$
- Soit $k > 1$ et $(i_1, \dots, i_k) \in [\![1,n]\!]^k$. La fonction f admet en $a \in \Omega$ une dérivée par rapport successivement aux $i_1, \dots, i_{k-1}$ et i_k-ièmes vecteurs de la base $\mathcal{B}$, notée $\partial_{i_k}\partial_{i_{k-1}} \dots \partial_{i_1} f(a)$, si et seulement si $\partial_{i_{k-1}} \dots \partial_{i_1} f(a)$ existe au voisinage (relatif dans Ω) de a et est dérivable selon b_{i_k} en a. Ainsi,

$$\partial_{i_k}\partial_{i_{k-1}} \dots \partial_{i_1} f(a) = \partial_{i_k}\left(\partial_{i_{k-1}} \dots \partial_{i_1} f\right)(a).$$

On trouve aussi les notations suivantes :

- $\partial_{i_1,\dots,i_k} f(a)$ (attention à l'ordre des indices !),

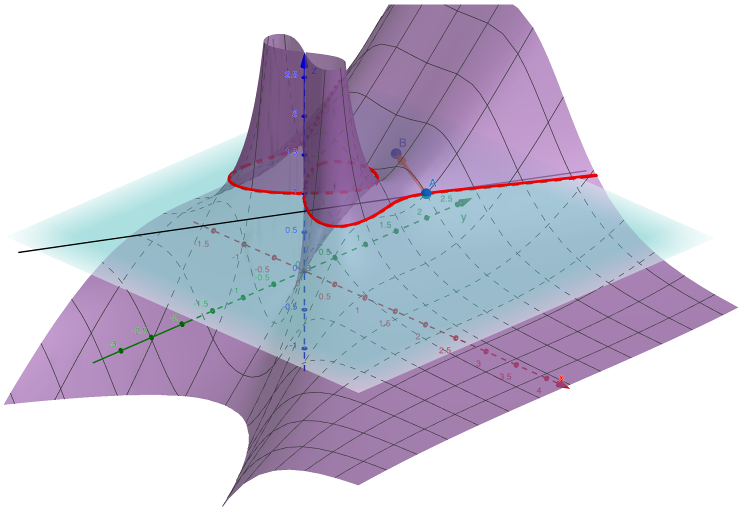

FIGURE 16.5 – Courbe de niveau $\dfrac{x^2+y^3}{x^4+y^2} = 1$, tangente et gradient en $A = (1,1,1)$.

- $\dfrac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}(a)$, si $(x_1, \dots, x_n)$ représentent les coordonnées génériques dans la base $\mathcal{B}$, en particulier si $E = \mathbb{R}^n$ muni de la base canonique.

On peut définir des différentielles d'ordre supérieur, mais cette notion est hors programme, et les objets manipulés deviennent de plus en plus complexes. Pour cette raison, nous définissons la classe $\mathcal{C}^k$ à l'aide des dérivées partielles d'ordre k. Cette définition est une généralisation de la caractérisation de la classe $\mathcal{C}^1$ par les dérivées partielles premières.

Définition 16.3.28 – Applications de classe $\mathcal{C}^k$

Soit $f : \Omega \to F$ une application définie sur l'ouvert Ω de E.

1. On dit que f est de classe $\mathcal{C}^k$ sur Ω si toutes les dérivées d'ordre k de f existent (il y en a 2^k), et sont continues sur Ω.
2. On dit que f est de classe $\mathcal{C}^k$ au voisinage de $a \in \Omega$, s'il existe un ouvert $V \subset \Omega$ tel que $f_{|V}$ soit de classe $\mathcal{C}^k$ sur V.
3. On note $\mathcal{C}^k(\Omega, F)$ l'ensemble des fonctions de classe $\mathcal{C}^k$ sur Ω, à valeurs dans F. Par convention, $\mathcal{C}^0(\Omega, F)$ est l'ensemble des fonctions continues.
4. On dit que f est de classe $\mathcal{C}^\infty$ sur Ω si elle est de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$.
5. On note $\mathcal{C}^\infty(\Omega, F)$ l'ensemble des fonctions de classe $\mathcal{C}^\infty$ sur Ω, à valeurs dans F.

Remarque 16.3.29

Si f est de classe $\mathcal{C}^k$, alors :
- f est de classe $\mathcal{C}^j$ pour tout $j \leqslant k$;
- plus généralement, si $\ell \leqslant k$ et $(i_1, \dots, i_\ell) \in [\![0, n]\!]^\ell$, $\partial_{i_1,\dots,i_\ell} f$ est de classe $\mathcal{C}^{k-\ell}$.

Pour étudier la classe $\mathcal{C}^k$, il y a donc *a priori* 2^k dérivées à étudier. Mais en pratique, cela en fait moins, car sous des conditions assez peu restrictives, on va retrouver plusieurs fois la même, dans le sens où l'ordre de dérivation (*i.e.* le choix de l'ordre des variables par rapport auxquelles on dérive) n'importe pas. C'est ce que dit le théorème suivant.

Théorème 16.3.30 – Théorème de Schwarz

Soit $f : \Omega \to F$, $\mathcal{B} = (b_1, \dots, b_n)$ une base de E, et $a \in \Omega$. Si f est de classe $\mathcal{C}^2$ au voisinage de a, alors

$$\forall (i,j) \in [\![1,n]\!], \quad \partial_i \partial_j f(a) = \partial_j \partial_i f(a).$$

◁ **Éléments de preuve.**

Démonstration non exigible.

En fixant toutes les autres variables, on est ramené au cas de fonctions de 2 variables. Exprimer un DL_2 de $f(x+h,y+k)-f(x,y)$ en $H=(h,k)$, de deux manières différentes, en intercalant $f(x+h,y)$, ou $f(x,y+k)$, et en se ramenant à la formule de Taylor-Young à l'ordre 2 pour les fonctions d'une variable (en fixant l'autre). Identifier les deux DL obtenus (justifier proprement cette identification). ▷

Corollaire 16.3.31 – Théorème de Schwarz pour la classe $\mathcal{C}^k$

Soit $f : \Omega \to F$, $\mathcal{B} = (b_1, \dots, b_n)$ une base de E, et $a \in \Omega$. Si f est de classe $\mathcal{C}^2$ au voisinage de a, alors pour tout $(i_1, \dots, i_k) \in [\![1,n]\!]^k$, et toute permutation $\sigma \in \mathfrak{S}_k$,

$$\partial_{i_1,\dots,i_k} f(a) = \partial_{i_{\sigma(1)},\dots,i_{\sigma(k)}} f(a)$$

◁ **Éléments de preuve.**

Les dérivées partielles successives commutent 2 à 2. On peut donc ramener d'abord i_1 en position initiale, par échange successif, pour faire ensuite une récurrence. Ou directement utiliser le fait que $\mathfrak{S}_n$ est engendré par les transpositions $(i\ i+1)$. ▷

Ainsi, l'ordre des variables de dérivation importe peu, dès lors que la fonction est de classe $\mathcal{C}^k$, ce qui permet de regrouper éventuellement les dérivées selon les variables, afin d'écrire ces dérivées sous la forme :

$$\frac{\partial^k f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}, \quad \alpha_1 + \cdots + \alpha_n = k.$$

Cette notation signifie qu'on a dérivé α_1 fois par rapport à x_1, α_2 fois par rapport à x_2 etc. Certains α_i peuvent être nuls et peuvent être omis dans la notation. On notera par exemple $\dfrac{\partial^3 f}{\partial x \partial z^2}$ la dérivée d'ordre 3 d'une fonction de 3 variables x, y, z, obtenue en dérivant une fois par rapport à x et 2 fois par rapport à z, mais aucune fois par rapport à y.

Proposition 16.3.32 – Opérations sur les applications de classe $\mathcal{C}^k$

Soit E, E' des $\mathbb{R}$-e.v.n. de dimension finie, et F, $F_1, \dots, F_q$ des $\mathbb{K}$-e.v.n. de dimension finie. Soit Ω, Ω' des ouverts de E et E' respectivement.

1. Soit $f, g : \Omega \to F$ de classe $\mathcal{C}^k$ et $\lambda \in \mathbb{K}$. Alors $f + \lambda g$ est de classe $\mathcal{C}^k$.
2. Soit pour $i \in [\![1,q]\!]$, $f_i : \Omega \to F_i$, et $M : F_1 \times \cdots \times F_q \to F$ une forme multilinéaire. Si les f_i sont de classe $\mathcal{C}^k$, alors $M(f_1, \dots, f_q)$ est de classe $\mathcal{C}^k$.
3. Soit $f : \Omega' \to \Omega$ et $g : \Omega \to F$. Si f et g sont de classe $\mathcal{C}^k$ sur Ω' et Ω respectivement, alors $g \circ f$ est de classe $\mathcal{C}^k$ sur Ω.
4. Si $f : \Omega \to \mathbb{K}$ ne s'annule pas et est de classe $\mathcal{C}^k$, alors $\frac{1}{f}$ aussi.
5. Soit pour $i \in [\![1,q]\!]$, $f_i : \Omega \to F_i$, et $f : \Omega \to F_1 \times \cdots \times F_q$ définie par

$$\forall x \in \Omega, \quad f(x) = (f_1(x), \dots, f_q(x)).$$

f est de classe $\mathcal{C}^k$ si et seulement si tout f_i est de classe $\mathcal{C}^k$.

Ces propriétés restent vraies pour la classe $\mathcal{C}^\infty$

◁ **Éléments de preuve.**

Démonstrations non exigibles, un peu fastidieuse mais sans difficulté. Faire une récurrence, en dérivant une première fois. ▷

Exemples 16.3.33

1. Les fonctions polynomiales $f : \mathbb{R}^n \to \mathbb{K}$, de n variables réelles et à coefficients dans $\mathbb{K}$, sont de classe $\mathcal{C}^\infty$.
2. Les fractions rationnelles aussi, sur leur domaine de définition.
3. $\det : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ est de classe $\mathcal{C}^\infty$.
4. $M \mapsto M^{-1}$ définie sur $\mathrm{GL}_n(\mathbb{R})$ est de classe $\mathcal{C}^\infty$.

Dans le cas de fonctions d'une seule variable réelle, on rajoute une propriété, concernant la dérivée itérée d'une forme bilinaire. Lorsqu'on applique cette formule au cas du produit de deux réels, on retrouve la formule de Leinbniz classique.

