Développements pour les oraux de l'agrégation externe de mathématiques

Romain GIUGE

Année 2012-2013

Ce document regroupe les développements que j'ai préparé pour les oraux de l'agrégation externe de mathématiques. Mon travail n'a pas été de me restreindre aux développements proprement dits. J'ai voulu aller plus loin en ajoutant à la suite de chaque développement des compléments pouvant concerner, selon les cas, des rappels de définitions et de théorèmes utilisés dans le développement, des ajouts de remarques, d'applications et d'exemples liés au développement. Le but recherché est que chaque développement se suffise à lui-même autant que possible, en permettant d'élargir un peu son champ et éventuellement d'anticiper des questions du jury. Un exemple parmi tant d'autres : dans le développement *"Densité des fonctions continues et nulle part dérivables"* sont donnés des exemples de fonctions continues et nulle part dérivables qu'il est bon d'avoir en tête lorsqu'on le présente.

Pour chaque développement, la partie à présenter s'arrête à la première barre horizontale. Les divers compléments se trouvent en-dessous, et à la fin une ou plusieurs références sont données dans la mesure du possible. Tous mes développements ont été relus de nombreuses fois au cours de l'année, j'espère donc que les coquilles et erreurs soient maintenant presque éradiquées. Il va néanmoins sans dire qu'un regard critique doit toujours être porté sur tout ce qui est écrit ici.

Je donne ma répartition de mes développements dans le chapitre "Couplages". Cette répartition est également sujette à des révisions. Les vaguelettes indiquent des choix très certainement douteux, et les développements barrés sont des développements qui pourraient aller mais dont je me suis finalement passé. Ces développements sont alors placés dans le chapitre "Développements bonus".

J'adresse pour finir un grand remerciement à notre formidable club agrég de Lyon pour cette super année de préparation : merci à Donatien Bénéat pour toutes ses méthodes bizarres sorties de son esprit ou de livres que personne n'a jamais pensé à emprunter ; à Abdelhakc Yakoub pour son rôle de délégué parfaitement réalisé avec une belle organisation des évènements de notre promo de l'université Lyon 1 ; à Nicolas Doyen pour notre course aux développements, pour ses 74 développements uniquement d'algèbre dont on n'oubliera pas la longueur interminable de chacun, et pour ses réponses précises à mes questions d'algèbre ; à Sa Sainte Omniscience Guillaume Delon pour ses corrections, ses éclaircissements des documents de Nicolas, et ses nombreux chocolats ; à Simon Boyer pour sa solution qui nous a tous fait gagner des points aux écrits et son talent pour des présentations de leçons pédagogiques, bien qu'interminables ; à Chloé Bourquard pour toutes les fois où elle a bien voulu manger avec nous et pour les leçons pénibles qu'elle a dû présenter dans l'année mais dont l'utilité a été certaine. Et merci à toute notre promo pour l'ambiance formidable qu'on a eu cette année !

Table des matières

Chapitre 1

Couplages

1.1 Leçons d'algèbre

1. 101 - Groupe opérant sur un ensemble. Exemples et applications.

- ➢ Théorème de Wedderburn.
- ➢ Topologie des orbites de l'action de Steinitz.
- ➢ ~~Théorème de Brauer.~~

2. 102 - Groupe des nombres complexes de module 1. Sous-groupes des racines de l'unité. Applications.

- ➢ Irréductibilité des polynômes cyclotomiques sur $\mathbb{Q}$.
- ➢ Dénombrement des solutions d'une équation diophantienne.
- ➢ Théorème de Kronecker.

3. 103 - Exemples et applications des notions de sous-groupe distingué et de groupe quotient.

- ➢ Simplicité de A_n.
- ➢ Théorème de Frobenius-Zolotarev.

4. 104 - Groupes finis. Exemples et applications.

- ➢ Simplicité de A_n.
- ➢ Théorème de Frobenius-Zolotarev.
- ➢ Théorème de Burnside.

5. 105 - Groupe des permutations d'un ensemble fini. Applications.

- ➢ Simplicité de A_n.
- ➢ Théorème de Frobenius-Zolotarev.
- ➢ ~~Théorème de Brauer.~~

6. 106 - Groupe linéaire d'un espace vectoriel de dimension finie E, sous-groupes de $\mathrm{GL}(E)$. Applications.

- ➢ Décomposition polaire dans $\mathrm{GL}_n(\mathbb{C})$.
- ➢ Théorème de Frobenius-Zolotarev.
- ➢ Théorème de Burnside.
- ➢ ~~Théorème de Brauer.~~

7. 107 - Représentations et caractères des groupes finis sur un espace vectoriel complexe.

- ➢ ✗
- ➢ ✗

8. 108 - Exemples de parties génératrices d'un groupe. Applications.