Théorème 16.3.34 – Formule de Leibniz vectorielle

Soit E, F, G des $\mathbb{K}$-evn de dimension finie, et $B : EtimesF \to G$ des $\mathbb{K}$-evn. Soit I un intervalle, et $f : I \to E$ et $g : I \to F$ deux applications. Si f et g sont de classe $\mathcal{C}^k$, alors $B(f,g) : t \mapsto B(f(t), g(t))$ est aussi de classe $\mathcal{C}^k$, et

$$\forall t \in I, \quad B(f,g)^{(k)}(t) = \sum_{i=1}^{k} \binom{k}{i} B(f^{(i)}(t), g^{(k-i)}(t))\ \mathrm{d}t.$$

III.4 Hessienne et formule de Taylor-Young à l'ordre 2

Définition 16.3.35 – Matrice hessienne de f

Soit Ω un ouvert de $\mathbb{R}^n$, et $f : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^2$. La matrice hessienne de f en $a \in \Omega$ est

$$H_f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{1 \leqslant i,j \leqslant n} = \begin{pmatrix} \dfrac{\partial^2 f}{\partial {x_1}^2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(a) & \cdots & \dfrac{\partial^2 f}{\partial {x_n}^2}(a) \end{pmatrix}.$$

Proposition 16.3.36 – Symétrie de $H_f(a)$

Sous les mêmes hypothèses, $H_f(a) \in \mathcal{S}_n(\mathbb{R})$.

◁ **Éléments de preuve.**

C'est une réexpression du théorème de Schwarz. ▷

Lemme 16.3.37 – Expression de la dérivée seconde par rapport à un vecteur

Soit Ω un ouvert de $\mathbb{R}^n$, et $f : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^2$. Soit $a \in \Omega$ et $H \in \mathbb{R}^n$. Alors, la dérivée seconde selon le vecteur U au point a s'exprime à l'aide de la hessienne :

$$f''_{a,U}(0) = U^T H_f(a) U = \sum_{i=1}^{n}\sum_{j=1}^{n} u_i u_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a), \quad \text{où} \quad U = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

La matrice hessienne va nous permettre d'exprimer de façon élégante la formule de Taylor-Young, c'est-à-dire l'expression d'un DL à l'ordre 2 à l'aide des dérivées d'ordre 1 et 2, sous les hypothèses idoines de régularité.

Nous avons déjà obtenu une première version de cette formule, lorsque f est une fonction de 2 variables de classe $\mathcal{C}^2$ au voisinage d'un point a, lors de la démonstration du théorème de Schwarz :

$$f(x+h,y+k) = f(x,y)+h\frac{\partial f}{\partial x}(x,y)+k\frac{\partial f}{\partial y}(x,y)+\frac{1}{2}h^2\frac{\partial^2 f}{\partial x^2}(x,y)+hk\frac{\partial^2 f}{\partial x \partial y}(x,y)+\frac{1}{2}k^2\frac{\partial^2 f}{\partial y^2}(x,y)+o(\|H\|^2),$$

où $H = \begin{pmatrix} h \\ k \end{pmatrix}$. À l'aide du lemme précédent, on obtient une réexpression synthétique de cette formule, en posant $X = \begin{pmatrix} x \\ y \end{pmatrix}$:

$$f(X+H) = f(X) + \langle \nabla f(X), H \rangle + \frac{1}{2} H^\top H_f(x) H + o(\|H\|^2).$$

Plus généralement :

Théorème 16.3.38 – Formule de Taylor-Young à l'ordre 2

Soit Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^2$ au voisinage de $a \in \Omega$. Alors f admet un DL à l'ordre 2 en a, s'exprimant :

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} h^\top H_f(a) h + o(\|h\|^2).$$

$\lhd$ **Éléments de preuve.**

Démonstration non exigible. On peut procéder comme dans le cas de 2 variables, en utilisant la formule de Taylor-Young pour les fonctions d'une variables, successivement sur chaque variable. On peut aussi exploiter la symétrie de $H_f(a)$ en considérant d'abord une b.o.n. de diagonalisation $(b_1, \dots, b_n)$, et en exploitanta la formule de Taylor à l'ordre 2 pour les fonctions partielles associées. Les exprimer par ε, et exprimer le cas général par linéairité et bilinéarité. $\rhd$

III.5 Équations aux dérivées partielles, EDP

Une équation aux dérivées partielles (EDP) est une équation d'inconnue une fonction f de plusieurs variables, reliant ses différentes dérivées partielles. L'ordre de l'EDP est l'ordre de dérivation le plus important intervenant dans cette équation. Par exemple, l'équation suivante, pour une fonction inconnue f de 2 variables, l'équation

$$\frac{\partial f^3}{\partial x^3} = \frac{\partial f^2}{\partial y^2} + xy\frac{\partial^2 f}{\partial x \partial y}$$

est une EDP d'ordre 3. La résolution d'équations aux dérivées partielles est un problème en général dur. Nous nous contentons ici de donner quelques pistes, sur des exemples simples.

Le principe va souvent être de se ramener à des équations simples. Pour cela, nous commençons par une forme simple d'equation de degré 1.

Proposition 16.3.39 – Résolution de $\dfrac{\partial f}{\partial x_i} = 0$

Soit $i \in [\![1, n]\!]$. Les propositions suivantes sont équivalentes:

(i) $f : \mathbb{R}^n \to \mathbb{R}$ est de classe $\mathcal{C}^1$ et vérifie $\dfrac{\partial f}{\partial x_i} = 0$ sur Ω

(ii) il existe $g \in \mathcal{C}^1(\mathbb{R}^{n-1}, \mathbb{R})$ tel que pour tout $(x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$f(x_1, \ldots, x_n) = g(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

En d'autres termes, les solutions de l'EDP $\dfrac{\partial f}{\partial x_i} = 0$ sont les fonctions ne dépendant pas de la vaiable x_i et de classe $\mathcal{C}^1$ par rapport aux autres variables.

Exemple 16.3.40

Résolution de $a\dfrac{\partial f}{\partial x} + b\dfrac{\partial f}{\partial y} = 0$, par changement de variable.

Nous donnons deux exemples d'EDP d'ordre 2 :

Exemples 16.3.41

1. Résoudre $\dfrac{\partial^2 f}{\partial x \partial y} = 0$ où $f \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$.
2. Résoudre $\dfrac{\partial^2 f}{\partial x^2} - a\dfrac{\partial^2 f}{\partial y^2} = 0$, où $f \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$. On cherchera un changement de variable linéaire.

IV Optimisation

IV.1 Principes généraux

L'optimisation consiste en l'étude des extrema d'une fonction f (globaux ou locaux).

Définition 16.4.1 – Extrema

Soit D un sous-ensemble de E, et $f : D \to \mathbb{R}$, et $a \in D$. On dit que :

- f admet un minimum global (resp. maximum global) en a si pour tout $x \in D$, $f(x) \geqslant f(a)$ (resp. $f(x) \leqslant f(a)$) ;
- f admet un minimum global strict (resp. maximum global strict) en a si pour tout $x \in D \backslash \{a\}$, $f(x) > f(a)$ (resp. $f(x) < f(a)$) ;
- f admet un minimum local (resp. maximum local) en a s'il existe $V \in \mathcal{V}(a)$ tel que pour tout $x \in D \cap V$, $f(x) \geqslant f(a)$ (resp. $f(x) \leqslant f(a)$) ;
- f admet un minimum local strict (resp. maximum local strict) en a s'il existe $V \in \mathcal{V}(a)$ tel que pour tout $x \in D \cap V \backslash \{a\}$, $f(x) > f(a)$ (resp. $f(x) < f(a)$).

Le but de cette section est de donner des principes de recherche des extrema (locaux ou globaux).
Les sous-sections suivantes donneront des méthodes pour trouver en quels points de D il existe un extremum, local ou global. Pour commencer, nous nous intéressons à l'existence d'un extremum global (ou local par restriction à un voisinage).
Nous rappelons le théorème de compacité, outil incontournable dans ce contexte.

Théorème 16.4.2 – Compacité

Si D est compact et $f : D \to \mathbb{R}$ est continue, alors f admet un minimum et un maximum globaux.

Il est fréquent d'étudier des fonctions qui ne sont pas définies sur des compacts. Mais il est possible de s'y ramener, soit en restreignant (si D n'est pas borné), soit en prolongeant (si D est borné).

Méthode 16.4.3 – Prouver l'existence d'un extremum global

On peut bien entendu utiliser les méthodes de localisation des extrema des paragraphes suivants et comparer les valeurs à toutes les autres, mais il peut être intéressant de savoir justifier *a priori* l'existence d'un extremum global de $f : D \to \mathbb{R}$.

1. Si D est compact et f continue, le théorème de compacité donne l'existence d'un maximum et et d'un minimum.
2. Si on parvient à trouver un compact $K \subset D$, et $a \in D$ tel que
$$\forall x \in D \backslash D, \quad f(x) \geqslant f(a),$$
alors f admet un minimum sur D, atteint dans K.
3. Si D est borné, $\overline{D}$ est borné. Si f admet une limite en tout point de $\overline{D} \backslash D$, et s'il existe $a \in D$ tel que
$$\forall b \in \overline{D} \backslash D, \quad \lim_{x \to b} f(x) \geqslant f(a),$$
alors f admet un minimum sur D.

Exemples 16.4.4 – figures 16.6 et 16.7

1. Montrer que $f : (x, y) \mapsto \dfrac{x^2 y - 3y + y^2}{(1 + x^2 + y^2)^2}$ admet un minimum et un maximum sur $\mathbb{R}^2$.
2. Montrer que $g : (x, y) \mapsto \dfrac{\mathrm{e} - \mathrm{e}^{xy}}{1 - (x^2 + y^2)}$ admet un minimum sur $\mathring{B}(0, 1)$.

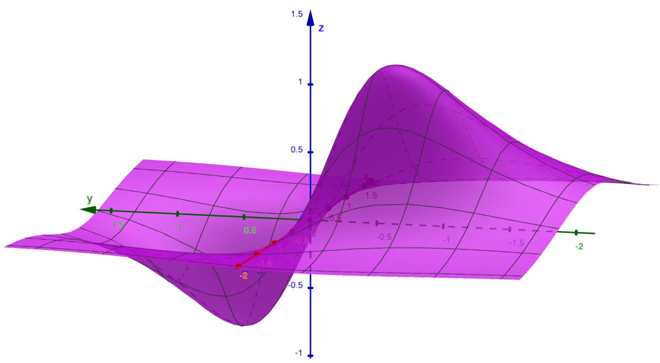

FIGURE 16.6 – $f : x \mapsto \dfrac{x^2 y - 3y + y^2}{(1 + x^2 + y^2)} = 1$

IV.2 Condition nécessaire du premier ordre

Théorème 16.4.5 – CN du premier ordre pour un extremum local

Soit $f : D \to \mathbb{R}$. Si f admet un extremum local en $a \in \mathring{D}$ et si f est différentiable en a, alors $\mathrm{d}f(a) = 0$.