- ➢ Théorème de Frobenius-Zolotarev.
- ➢ Théorème de Cartan-Dieudonné.

9. 109 - Représentations des groupes finis de petit cardinal.

- ➢ ✗
- ➢ ✗

10. 120 - Anneau $\mathbb{Z}/n\mathbb{Z}$. Applications.

- ➢ Irréductibilité des polynômes cyclotomiques sur $\mathbb{Q}$.
- ➢ Théorème des deux carrés.

11. 121 - Nombres premiers. Applications.

- ➢ Irréductibilité des polynômes cyclotomiques sur $\mathbb{Q}$.
- ➢ Théorème des deux carrés.

12. 122 - Anneaux principaux. Applications.

- ➢ Théorème des deux carrés.
- ➢ Algorithme pour le calcul des facteurs invariants.

13. 123 - Corps finis. Applications.

- ➢ Dénombrement des polynômes irréductibles sur un corps fini.
- ➢ Théorème de Chevalley-Warning.
- ➢ Théorème de Wedderburn.

14. 124 - Anneau des séries formelles. Applications.

- ➢ Dénombrement des solutions d'une équation diophantienne.
- ➢ Nombres de Bell (la version adaptée).

15. 125 - Extensions de corps. Exemples et applications.

- ➢ Dénombrement des polynômes irréductibles sur un corps fini.
- ➢ Théorème de Wantzel.
- ➢ Théorème de Rothstein-Trager.

16. 140 - Corps des fractions rationnelles à une indéterminée sur un corps commutatif. Applications.

- ➢ Dénombrement des solutions d'une équation diophantienne.
- ➢ Théorème de Rothstein-Trager. (décomposition en éléments simples, primitives de fractions rationnelles)

17. 141 - Polynômes irréductibles à une indéterminée. Corps de rupture. Exemples et applications.

- ➢ Dénombrement des polynômes irréductibles sur un corps fini.
- ➢ Irréductibilité des polynômes cyclotomiques sur $\mathbb{Q}$.

18. 142 - Algèbre des polynômes à n indéterminées. Applications.

- ➢ Théorème de Chevalley-Warning.
- ➢ Théorème de Kronecker.

19. 143 - Résultant. Applications.

- ➢ Théorème de Kronecker.
- ➢ Théorème de Rothstein-Trager.

20. 144 - Racines d'un polynômes, fonctions symétriques élémentaires. Localisation des racines dans les cas réel et complexe.

- ➢ Comptage de racines par les formes quadratiques.
- ➢ Théorème de Kronecker.
- ➢ Ellipse de Steiner et application.

21. 150 - Exemples d'actions de groupes sur les espaces de matrices.

- ➢ Décomposition polaire dans $\mathrm{GL}_n(\mathbb{C})$ (l'énoncé adapté).
- ➢ Topologie des orbites de l'action de Steinitz.
- ➢ ~~Théorème de Brauer (l'énoncé adapté).~~

22. 151 - Dimension d'un espace vectoriel (on se limitera au cas de dimension finie). Rang. Exemples et applications.

- ➢ Théorème des extrema liés.
- ➢ Comptage de racines par les formes quadratiques.
- ➢ ~~Sous-espaces de dimension finie de $\mathcal{C}(\mathbb{R}, \mathbb{C})$ stables par translations.~~

23. 152 - Déterminant. Exemples et applications.

- ➢ Théorème de Frobenius-Zolotarev.
- ➢ Ellipsoïde de John-Loewner.
- ➢ Topologie des orbites de l'action de Steinitz.
- ➢ ~~Théorème de Brauer.~~

24. 153 - Polynômes d'endomorphismes en dimension finie. Application à la réduction d'un endomorphisme en dimension finie.

- ➢ Décomposition de Dunford.
- ➢ Théorème de Burnside.
- ➢ Décomposition polaire dans $\mathrm{GL}_n(\mathbb{C})$.

- ➢ Surjectivité de l'exponentielle.

25. 154 - Sous-espaces stables d'un endomorphisme ou d'une famille d'endomorphismes en dimension finie. Applications.

- ➢ Réduction des isométries d'un espace euclidien.
- ➢ Décomposition de Dunford.
- ➢ ~~Sous-espaces de dimension finie de $\mathcal{C}(\mathbb{R},\mathbb{C})$ stables par translations (stabilité par l'endomorphisme de dérivation).~~

26. 155 - Endomorphismes diagonalisables en dimension finie.

- ➢ Théorème de Burnside.
- ➢ Décomposition de Dunford.

27. 156 - Exponentielle de matrices. Applications.

- ➢ Surjectivité de l'exponentielle.
- ➢ Théorème de Liapounov.
- ➢ ~~Différentielle d'une limite et application exponentielle.~~

28. 157 - Endomorphismes trigonalisables. Endomorphismes nilpotents.