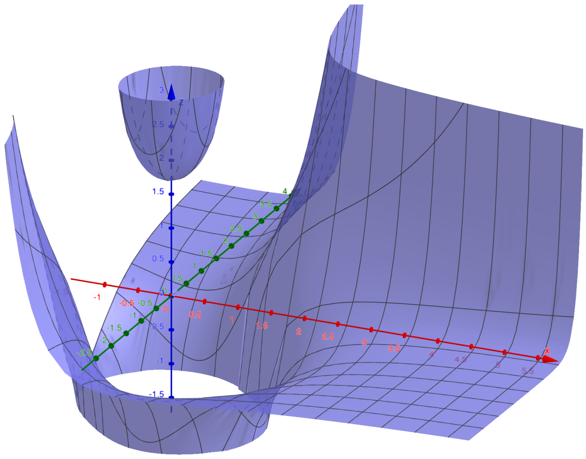

FIGURE 16.7 – $g : x \mapsto \dfrac{\mathrm{e} - \mathrm{e}^{xy}}{1 - (x^2 + y^2)}$

◁ **Éléments de preuve.**
Se ramener à la CN similaire pour les fonctions d'une variable réelle, par l'étude de toutes les dérivées directionnelles $\mathrm{d}f(a) \cdot h$. Ou alors, montrer directement la proposition 16.4.15 qui est une version plus générale de ce théorème. ▷

Ce théorème motive la définition suivante, comme dans le cas des fonctions d'une variable réelle.

Définition 16.4.6 – Point critique

Soit $f : D \to \mathbb{R}$. On dit que $a \in E$ est un point critique de f si $a \in \mathring{D}$ et $\mathrm{d}f(a) = 0$.

Remarque 16.4.7

Si f est une fonction d'une variable réelle, $a \in \mathring{D}$ est point critique si et seulement si f est dérivable en a et l'application linéaire $\mathrm{d}f(a) : h \mapsto f'(a)h$ est nulle, c'est-à-dire si et seulement si $f'(a) = 0$. On retrouve la notion de point critique définie en MPSI.

Corollaire 16.4.8 – CN portant sur le gradient dans le cas euclidien

Soit E un espace euclidien, $D \subset E$, et $f : D \to \mathbb{R}$. Si f admet un extremum en $a \in \mathring{D}$ en lequel f est différentiable, alors $\nabla f(a) = 0$.

Exemple 16.4.9

Rechercher les points critiques de $g : (x, y) \mapsto \dfrac{\mathrm{e} - \mathrm{e}^{xy}}{1 - x^2 - y^2}$ sur $\mathring{B}(0, 1)$, et en déduire l'existence d'un minimum global.

Remarque 16.4.10

Si le nombre de points critiques est peu élevé et si on a pu montrer l'existence d'extrema locaux, comme dans le paragraphe précédent, l'étude de la CN du premier ordre peut être suffisante pour conclure.

Avertissement 16.4.11

Comme dans le cas d'une fonction d'une variable réelle, $\mathrm{d}f(a) = 0$ n'est pas une CN pour avoir un extremum local.

Exemple 16.4.12 – figure 16.8

Quels sont les points critiques de $f : (x, y) \mapsto xy$? La fonction f admet-elle un extremum local ? Cet exemple montre au passage que le fait que les restrictions de f à toutes les droites passant par a admettent un extremum local en a n'est pas suffisant pour que f admette un extremum local en a.

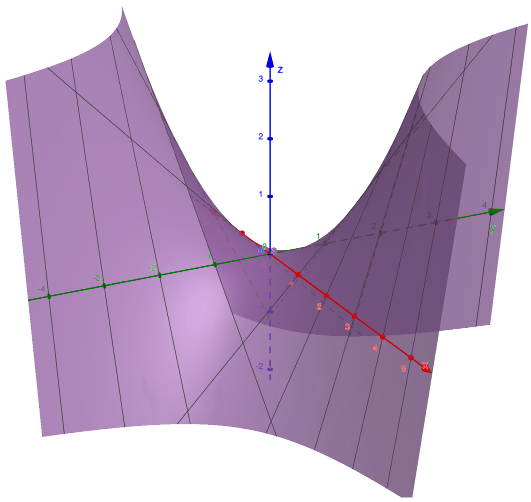

Figure 16.8 – $f : x \mapsto xy$

Méthode 16.4.13 – Rechercher les extrema locaux sur un ouvert

On cherche les points critiques de f, puis on étudie chaque point critique a :

- soit par étude directe du signe de $f(x) - f(a)$,
- soit en utilisant la CN d'ordre 2 qui fait l'objet d'un paragraphe ultérieur.

Méthode 16.4.14 – Rechercher les extrema globaux

Soit $f : D \to \mathbb{R}$, différentiable sur $\mathring{D}$.

- Rechercher les points critiques sur $\mathring{D}$. Les extrema globaux, s'ils existent, sont un point critique, ou un point de $D \backslash \mathring{D}$.
- Si on a pu justifier l'existence d'un extremum global, on peut le trouver en comparant les valeurs prises aux différents points critiques, ainsi que le maximum et/ou le minimum de f restreinte à $D \backslash \mathring{D}$.

Le paragraphe suivant donne justement des techniques en vue de l'étude des extrema d'une fonction retreinte. Cela permet notamment l'étude aux bords (si D lui-même peut être plongé dans un ouvert Ω de sorte que $D \backslash \mathring{D}$ soit une partie de Ω).

IV.3 Optimisation sous contrainte

Soit $f : \Omega \to \mathbb{R}$, et $X \subset D$. Optimiser f sous la contrainte X consiste à étudier les extrema de f en se retreignant à X (la contrainte). Il s'agit donc d'étudier les extrema d'une restriction. Nous commençons par une CN similaire à celle qu'on a trouvé dans le cas global (qui d'ailleurs peut être vu comme conséquence de celle-ci).

Proposition 16.4.15 – CN du premier ordre pour une restriction

Soit Ω un ouvert de E, X une partie de Ω, et $f : \Omega \to \mathbb{R}$. Si $f_{|X}$ admet un extremum local en $a \in X$, alors

$$\forall h \in T_aX, \quad \mathrm{d}f(a) \cdot h = 0.$$

En d'autres termes, $\mathrm{d}(a)$ s'annule en tout vecteur tangent à X, c'est-à-dire $T_aX \subset \operatorname{Ker} \mathrm{d}f(a)$.

◁ **Éléments de preuve.**
Dériver $f(\alpha(t))$, où $\alpha = [-\varepsilon, \varepsilon] \to X$ vérifie $\alpha(0) = a$. ▷

Remarques 16.4.16

1. En prenant $X = \mathring{D}$, puisque X est un voisinage de a, $T_aX = E$, et on retrouve donc le théorème 16.4.5.
2. Si E est euclidien, la proposition 16.4.15 affirme que si $f_{|X}$ admet un extremum local en a, $\nabla f(a) \perp T_aX$.

Si l'ensemble X est défini par une équation $g(x) = 0$, on peut combiner le résultat précédent et le théorème 16.3.21.

Théorème 16.4.17 – Optimisation sous contrainte

Soit $f, g : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^1$, et $X = \{x \in \Omega,\ g(x) = 0\}$. Si $f_{|X}$ admet un extremum local en $a \in X$ tel que $\mathrm{d}g(a) \neq 0$, alors $\mathrm{d}f(a)$ et $\mathrm{d}g(a)$ sont colinéaires.

Exemple 16.4.18

Après avoir montré leur existence, déterminer les extrema globaux de $f : (x, y) \mapsto x^2 + xy - 2y^2$ sur $\overline{B}(0, 1)$.

IV.4 Étude à l'ordre 2

Si $f : I \to \mathbb{R}$ est une fonction définie sur un intervalle ouvert de $\mathbb{R}$, et si $a \in I$ est un point critique de f, on peut déterminer s'il s'agit effectivement d'un extremum par un argument de convexité locale, ce qui, en cas de régularité suffisante, se ramène à l'étude du signe de $f''(a)$. Nous recherchons une condition similaire lorsque f est une fonction définie sur un ouvert de $\mathbb{R}^n$, à l'aide de la matrice hessienne $H_f(a)$.

Proposition 16.4.19 – CN du second ordre

Soit $f : \Omega \to \mathbb{R}$ (où Ω est un ouvert de $\mathbb{R}^n$) de classe $\mathcal{C}^2$. Si f admet un minimum local (resp. maximum local) en $a \in \Omega$, alors a est un point critique de f et $H_f(a) \in \mathcal{S}^+(\mathbb{R})$ (resp. $-H_f(a) \in \mathcal{S}^+(\mathbb{R})$).

◁ **Éléments de preuve.**
D'après la formule de Taylor-Young, le signe de $f(a + h) - f(a)$ est celui de $h^\top H_f(a)h$. ▷

Théorème 16.4.20 – CS du second ordre

Soit $f : \Omega \to \mathbb{R}$ (où Ω est un ouvert de $\mathbb{R}^n$) de classe $\mathcal{C}^2$, et soit $a \in \Omega$. Si a est un point critique de f et si $H_f(a) \in \mathcal{S}^{++}(\mathbb{R})$ (resp. $-H_f(a) \in \mathcal{S}^{++}(\mathbb{R})$), alors f admet un minimum local strict (resp. maximum local strict) en a.

◁ **Éléments de preuve.**
Toujours le contrôle du signe par Taylor-Young, en travaillant avec une norme adaptée à $H_f(a)$. ▷

Avertissement 16.4.21

- La CN du second ordre n'est pas une CS.
- La CS du second ordre n'est pas une CN

Exemples 16.4.22

Étudier les CN et CS du second ordre pour les fonctions suivantes, en 0.

1. $f : (x, y) \mapsto (x^2 + y^2)^2$.
2. $g : (x, y) \mapsto x^3 y^3$.

Méthode 16.4.23 – Étude d'un point critique

Pour savoir si un point critique correspond à un extremum local, on étudiera le signe de la hessienne,

- soit directement par mises sous forme canonique successives de la forme quadratique $h \mapsto h^\top H_f(a)h$,
- soit par étude des valeurs propres de $H_f(a)$.

On peut alors conclure :

- Si $H_f(a)$ n'est pas positive (*i.e.* s'il existe une valeur propre strictement négative, ou s'il existe h tel que $h \mapsto h^\top H_f(a)h < 0$, alors f n'admet pas d'extremum local en a
- Si $H_f(a)$ est strcitement positive (si $\mathrm{Spec}(H_f(a)) \subset \mathbb{R}_+^*$), alors f admet un minimum local.

On ne peut pas conclure de cette façon si ces conditions ou les conditions symétriques (pour un maximum) ne sont pas satisfaites, à savoir dans le cas où $0 \in \mathrm{Spec}(H_f(a))$.

Corollaire 16.4.24 – Explicitation pour $n = 2$

Soit Ω un ouvert de $\mathbb{R}^2$, $f : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^2$, et a un point critique de f. Alors :

- Si $\det(H_f(a)) > 0$ et $\mathrm{tr}(H_f(a)) > 0$, alors f admet un minimum local strict en a
- Si $\det(H_f(a)) > 0$ et $\mathrm{tr}(H_f(a)) < 0$, alors f admet un maximum local strict en a
- si $\det(H_f(a)) < 0$, alors f n'admet pas d'extremum local en a.

Exemples 16.4.25

Étudier la nature du point critique $(0, 0)$ dans les 3 cas suivants :

1. $f_1(x, y) = y^4 + x^2 y + x^2 + 2xy + 2y^2$
2. $f_2(x, y) = x^3 y + x^2 - 4xy + y^2$
3. $f_3(x, y) = xy^2 + x^3 - x^2 + 2xy - 2y^2$.

17

Intégration vectorielle et équations différentielles linéaires

Ce chapitre consiste en un certain nombre de rappels concernant la résolution des équations différentielles linéaires, ainsi qu'un certain nombre de compléments. En effet, le contexte des e.v.n., et les techniques matricielles (notamment la diagonalisation et l'utilisation de l'exponentielle complexe) nous permettent d'avoir des outils efficaces pour gérer des situations bien plus générales que ce que vous avez vu en MPSI.

Les fonctions que nous considérerons sont des fonctions $y : I \to E$, définies sur un intervalle I de $\mathbb{R}$, et à valeurs dans un e.v.n. E de dimension finie sur le corps $\mathbb{K}$, égal à $\mathbb{R}$ ou $\mathbb{C}$. Une équation différentielle est une équation reliant les valeurs de f et de ses dérivées successives, évaluées au même point $t \in I$. Nous nous limitons dans ce chapitre à une forme bien particulière d'équations différentielles : les équations différentielles linéaires. Les autres types d'équations différentielles sont bien plus complexes à étudier, et il n'est pas toujours possible d'expliciter les solutions.

Les équations différentielles apparaissent de façon naturelle dans de nombreux contextes, en particulier en physique. Le cas linéaire est souvent suffisant, les physiciens étant des habitués de la linéarisation, donnant souvent (mais pas toujours) des approximations suffisamment précises pour être pertinentes pour la situation étudiée.

I Compléments d'intégration

On présente dans cette section quelques rappels et compléments sur les intégrales de fonctions d'une variable réelle, à valeurs dans un $\mathbb{K}$-e.v.n. F de dimension finie n (où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). Les fonctions seront définie sur un segment $I = [a, b]$.

I.1 Définition et propriétés générales

Dans cette sous-partie, f désigne une application d'un segment I dans F. Comme dans le cas d'une application à valeurs réelles, on dit que f est continue par morceaux si I se subdivise en un nombre fini de segments sur l'intérieur desquels f est continue, et au bord desquels elle a une limite dans F.
Pour toute base $\mathcal{B}$ de F, on note $(f_{\mathcal{B},1}, \dots, f_{\mathcal{B},n})$ les coordonnées de f dans $\mathcal{B}$, c'est-à-dire les fonctions $f_{\mathcal{B},i} : I \to \mathbb{K}$ telles que pour tout $t \in I$,

$$f(t) = \sum_{i=1}^{n} f_{\mathcal{B},i}(t) b_i.$$

S'il n'y a pas d'ambiguïté sur la base, on notera simplement f_i au lieu de $f_{\mathcal{B},i}$.

Proposition 17.1.1 – Caractérisation de la continuité par morceaux via les coordonnées

Les propositions suivantes sont équivalentes: :

(i) f est continue par morceaux sur I ;

(ii) pour toute base $\mathcal{B}$ de F et tout $i \in [\![1, n]\!]$, $f_{\mathcal{B},i}$ est continue par morceaux ;

(iii) il existe une base $\mathcal{B}$ de F telle que pour tout $i \in [\![1, n]\!]$, $f_{\mathcal{B},i}$ soit continue par morceaux.

◁ **Éléments de preuve.**

Découle de la caractérisation coordonnée par coordonnée de la continuité et des limites. ▷

Proposition/Définition 17.1.2 – Définition de l'intégrale vectorielle

Soit $f : I = [a, b] \to F$ une fonction continue par morceaux. Le vecteur

$$\sum_{k=1}^{n} \left(\int_a^b f_{\mathcal{B},i}(t) \, \mathrm{d}t \right) b_i$$

est bien définie, et indépendante du choix de la base $\mathcal{B} = (b_1, \ldots, b_n)$ de F. On définit alors

$$\int_a^b f(t) \, \mathrm{d}t = \sum_{k=1}^{n} \left(\int_a^b f_{\mathcal{B},i}(t) \, \mathrm{d}t \right) b_i,$$

la valeur commune de ces expressions. On trouve aussi les notations

$$\int_a^b f, \qquad \int_{[a,b]} f(t) \, \mathrm{d}t \qquad \text{ou} \qquad \int_{[a,b]} f.$$

◁ **Éléments de preuve.**

Exprimer le changement de base matriciellement, et utiliser la linéarité de l'intégrale scalaire. ▷

En travaillant coordonnées par coordonnées, on transfert de façon quasi-immédiate une grande partie des propriétés usuelles de l'intégrale scalaire.

Proposition 17.1.3 – Linéarités

Soit $f, g : I \to F$ continues par morceaux sur $I = [a, b]$.

1. Linéarité : $\displaystyle\int_a^b (f + \lambda g) = \int_a^b f + \lambda \int_a^b g$;
2. si $\alpha : I \to \mathbb{R}$ est c.p.m. et $u \in F$, alors $\displaystyle\int_a^b \alpha(t) u \, \mathrm{d}t = \int_a^b \alpha(t) \, \mathrm{d}t \cdot u$.
3. Pour tout $L \in \mathcal{L}(F, G)$, $\displaystyle\int_a^b L \circ f = L \left(\int_a^b f \right)$;

◁ **Éléments de preuve.**

Le point 1 se ramène à la propriété similaire pour les fonctions scalaires, par caractérisation par les coordonnées.

Le point 2 s'obtient en choisissant une base contenant u.

Le point 3 est conséquence des 2 premiers, en tranvaillant encore dans une base $\mathcal{B}$. On utilise d'abord la linéarité de L pour entrer L dans la somme, puis la linéarité de l'intégrale pour sortir L de l'intégrale, et encore la linéarité de L pour entrer l'intégrale dans L ! ▷

Proposition 17.1.4 – Relation de Chasles

Soit $f : I \to F$, c.p.m. sur $I = [a,b]$ et $c \in [a,b]$. Alors

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

◁ **Éléments de preuve.**

Encore les coordonnées. ▷

Le point suivant se généraliserait sur des subdivisions non régulières dont le pas tend vers 0, mais seul le cas des subdivisions régulières est au programme.

Théorème 17.1.5 – Sommes de Riemann

Soit f une fonction continue par morceaux sur $[a,b]$ à valeurs dans F. Alors

$$\frac{b-a}{n}\sum_{k=0}^{n-1} f\left(a + k\cdot\frac{b-a}{n}\right) \longrightarrow \int_a^b f(t)\ \mathrm{d}t.$$

◁ **Éléments de preuve.**

On peut voir F comme un $\mathbb{R}$-ev (même si au départ il est complexe). Se ramener au cas de fonctions à valeurs réelles en travaillant dans une $\mathbb{R}$-base. ▷

Théorème 17.1.6 – Inégalité triangulaire

On suppose que F est un $\mathbb{K}$-e.v.n. Soit $f : I = [a,b] \to F$ une application c.p.m., avec $a < b$. Alors $\|f\|_F$ est c.p.m. et

$$\left\|\int_a^b f(x)\ \mathrm{d}x\right\|_F \leqslant \int_a^b \|f(x)\|_F\ \mathrm{d}x.$$

◁ **Éléments de preuve.**

Par sommes de Riemann. ▷

Même si l'inégalité est vraie quelle que soit la norme de F choisie, les valeurs de ces intégrales dépendent de la norme.

Avertissement 17.1.7

1. Il faut évidemment utiliser la même norme de F dans les deux membres de l'inégalité !
2. Attention ici à l'hypothèse $a < b$. Si $b < a$, il faut rajouter un signe pour contrer l'inversion des bornes. Ou rajouter une valeur absolue externe.

Remarque 17.1.8

Le cas des fonctions $f : I \to \mathbb{C}$ est déjà connu (programme de MPSI). Cependant, on peut remarquer qu'on le retrouve ici, à partir du cas $f : I \to \mathbb{R}$, en voyant $\mathbb{C}$ comme un $\mathbb{R}$-e.v.n..

I.2 Théorème fondamental

Pour terminer ce passage en revue rapide des intégrales vectorielles, on généralise le théorème fondamental du calcul intégral, et ses conséquences pour l'obtention de l'inégalité des accroissements finis, puis ses généralisations usuelles à des ordres de dérivation supérieure que sont les inégalités de Taylor-Lagrange. Comme dans le cas numérique, le calcul d'une intégrale se ramène au calcul de la variation d'une primitive entre les deux bornes.

Définition 17.1.9 – Primitive de g

Soit $g, G : I \to F$. On dit que G est une primitive de g si G est dérivable sur I et $G' = g$.

Proposition 17.1.10 – Comparaison de deux primitives

Soit $g : I \to F$, continue sur l'intervalle I, et G_1 et G_2 deux primitives de g. Alors $G_2 - G_1$ est constante.

◁ **Éléments de preuve.**

Par linéarité de la dérivation, et caractérisation des fonctions contantes. ▷

Théorème 17.1.11 – Théorème fondamental de l'analyse, version vectorielle

1. Soit $g : I \to F$ une application continue sur $I = [a, b]$, et $a \in I$. Alors :

$$G_0 : x \mapsto \int_a^x g(x) \, \mathrm{d}x$$

est une primitive de g. De plus, G_0 est alors de classe $\mathcal{C}^1$.

2. Soit $g : I \to F$ et G une primitive de g. Alors

$$\int_a^b g(x) \, \mathrm{d}x = \Big[G(x)\Big]_a^b = G(b) - G(a).$$

◁ **Éléments de preuve.**

1. Se ramener au cas scalaire par choix d'une base.
2. G_0 diffère de G d'une constante, donc $G(b) - G(a) = G_0(b) - G_0(a)$.

▷

Voici une réexpression classique du théorème fondamental.

Corollaire 17.1.12 – Intégrale d'une dérivée

Soit $f \in \mathcal{C}^1(I, F)$. Alors, pour tout $x \in I$,

$$f(x) = f(a) + \int_a^x f'(t) \, \mathrm{d}t.$$

Corollaire 17.1.13 – Inégalité des accroissements finis

Soit $f \in \mathcal{C}^1(I, F)$, telle que f' soit bornée. Alors, pour tout $x \in I$,

$$\|f(b) - f(a)\| \leqslant |b - a| \cdot \|f'\|_\infty.$$

Les propriétés calculatoires (formule de changement de variable, formule d'intégration par parties) peuvent s'obtenir de même que dans le cas scalaire, à l'aide du théorème fondamental. Ces formules ne sont pas au programme, et seront établies en exercice seulement.
Le théorème fondamental de l'analyse permet aussi de dériver des intégrales vectorielles dépendant de façon $\mathcal{C}^1$ de leur borne, en se ramenant à des dérivées de composées.

I.3 Formules de Taylor

Enfin, nous généralisons les formules de Taylor dans le cadre de fonctions d'une variable à valeurs vectorielles.

Théorème 17.1.14 – Formule de Taylor vectorielle avec reste intégral

Soit $f \in \mathcal{C}^{n+1}([a,b],F)$, à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. Alors

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

◁ **Éléments de preuve.**

Se ramener à la formule de Taylor scalaire appliquée aux fonctions coordonnées. On peut aussi procéder directement vectoriellement, en faisant une IPP itérée, à condition de démontrer d'abord la formule d'IPP vectorielle (voir exercices). ▷

Remarque 17.1.15

Cette formule ne nécessite pas d'hypothèse $a < b$. Elle reste valide lorsque $b \leqslant a$.

On en déduit, comme dans le cas scalaire :

Théorème 17.1.16 – Inégalité de Taylor-Lagrange vectorielle

Soit $a < b$, et f une fonction de classe $\mathcal{C}^{n+1}$ sur $[a,b]$, à valeurs dans F. Alors $f^{(n+1)}$ est bornée sur $[a,b]$, et

$$\left| f(b) - \sum_{k=0}^{n} f^{(k)}(a)\frac{(b-a)^k}{k!} \right| \leqslant \|f^{(n+1)}\|_\infty \cdot \frac{|b-a|^{n+1}}{(n+1)!}$$

En pratique, on cherchera explicitement à majorer $\|f^{(n+1)}\|$ par M, pour obtenir :

$$\left| f(b) - \sum_{k=0}^{n} f^{(k)}(a)\frac{(b-a)^k}{k!} \right| \leqslant M \cdot \frac{|b-a|^{n+1}}{(n+1)!}$$

Enfin, la formule de Taylor-Young passe aussi au cas vectoriel.

Théorème 17.1.17 – Formule de Taylor-Young à l'ordre n au point x_0

Soit I un intervalle ouvert de $\mathbb{R}$, $x_0 \in I$, et $f : I \to F$ une fonction de classe $\mathcal{C}^n$ au voisinage de x_0. Alors :

$$f(x) \underset{x_0}{=} \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + o((x-x_0)^n).$$

◁ **Éléments de preuve.**

Coordonnée par coordonnée, les dérivées à tout ordre et les o se caractérisant bien ainsi. ▷

II Propriétés générales des EDL

II.1 Notations et définitions

Dans toute la suite, I désigne un intervalle de $\mathbb{R}$, et E un e.v.n. de dimension finie sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Notation 17.2.1 – Évaluation d'un endomorphisme

Soit $a \in \mathcal{L}(E)$, et $x \in E$. Ainsi qu'on l'a déjà fait dans l'étude des différentielles, on notera dans ce chapitre $a \cdot x$ l'évaluation de a en x. Ainsi, par définition,

$$a \cdot x = a(x) \in E.$$

Les équations différentielles linéaires font intervenir des fonctions $a : I \to \mathcal{L}(E)$. Ainsi, pour $t \in I$, $a(t)$ est un endomorphisme de E. La notation s'explique alors par le fait qu'elle offre un peu plus de clarté dans les écritures, puisqu'elle permet d'écrire $a(t) \cdot x$ au lieu du très lourd $a(t)(x)$.
Par ailleurs, dans le cas courant où $E = \mathbb{K}$, un endomorphisme consiste en la multiplication par un scalaire. Via l'identification des endomorphismes aux scalaires de $\mathbb{K}$, l'évaluation $a(x)$ correspond au produit ax, ce qui rend la notation ci-dessus cohérente et compatible avec cette situation.

Définition 17.2.2 – Équation différentielle linéaire

1. Une équation différentielle linéaire d'ordre r est une équation de fonction inconnue $y : I \to E$, où I est un intervalle de $\mathbb{R}$ et E un e.v.n. de dimension finie sur $\mathbb{K}$, de la forme :
$$\forall t \in I, \quad a_r(t) \cdot y^{(r)}(t) + \cdots + a_0(t) \cdot y(t) = b(t), \qquad \text{soit:} \qquad \sum_{k=0}^{r} a_k(t) \cdot y^{(k)}(t) = b(t),$$
où :
   - pour tout $k \in [\![0, r]\!]$, $a_k \in \mathcal{C}^0(I, \mathcal{L}(E))$;
   - $b \in \mathcal{C}^0(I, E)$.
2. L'équation différentielle sera souvent notée de façon abrégée et abusive :
$$(E) \ : \ a_r(t) \cdot y^{(r)} + \cdots + a_0(t) \cdot y = b(t).$$
3. On dit qu'une fonction $\varphi : I \to E$ est solution de l'équation E si pour tout $t \in I$,
$$a_r(t) \cdot f^{(r)}(t) + \cdots + a_0(t) \cdot y(t) = b(t).$$

Définition 17.2.3 – Équation différentielle scalaire

On dit qu'un EDL d'ordre r est scalaire si $E = \mathbb{K}$. Ainsi, $\mathcal{L}(E)$ peut être assimilée à $\mathbb{K}$, et les a_k sont donc des fonctions continues de I dans $\mathbb{K}$. La notation $a_k(t) \cdot y^{(k)}(t)$ désigne alors le produit dans $\mathbb{K}$, au sens usuel.

II.2 Structure de l'ensemble des solutions

Théorème 17.2.4 – Structure de l'ensemble des solutions d'une EDL

Soit l'équation différentielle
$$(E) \ : \ a_r(t) \cdot y^{(r)} + \cdots + a_0(t) \cdot y = b(t),$$
et y_0 une solution particulière de (E). Alors l'ensemble $\mathcal{S}$ des solutions de (E) est un sous-espace affine de (E), passant par y_0, et de direction l'ensemble des solutions $\mathcal{S}_H$ de l'équation homogène associée
$$(EH) \ : \ a_r(t) \cdot y^{(r)} + \cdots + a_0(t) \cdot y = 0.$$
Ainsi,
$$\mathcal{S} = y_0 + \mathcal{S}_H.$$

En appelant $\Phi \in \mathcal{L}(\mathcal{C}^r(I, E), \mathcal{C}^0(I, E))$ l'endomorphisme tel que
$$\forall t \in I, \quad \Phi(y)(t) = a_r(t) \cdot y^{(r)}(t) + \cdot + a_0(t) \cdot y(t),$$
l'espace affine $\mathcal{S}$ correspond à la *fibre* de Φ passant par b, c'est-à-dire à l'image réciproque de b, qui est un sous-espace affine passant par un antécédent quelconque y_0, et dirigée par le noyau de Φ, qui n'est autre que $\mathcal{S}_H$.

Méthode 17.2.5 – Résoudre une EDL

On s'y prendra souvent en 2 temps :

1. Résoudre l'équation homogène associée
2. Trouver une solution particulière.

On pourra remarquer que la connaissance des solutions de l'équation homogène aide souvent à la recherche d'une solution particulière, grâce aux méthodes de variation des constantes.

Pour la recherche d'une solution particulière, on peut parfois « isoler les difficultés » en séparant le second membre en somme de plusieurs fonctions pour lesquelles une solution particulière est plus simple à exprimer. On reconstitue alors une solution particulière de l'équation initiale grâce au résultat ci-dessous.

Proposition 17.2.6 – Principe de superposition

Si $b = b_1 + b_2$, pour trouver une solution particulière y_0 de l'équation (E), il suffit de trouver :

- une solution particulière y_1 de (E_1) : $a_r(x)y^{(r)} + \cdots + a_0(x)y = b_1(x)$;
- une solution particulière y_2 de (E_2) : $a_r(x)y^{(r)} + \cdots + a_0(x)y = b_2(x)$

Une solution particulière de (E) est alors $y_1 + y_2$.
Cela se généralise de façon immédiate à une décomposition de b en une somme finie quelconque de fonctions continues.

◁ **Éléments de preuve.**
Vérification facile, par linéarité de la dérivation. ▷

Méthode 17.2.7 – Recherche de solutions particulières

On suppose ici que (E) est scalaire.

1. Si les a_i sont constantes, et $b(t) = P(t)\mathrm{e}^{\lambda t}$, où P est un polynôme, on pourra rechercher une solution particulière sous la même forme. On précisera un peu plus loin.
2. Si les a_i et b sont polynomiaux, on peut essayer de rechercher une solution particulière sous forme polynomiale, ou (au moins localement) sous forme d'une série entière. De la sorte, en posant les équations sur les coefficients, on peut déterminer toutes les solutions développables en séries entières. Le cas polynomial n'en est qu'un cas particulier.
3. Cela peut se généraliser au cas où les a_i et b sont seulement développables en séries entière, mais c'est plus technique, et cela nécessite de faire des produits de Cauchy.
4. Attention à bien justifier que le rayon de convergence obtenu est non nul pour obtenir une solution locale !

Exemples 17.2.8

1. Déterminer toutes les solutions développables en séries entières au voisinage de 0, de l'équation
$$(x^2+1)y' + x^3 y = \frac{1}{1-x}.$$
On se contentera de donner une relation de récurrence permettant de calculer les coefficients de ces solutions, mais sans les expliciter. On ne demande pas non plus d'écrire ces solutions avec les fonctions usuelles. En revanche, on vérifiera que le rayon de convergence est non nul.
2. Montrer qu'il n'existe qu'une solution de l'équation
$$\ln(1+x)y' + y = x$$
développable en série entière au voisinage de 0. Exprimer une relation permettant de calculer

les coefficients de cette solution de proche en proche.

II.3 Forme normalisée

Définition 17.2.9 – Forme normalisée d'une EDL

Une équation différentielle linéaire d'ordre r est sous forme normalisée si elle est de la forme :

$$y^{(r)} = \alpha_{r-1}(t) \cdot y^{(r-1)} + \cdots + \alpha_0(t) \cdot y + \beta(t).$$

Remarque 17.2.10

Si dans l'équation (E) du paragraphe I-2, pour tout $t \in I$, $a_r(t) \in \mathrm{GL}(E)$, alors on peut mettre (E) sous forme normale en posant

$$\forall k \in [\![0, r-1]\!], \forall t \in I, \quad \alpha_k(t) = -a_r(t)^{-1} \circ a_k(t) \qquad \text{et} \qquad b(t) = -a_r(t)^{-1} \cdot b(t).$$

Méthode 17.2.11 – Résoudre une EDL scalaire non normalisée

On suppose que l'équation (E) est scalaire, et que a_r ne s'annule qu'en un nombre fini $t_1 < \cdots < t_{k-1}$ de réels de l'intervalle I.

- Sur chaque intervalle délimité par les bornes de I et les t_k, mettre (E) sous forme normale et la résoudre avec les méthodes usuelles
- Etudier l'existence d'une solution globale par raccordement des fonctions obtenues sur chacun des intervalles. Au point de raccordement, la fonction doit être r fois dérivable. On pourra étudier la dérivabilité à gauche et à droite jusqu'à l'ordre r, soit de façon directe (avec les expressions), soit en s'aidant du théorème de la classe $\mathcal{C}^r$ par prolongement.

On verra des exemples plus tard, lorsqu'on aura revu les méthodes usuelles de résolution des EDL normalisées.

III EDL d'ordre 1

III.1 Situation

Dans cette partie, nous étudions les équations différentielles d'ordre 1, de fonction inconnue $y : I \to E$, où E est un $\mathbb{K}$-e.v.n. de dimension finie, sous forme normalisée. Ainsi, ce sont des équations de la forme

$$(E) \ : \ y' = a(t) \cdot y + b(t),$$

où $a \in \mathcal{C}^0(I, \mathcal{L}(E))$, et $b \in \mathcal{C}^0(I, E)$. L'équation homogène associée est

$$(EH) \ : \ y' = a(t) \cdot y.$$

Proposition 17.3.1 – Trois points de vue équivalents

Soit I un intervalle de $\mathbb{R}$. Les 3 points de vues suivants sur les EDL d'ordre 1 sont équivalents :

(i) équation vectorielle $y' = a(t) \cdot y + b(t)$, où $a \in \mathcal{C}^0(I, \mathcal{L}(E))$, $b \in \mathcal{C}^0(I, E)$, E étant un $\mathbb{K}$-e.v.n. de dimension n ;

(ii) équation matricielle $Y' = A(t) \cdot Y + B(t)$, où $A \in \mathcal{C}^0(I, \mathcal{M}_n(\mathbb{K}))$, $B \in \mathcal{C}^0(I, \mathbb{K}^n)$;

(iii) système de n équations différentielles scalaires d'ordre 1

$$\begin{cases} y_1' &= a_{1,1}(t)y_1 + \cdots + a_{1,n}(t)y_n + b_1(t) \\ y_2' &= a_{2,1}(t)y_1 + \cdots + a_{2,n}(t)y_n + b_2(t) \\ & \vdots \\ y_n' &= a_{n,1}(t)y_1 + \cdots + a_{n,n}(t)y_n + b_n(t) \end{cases}$$

où les $a_{i,j}$ et b_i sont des applications continues de I dans $\mathbb{K}$.

◁ **Éléments de preuve.**

- L'équivalence entre (i) et (ii) provient du choix d'une base dans un sens, et du fait que (ii) est de la forme de (i) dans $E = \mathcal{M}_n(\mathbb{K})$. Ainsi, (i) est une forme plus générale de (ii), mais pouvant toujours se ramener à (ii) pâr choix d'une base.
- L'équivalence entre (ii) et (iii) est claire en explicitant le produit matriciel coefficient par coefficient.

On vérifiera que les propriétés de continuité sont bien préservés à chaque implication. ▷

Nous adopterons le plus souvent le point de vue matriciel.

III.2 Problème de Cauchy et espace des solutions

On se donne une EDL d'ordre 1, sous forme matricielle

$$(E) \ : \ Y' = A(t)Y + B(t),$$

où $A \in \mathcal{C}^0(I, \mathcal{M}_n(\mathbb{K}))$ et $B \in \mathcal{C}^0(I, \mathbb{K}^n)$.
D'après les résultats généraux l'ensemble des solutions de (E) est un espace affine dirigé par l'espace vectoriel des solutions de l'équation homogène (EH). Nous nous intéressons ici notamment à la dimension de l'espace des solutions de (EH).

Définition 17.3.2 – Problème de Cauchy pour une EDL d'ordre 1

Soit $t_0 \in I$, considéré comme « temps initial ».

1. Un problème de Cauchy pour l'équation (E) en t_0 est la donnée de conditions initiales en t_0 de la forme $Y(t_0) = Y_0$, où $Y_0 \in \mathbb{K}^n$ est fixé.
2. Une solution $Y : I \to \mathbb{K}^n$ de (E) est solution du problème de Cauchy $Y(t_0) = Y_0$ si

$$\forall t \in I, \quad Y'(t) = A(t)Y(t) + B(t) \qquad \text{et} \qquad Y(t_0) = Y_0.$$

Proposition 17.3.3 – Forme intégrale d'un problème de Cauchy

Les propositions suivantes sont équivalentes:

- Y est solution du problème de Cauchy $Y(t_0) = Y_0$ associé à l'équation (E) ;
- Pour tout $t \in I$,

$$Y(t) = Y_0 + \int_{t_0}^{t} A(u)Y(u)\ \mathrm{d}u.$$

Remarque 17.3.4

Traduire la définition et la propriété qui suit dans les points de vue vectoriel et système linéaire.

Voici un résultat d'une importance capitale, qui peut se généraliser, au moins localement, sous certaines hypothèses dans un contexte non linéaire.

Théorème 17.3.5 – Théorème de Cauchy linéaire pour l'ordre 1

Soit $Y_0 \in \mathbb{K}^n$. L'équation différentielle (E) admet une et une seule solution répondant au problème de Cauchy $Y(t_0) = Y_0$.

◁ **Éléments de preuve.**

Démonstration non exigible (mais pas hors-programme).

- Existence. On définit $\Phi : \mathcal{C}^0(I, \mathbb{K}^n) \to \mathcal{C}^0(I, \mathbb{K}^n)$ qui à Y associe le membre de droite de l'expression intégrale du problème de Cauchy. On en cherche un point fixe. Considérer $(F_n)_{n\in\mathbb{N}}$ définie par la récurrence
$$F_{n+1} = \Phi(F_n),$$
et F_0 constante de valeur Y_0. Montrer que $\sum(F_{n+1} - F_n)$ converge normalement sur tout compact K de I. On pourra pour cela majorer sa norme par récurrence par le terme général d'une série exponentielle de paramètre $\ell(K)$.
- Unicité. Si F et G vérifient tous deux les conditions de Cauchy, adapter le raisonnement précédent pour montrer que $F - G$ est nécessairement nul.

▷

Corollaire 17.3.6 – Dimension de l'espace des solutions de (EH)

Soit $t_0 \in I$. On note $\mathcal{S}_H$ l'espace vectoriel des solutions de l'équation homogène (EH).

1. L'application $\mathrm{ev}_{t_0} : Y \mapsto Y(t_0)$ est un isomorphisme de $\mathcal{S}_H$ dans $\mathbb{K}^n$.
2. $\dim \mathcal{S}_H = n$.

◁ **Éléments de preuve.**

La linéarité est évidente. La bijectivité provient du théorème de Cauchy linéaire, donnant l'existence et l'unicité d'un antécédent de tout élément de $\mathbb{K}^n$. ▷

Corollaire 17.3.7 – Caractérisation des bases de (EH)

Soit $(Y_1, \dots, Y_n)$ des solutions de l'équation homogène (EH). Les propositions suivantes sont équivalentes:

(i) $(Y_1, \dots, Y_n)$ est une base de $\mathcal{S}_H$;

(ii) pour tout $t_0 \in I$, $(Y_1(t_0), \dots, Y_n(t_0))$ est une base de $\mathbb{K}^n$;

(iii) il existe $t_0 \in I$, $(Y_1(t_0), \dots, Y_n(t_0))$ est une base de $\mathbb{K}^n$;

◁ **Éléments de preuve.**

L'isomorphisme ev_{t_0} envoie base sur base, ainsi que sa réciproque. ▷

Un autre résultat qui peut se voir comme conséquence du théorème de Cauchy est une propriété importante de l'exponentielle matricielle, qu'on a déjà démontrée de façon un peu laborieuse par produit de Cauchy. La démonstration recommandée par le programme est celle ci-dessous.

Théorème 17.3.8 – Propriété remarquable de l'exponentielle matricielle

Soit A et B deux éléments de $\mathcal{M}_n(\mathbb{K})$ telles que $AB = BA$. Alors
$$\exp(A + B) = \exp(A)\exp(B) = \exp(B)\exp(A).$$

◁ **Éléments de preuve.**

Démonstration recommandée par le programme : considérer $f(t) = \exp(t(A+B))$ et $g(t) = \exp(tA)\exp(tB)$. Vérifier que f et g sont solutions du même problème de Cauchy. Il faut aussi revoir la preuve de la dérivabilité de $t \mapsto \exp(At)$ qu'on avait fait reposer sur ce théorème : on peut la voir par dérivation d'une série de fonction normalement convergente. Là encore, c'est la démonstration recommandée par le programme. ▻

III.3 Méthodes de résolution dans le cas scalaire

Nous rappelons ici les méthodes vues en MPSI pour la résolution des équations différentielles scalaires du premier ordre, de la forme

$$y' = a(t)y + b(t),$$

où a et b sont des fonctions de $\mathcal{C}^0(I, \mathbb{K})$.

Théorème 17.3.9 – Résolution de l'équation $y' = a(x)y$ (a continue)

L'ensemble des solutions de l'équation homogène $y' = a(x)y$ est :

$$\mathcal{S}_0 = \{y : x \mapsto C\mathrm{e}^{A(x)},\ C \in \mathbb{K}\},$$

où A est une primitive de la fonction (continue) a, et C est une constante.

◅ **Éléments de preuve.**
Poser $z(x) = y(x)\mathrm{e}^{-A(x)}$ et dériver z. ▻

Remarque 17.3.10 – Comment retrouver cette formule si on l'a oubliée

- Au brouillon, on s'autorise des divisions par y (rigoureusement incorrect si on n'a pas justifié que la fonction ne s'annule pas !) L'équation s'écrit alors $\frac{y'}{y} = a(x)$.
- On reconnait en $\frac{y'}{y}$ la dérivée de $\ln |y|$ (appelée dérivée logarithmique de y). On primitive, on passe à l'exponentielle et le tour est joué.
- Au propre, il est préférable d'utiliser directement la formule du cours, pour éviter les problèmes de justification issus de la division par y.

Exemples 17.3.11

1. Résolution de $y' = ay$ (a constant)
2. Résolution de $y' = yx^\alpha$ sur $\mathbb{R}$ si $\alpha \geqslant 0$, sur $\mathbb{R}_+^*$ sinon.

Méthode 17.3.12 – Recherche d'une solution particulière

- Isolez les difficultés grâce au principe de superposition.
- Demandez-vous ensuite s'il existe une solution particulière simple (constante, exponentielle, polynomiale de petit degré)...
- Dans le cas où a et constante et $b(t) = P(t)\mathrm{e}^{\lambda t}$, où P est un polynôme, on peut trouver une solution sous la même forme $Q(t)\mathrm{e}^{\lambda t}$. Plus précisément :
  * si $\lambda \neq a$, on cherche Q tel que $\deg(Q) = \deg(P)$
  * si $\lambda = a$, on cherche Q sous la forme $Q(t) = tR(t)$ ou $\deg(R) = \deg(P)$.

  On procède alors par identification des coefficients.
- Si a est constante et réelle, cette méthode s'adapte au cas où $b(t) = P(t)\cos(\lambda t)$ ou $b(t) = P(t)\sin(\lambda t)$ en passant en complexe.
- Si a et b sont polynomiales, ou admettent un développement en série entière simple, on peut aussi rechercher une solution particulière développable en série entière, sur le même principe d'identification. Si les coefficients sont des fractions rationnelles, on pourra avoir intérêt à passer au cadre polynomiale, en revenant à des équations non normalisées.

- Dans les autres situations, la méthode de variation de la constante donne une méthode systématique, qui permet d'avoir une expression d'une solution particulière sous forme intégrale.

Exemples 17.3.13

1. Résoudre $y' = 2y + \sin(x) + \mathrm{e}^x + x$ sur $\mathbb{R}$.
2. Résoudre $y' = y + x\mathrm{e}^x$.
3. Résoudre $y' = \frac{2}{x}y + \frac{x^2}{1+x^2}$ sur $\mathbb{R}_+^*$.

Méthode 17.3.14 – Méthode de variation de la constante

1. Les solutions de l'équation homogène étant de la forme $x \mapsto C\mathrm{e}^{A(x)}$, on recherche une solution particulière de l'équation non homogène sous la forme $x \mapsto C(x)\mathrm{e}^{A(x)}$ (on « rend la constante variable »)
2. En remplaçant dans l'équation différentielle, $x \mapsto C(x)$ s'obtient par primitivation :
$$C(x) = \int b(x)\mathrm{e}^{-A(x)}\, \mathrm{d}x.$$

Exemples 17.3.15

1. Retrouver les solutions des deux premières équations de l'exemple précédent par la méthode de variation de la constante.
2. Résoudre $y' = -\frac{y}{x} + \mathrm{Arctan}(x)$ sur $\mathbb{R}_+^*$.

Remarque 17.3.16

On peut retrouver le théorème de Cauchy linéaire dans le cas scalaire d'ordre 1 de façon élémentaire, grâce à l'explicitation donnée par la méthode de variation de la constante.

Pour terminer cette étude, on étudier le problème des raccordements, qui intervient lors de l'étude d'équations non normalisées.

Méthode 17.3.17 – Raccordements

Pour l'étude d'une ED $a(x)y' + b(x)y = c(x)$, où a, b, $c : I \to \mathbb{K}$ sont continues, et où a s'annule un en nombre fini des points $t_1 < \cdots < t_n$:

- considérer les intervalles $I_1, I_2, \ldots, I_{n+1}$ ouverts en $x_1, \ldots, x_n$, et dont l'union fait $I \setminus \{x_1, \ldots, x_n\}$;
- sur chacun de ces intervalles, on peut normaliser l'ED pour se ramener aux méthodes précédentes ;
- une solution sur I définissant par restriction des solutions sur chaque I_k, elle s'obtient en « recollant » les morceaux obtenus sur chaque intervalle ;
- ces recollements (ou raccordements) doivent être continus (ce qui donne en général soit des obstructions fortes, soit des relations entre les différentes constantes) ;
- ils doivent aussi être dérivables au point de raccordement : le vérifier par taux d'accroissement, par dérivée à gauche et droite, ou encore en utilisant le théorème de la limite de la dérivée, ce qui permet d'exploiter de façon efficace l'expression de l'équation différentielle ;
- vérifier rapidement que les valeurs de f et f' au point de raccordement vérifient bien l'ED.

Exemple 17.3.18

Résoudre sur $\mathbb{R}$ les équations différentielles suivantes :

1. $\sqrt{|x|}y' = y$
2. $xy' = y$.

III.4 Méthodes de résolution dans le cas vectoriel à coefficients constants

On se replace dans le contexte vectoriel $y' = a(t) \cdot y + b(t)$, et on suppose ici que $a : I \to \mathcal{L}(E)$ est constante. De façon équivalente, cela revient à étudier l'équation différentielle matricielle

$$Y' = AY + B(t),$$

où A est une matrice de $\mathcal{M}_n(\mathbb{K})$ indépendante de t.
On considère toujours I un intervalle de $\mathbb{R}$.

Théorème 17.3.19 – Résolution de l'équation homogène matricielle $Y' = AY$

Soit $A \in \mathcal{M}_n(\mathbb{K})$, $t_0 in I$ et $Y_0 \in \mathcal{M}_{n,1}(\mathbb{K})$. Le problème de Cauchy

$$Y' = AY \qquad \text{et} \qquad Y(t_0) = Y_0$$

admet comme unique solution sur I la fonction vectorielle définie par

$$\forall t \in I, \quad Y(t) = \exp((t - t_0)A)Y_0.$$

Avertissement 17.3.20

La position de Y_0 par rapport à l'exponentielle est importante. D'ailleurs, dans l'autre sens, cela pose un problème de compatibilité des formats.

Remarque 17.3.21

Donner la traduction dans le langage vectoriel de ce résultat, lorsqu'on considère l'équation $y' = a \cdot y$, où $a \in \mathcal{L}(E)$; c'est-à-dire

$$\forall t \in I, \quad y'(t) = a(y(t)).$$

Ainsi, la résolution d'une équation différentielle matricielle à coefficients constants (donc d'un système linéaire différentiel à coefficients constants) se ramène au calcul d'une exponentielle matricielle $\exp(sA)$ ($s = (t - t_0)$). On rappelle les situations favorables à ce calcul, déjà vues dans un chapitre antérieur.

Méthode 17.3.22 – Calcul de $\exp(sA)$

- Si A est diagonalisable, $A = PDP^{-1}$, avec $D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$. Alors

$$\exp(A) = P\exp(D)P^{-1} = P\,\mathrm{Diag}(\mathrm{e}^{\lambda_1}, \ldots, \mathrm{e}^{\lambda_n})P^{-1}.$$

- Ce calcul s'adapte au cas où A est diagonale par blocs.
- Si $A = D + N$, avec :
  (i) D diagonalisable
  (ii) N nilpotente
  (iii) $DN = ND$,

  alors $\exp(A) = \exp(D)\exp(N)$, et $\exp(D)$ se calcule par diagonalisation, par la méthode ci-dessus, et $\exp(N)$ est une somme finie.
- Une décomposition de Dunford, ou une jordanisation donnent une décomposition telle que ci-dessus.

Exemples 17.3.23

1. Résoudre le système
$$\begin{cases} x' &= 6x - 3y + 3z \\ y' &= 3x - 2y + 5z \\ z' &= 3x + y + 2z \end{cases}$$
2. Résoudre le système
$$\begin{cases} x' &= -x + y \\ y' &= -4x + 3y \end{cases}$$

Méthode 17.3.24 – Rappel : décomposition de Dunford

Elle est obtenue en calculant une base de chaque espace caractéristique pour former une base de E. On obtient alors une matrice diagonale par blocs, chaque bloc correspondant à un espace caractéristique. D'après le point 2 de la méthode précédente, il suffit de calculer l'exponentielle de chaque bloc. On obtient une décomposition de Dunford en retranchant λI_k, où λ est la valeur propose associé à l'espace caractéristique considéré.

IV EDL scalaires d'ordre supérieur

IV.1 Comment se ramener à l'ordre 1

On considère
$$(E)\ :\ y^{(r)} = a_{r-1}(t)y^{(r-1)} + \cdots + a_0(t)y + b(t)$$
une équation différentielle scalaire, de fonction inconnue $y \in \mathcal{D}^r(I, \mathbb{K})$, les fonctions a_k et b étant dans $\mathcal{C}^0(I, \mathbb{K})$.

Une première remarque importante, qui permet d'adapter à cette situation tous les résultats théoriques et pratiques de la partie précédente, est le fait que toute équation de ce type se ramène à une équation différentielle matricielle d'ordre 1.

Théorème 17.4.1 – Réduction à une ED matricielle d'ordre 1

Soit $y \in \mathcal{D}^r(I, K)$, et $Y : I \to \mathbb{K}^r$ telle que
$$\forall t \in I,\quad Y(t) = \begin{pmatrix} y(t) \\ y'(t) \\ \vdots \\ y^{(r-1)}(t) \end{pmatrix}.$$
Les propositions suivantes sont équivalentes:

(i) y est solution de l'équation différentielle scalaire $(E)\ :\ y^{(r)} = a_{r-1}y^{(r-1)} + \cdots + a_0 y + b$;

(ii) Y est solution de l'équation différentielle matricielle $(E')\ :\ Y' = A(t)Y + B(t)$, où
$$\forall t \in I,\quad A(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ a_0 & a_1 & \cdots & a_{r-2} & a_{r-1} \end{pmatrix} \quad \text{et} \quad B(t) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}.$$

Un problème de Cauchy associé à la fonction vectorielle Y s'écrit alors, en t_0 :

$$\begin{pmatrix} y(t_0) \\ y'(t_0) \\ \vdots \\ y^{(r-1)}(t_0) \end{pmatrix} = Y(t_0) = Y_0.$$

On définit donc un problème de Cauchy pour les équations scalaires d'ordre r de cette façon.

Définition 17.4.2 – Problème de Cauchy pour une EDL scalaire d'ordre r

Soit $t_0 \in I$.

1. Un problème de Cauchy associé à l'équation (E) est la donnée de conditions initiales imposées :
$$y(t_0) = y_0,\ y'(t_0) = y_1,\ \dots,\ y^{(r-1)}(t_0) = y_{r-1},$$
où $y_0, \dots, y_{r-1} \in \mathbb{K}$.
2. On dit que y est une solution du problème de Cauchy ci-dessus si y est solution de (E), et vérifie les conditions de Cauchy
$$y(t_0) = y_0,\ y'(t_0) = y_1,\ \dots,\ y^{(r-1)}(t_0) = y_{r-1}.$$

Via le système vérifié par Y, le théorème de Cauchy linéaire s'adapte.

Théorème 17.4.3 – Théorème de Cauchy linéaire d'ordre r

Soit $(E)\ :\ y^{(r)} = a_{r-1}(t)y^{(r-1)} + \cdots + a_0(t)y + b(t)$ une EDL d'ordre r sur I, les a_i et b étant continues.
Pour tout $(y_0, \dots, y_{r-1}) \in \mathbb{K}^r$, le problème de Cauchy

$$y(t_0) = y_0,\ y'(t_0) = y_1,\ \dots,\ y^{(r-1)}(t_0) = y_{r-1}.$$

associé à E admet une unique solution définie sur I.

Le corollaire sur l'espace des solutions de l'equation homogène

$$(EH)\ :\ y^{(r)} = a_{r-1}(t)y^{(r-1)} + \cdots + a_0(t)y$$

s'adaptent aussi.

Corollaire 17.4.4 – Dimension de l'espace des solutions de (EH)

Soit $t_0 \in I$. On note $\mathcal{S}_H$ l'espace vectoriel des solutions de l'équation homogène (EH).

1. L'application $\mathrm{ev}_{r,t_0} : y \mapsto (y(t_0), \dots, y^{(r-1)}(t_0))$ est un isomorphisme de $\mathcal{S}_H$ dans $\mathbb{K}^r$.
2. $\dim \mathcal{S}_H = r$.

$\lhd$ **Éléments de preuve.**
La linéarité est évidente. La bijectivité provient du théorème de Cauchy linéaire. $\rhd$

IV.2 Équations à coefficients constants

On considère dans ce paragraphe une équation linéaire d'ordre r de la forme

$$(E)\ :\ y^{(r)} = a_{r-1}y^{(r-1)} + \cdots + a_0 y + b(t),$$

où $a_{r-1}, \dots, a_0$ sont des constantes de $\mathbb{K}$, et $b : I \to \mathbb{K}$ est continue. On note (EH) l'équation homogène associée.
On note $\chi(X) = X^r - a_{r-1}X^{r-1} - \cdots - a_0$ le polynôme caractéristique de l'équation différentielle (E).

Théorème 17.4.5 – Résolution d'une EDL homogène à coefficients constants, HP si $r > 2$

Soit $\lambda_1, \dots, \lambda_k$ les racines 2 à 2 distinctes dans $\mathbb{C}$ de χ et $\alpha_1, \dots, \alpha_k$ leurs multiplicités.

1. Pour tout $j \in [\![1, k]\!]$, et tout $P \in \mathbb{C}_{\alpha_j - 1}[X]$, la fonction $f_j : t \mapsto P(t)\mathrm{e}^{\lambda_j t}$ est solution de (EH).
2. L'ensemble des solutions de (EH) est

$$\mathcal{S}_H = \{t \mapsto \sum_{j=1}^{k} P_j(t)\mathrm{e}^{\lambda_j t}, \quad P_j \in \mathbb{C}_{\alpha_j - 1}[X]\}.$$

◁ **Éléments de preuve.**

Cette démonstration ne sera pas faite en cours, mais fait l'objet d'un exercice du TD.

Écrire $\chi(D) = (D - \lambda_1 \mathrm{Id})^{\alpha_1} \circ \cdots \circ (D - \lambda_k \mathrm{Id})^{\alpha_k}$ pour calculer $\chi(D)(f_j)$. ▷

Proposition 17.4.6 – Solution particulière si $b(t) = P(t)\mathrm{e}^{\lambda t}$, HP si $r > 2$

On note toujours $\lambda_1, \dots, \lambda_k$ les racines 2 à 2 distinctes dans $\mathbb{C}$ de χ et $\alpha_1, \dots, \alpha_k$ leurs multiplicités. On suppose que b vérifie :

$$\forall t \in I, \quad b(t) = P(t)\mathrm{e}^{\lambda t}, \text{ avec } P \in \mathbb{C}[X] \text{ et } \lambda \in \mathbb{C}.$$

Alors il existe une solution particulière de la forme

$$y(t) = t^{\alpha} Q(t)\mathrm{e}^{\lambda t},$$

où Q est un polynôme de même degré que P et α est la multiplicité de λ comme racine de χ.

◁ **Éléments de preuve.**

Cette démonstration ne sera pas faite en cours, mais fait l'objet d'un exercice du TD.

Encore une fois, il s'agit d'écrire $\chi(D)$ sous forme factorisée, pour comprendre son effet sur les polynômes précédant l'exponentielle. ▷

IV.3 EDL d'ordre 2

On s'intéresse maintenant aux EDL scalaires d'ordre 2, à coefficients non nécessairement constants :

$$(E) \; : \; y'' = a_1(t)y' + a_0(t)y + b(t),$$

où $a_0, a_1, b : I \to \mathbb{K}$ sont des fonctions continues.

Le cas où a_1 et a_0 sont constants, et b de la forme de la proposition 17.4.6. Nous étendons ici un peu la panoplie des méthodes, afin de pourvoir résoudre une plus grande variété de situations.

Nous rappelons qu'en vertu du théorème de Cauchy, l'espace des solutions de l'équation homogène (EH) est de dimension 2, et que si $t_0 \in I$, pour tout (y_0, y_1), il existe une unique solution de (E) telle que

$$y(t_0) = y_0 \qquad \text{et} \qquad y'(t_0) = y_1.$$

On peut remarquer alors qu'en appliquant cette propriété d'unicité à (EH), si z est une solution de (EH) telle que

$$z(t_0) = z_0 \qquad \text{et} \qquad z'(t_0) = z_1,$$

alors λz est encore solution de (EH), mais répondant cette fois au problème de Cauchy suivant :

$$\lambda z(t_0) = \lambda z_0 \qquad \text{et} \qquad z'(t_0) = \lambda z_1.$$

Cet argument montre que si les deux vecteurs $\begin{pmatrix} z_0 \\ z_1 \end{pmatrix}$ et $\begin{pmatrix} z_0' \\ z_1' \end{pmatrix}$ sont liés, les deux solutions qu'elles définissent le sont aussi, et réciproquement. Cela se comprend aussi bien plus formellement par le fait que l'application qui à une solution associe les conditions de Cauchy qu'elle vérifie est un isomorphisme, donc préserve la liberté.
Afin d'étudier la liberté d'un couple de solutions, on introduit donc naturellement un objet permettant de caractériser la colinéarité des conditions de Cauchy en un point t qu'on pourra ensuite faire varier.

Définition 17.4.7 – Wronskien d'un couple de solutions de l'équation homogène

Soit y et z deux solutions de (EH). On définit leur wronskien $W_{y,z} : I \to \mathbb{K}$ par :

$$\forall t \in I, \quad W_{y,z}(t) = \begin{vmatrix} y(t) & z(t) \\ y'(t) & z'(t) \end{vmatrix}$$

Proposition 17.4.8 – Équation différentielle satisfaite par $W_{y,z}$

Quel que soit le couple (y, z) de solutions, le wronskien est solution de l'équation différentielle

$$x' = a_1(t)x.$$

Ainsi, on sait déterminer tous les wronskiens possibles, à une constante multiplicative près, par les méthodes usuelles de résolution des équations d'ordre 1. C'est cette capacité calculatoire sur le wronskien qui en donne l'intérêt.
Par définition même, il permet de caractériser la liberté d'un couple de solutions de (EH). Comme l'espace des solutions est de dimension 2, cela donne une caractérisation des bases de l'espace des solutions de (EH).

Proposition 17.4.9 – Caractérisation des bases de (EH)

Soit (y, z) un couple de solutions de (EH), et W leur wronskien. Les propositions suivantes sont équivalentes:

(i) (y, z) est une base de $\mathcal{S}_H$;

(ii) il existe $t_0 \in I$ tel que $W(t_0) \neq 0$;

(iii) pour tout $t \in I$, $W(t) \neq 0$.

◁ **Éléments de preuve.**
C'est un peu ce qu'on a dit depuis le début ! ▷

Méthode 17.4.10 – Trouver $\mathcal{S}_H$ à partir d'une solution non nulle de (EH)

Il est important de constater que la connaissance d'une solution $\varphi_0 \neq 0$ de l'équation homogène permet de trouver $\mathcal{S}_H$ tout entier. On propose pour cela deux méthodes :

1. Méthode 1 : méthode du wronskien. On se place sur un intervalle sur lequel φ_0 ne s'annule pas. Si φ est solution, alors
$$\left(\frac{\varphi}{\varphi_0}\right)' = \frac{W}{\varphi_0}^2.$$
Ainsi, W étant connu à un paramètre près, on trouve φ par primitivation, avec 2 paramètres (ce qui correspond à la dimension de l'espace des solutions).
2. Méthode 2 : variation de la constante. Toujours sur un intervalle sur lequel φ_0 ne s'annule pas, rechercher $\varphi(t)$ sous la forme $\lambda(t)\varphi_0(t)$. En injectant dans l'équation, il vient
$$\lambda''\varphi_0 + \lambda'(2\varphi_0' - a_1\varphi_0) = 0.$$

On résout l'équation en λ' puis on primitive.

Remarque 17.4.11

En fait, les deux méthodes sont complètement équivalentes. La deuxième méthode ne fait que contourner le wronskien, mais l'ED obtenue correspond à l'ED $W' = a_1 W$ pour le wronskien du couple (φ, φ_0). Ainsi, en vue d'avoir des calculs plus structurés (et de pouvoir éventuellement réutiliser le wronskien par la suite), la méthode 1 est préférable.

Exemple 17.4.12

Résoudre l'équation $y' + \frac{1}{t}y = 0$ par la méthode du wronskien.

Ainsi, la résolution d'une ED scalaire homogène d'ordre 2 se ramène le plus souvent à la recherche d'une solution particulière.

Méthode 17.4.13 – Solutions particulières de (EH)

Il n'y a malheureusement pas de méthode systématique. Il faut donc deviner, s'il en existe, des solutions simples :

- chercher parmi les fonctions usuelles, éventuellement des compositions simples de fonctions usuelles ;
- solutions polynomiales : pensez à déterminer le degré auparavant en raisonnant sur le monôme dominant, afin d'éviter l'excès de technicité ;
- solutions développables en séries entières : poser le développement et écrire les relations sur les coefficients. Essayer d'expliciter si possible et montrer que le rayon est non nul, pour avoir au moins une solution locale.

Exemple 17.4.14

1. Trouver une solution de $y'' = y' + e^{2t}y$, en essayant de deviner une solution. Résoudre ensuite l'équation.
2. Faire un changement de variable $u = e^t$ dans l'équation précédente, et retrouver la solution particulière.
3. Trouver une solution particulière de $y'' = \frac{1}{t^2+t+1}(y' + 2y)$

On termine enfin cette étude par la recherche d'une solution particulière de l'équation non homogène (E), connaissant une base (φ, ψ) de l'espace des solutions de (EH).

Proposition 17.4.15 – Variation des constantes

Soit $(E)\ :\ y'' = a_1(t)y' + a_0(t) + b(t)$, et (φ, ψ) une base de $\mathcal{S}_H$. Alors, pour que soit solution de (E), il suffit que

$$\forall t \in I,\ f(t) = \lambda(t)\varphi(t) + \mu(t)\psi(t),\ \text{avec}\ \begin{cases} \lambda'\varphi + \mu'\psi = 0 \\ \lambda'\varphi' + \mu'\psi' = 0. \end{cases}$$

Méthode 17.4.16 – Utilisation de la méthode de variation des constantes

Poser le système précédent, résoudre en λ' et μ', et intégrer.

Exemple 17.4.17

Résoudre le système $y'' + y = \tan(t)$.