- ➢ Théorème de Burnside.
- ➢ Décomposition de Dunford.

29. 158 - Matrices symétriques réelles, matrices hermitiennes.

- ➢ Méthode de Gauss-Seidel.
- ➢ Lemme de Morse.
- ➢ Décomposition polaire dans $\mathrm{GL}_n(\mathbb{C})$.
- ➢ ~~Points extrémaux de la boule unité de $\mathcal{M}_n(\mathbb{R})$.~~

30. 159 - Formes linéaires et hyperplans en dimension finie. Exemples et applications.

- ➢ Théorème de Cartan-Dieudonné.
- ➢ Théorème de Krein-Milman.
- ➢ Comptage de racines par les formes quadratiques.
- ➢ Théorème des extrema liés.

31. 160 - Endomorphismes remarquables d'un espace vectoriel euclidien de dimension finie.

- ➢ Théorème de Cartan-Dieudonné.
- ➢ Décomposition polaire dans $\mathrm{GL}_n(\mathbb{C})$.
- ➢ Réduction des isométries d'un espace euclidien.

32. 161 - Isométries d'un espace affine euclidien de dimension finie. Formes réduites. Applications en dimensions 2 et 3.

- ➢ Réduction des isométries d'un espace euclidien.
- ➢ Théorème de Cartan-Dieudonné.
- ➢ ~~Points extrémaux de la boule unité de $\mathcal{M}_n(\mathbb{R})$.~~

33. 162 - Systèmes d'équations linéaires. Opérations, aspects algorithmiques et conséquences théoriques.

- ➢ Méthode de Gauss-Seidel.
- ➢ Décomposition QR.

34. 170 - Formes quadratiques sur un espace vectoriel de dimension finie. Orthogonalité, isotropie. Applications.

- ➢ Comptage de racines par les formes quadratiques.
- ➢ Théorème de Liapounov.
- ➢ Lemme de Morse.

35. 171 - Formes quadratiques réelles. Exemples et applications.

- ➢ Comptage de racines par les formes quadratiques.
- ➢ Théorème de Liapounov.
- ➢ Lemme de Morse.

36. 180 - Coniques. Applications.

- ➢ Ellipse de Steiner et application.
- ➢ Théorème de Pascal.

37. 181 - Barycentres dans un espace affine réel de dimension finie, convexité. Applications.

- ➢ Sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$.
- ➢ Théorème de Krein-Milman.
- ➢ ~~Points extrémaux de la boule unité de $\mathcal{M}_n(\mathbb{R})$.~~

38. 182 - Application des nombres complexes à la géométrie.

- ➢ Ellipse de Steiner et application.
- ➢ Quelques propriétés des homographies.

39. 183 - Utilisation des groupes en géométrie.

- ➢ Quelques propriétés des homographies.
- ➢ Ellipse de Steiner et application (à adapter avec l'action du groupe affine).

40. 190 - Méthodes combinatoires, problèmes de dénombrement.

- ➢ Dénombrement des solutions d'une équation diophantienne.
- ➢ Dénombrement des polynômes irréductibles sur un corps fini.
- ➢ Nombres de Bell.
- ➢ ~~Théorème de Brauer.~~

1.2 Leçons d'analyse

1. 201 - Espaces de fonctions. Exemples et applications.
- ➢ Densité des fonctions continues et nulle part dérivables.
- ➢ Complétude de L^p.

2. 202 - Exemples de parties denses et applications.
- ➢ Densité des fonctions continues et nulle part dérivables.
- ➢ Densité des polynômes orthogonaux.
- ➢ Sous-espaces vectoriels fermés de L^p (application de $\mathbb{C}^n$ est séparable).

3. 203 - Utilisation de la notion de compacité.
- ➢ Sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$.
- ➢ Théorème de Brouwer.
- ➢ Décomposition polaire dans $\mathrm{GL}_n(\mathbb{C})$.

4. 204 - Connexité. Exemples et applications.
- ➢ Surjectivité de l'exponentielle.
- ➢ Théorème de Brouwer.

5. 205 - Espaces complets. Exemples et applications.
- ➢ Complétude de L^p.
- ➢ Densité des fonctions continues et nulle part dérivables (utilisation du théorème de Baire).

6. 206 - Théorèmes de point fixe. Exemples et applications.
- ➢ Sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$.
- ➢ Théorème de Brouwer.

7. 207 - Prolongement de fonctions. Exemples et applications.
- ➢ Densité des polynômes orthogonaux (prolongement analytique).
- ➢ Maximalité et globalité des solutions de $y' = f(t, y)$.

8. 208 - Espaces vectoriels normés, applications linéaires continues. Exemples.
- ➢ Sous-espaces vectoriels fermés de L^p.

- ➢ Projection sur un convexe fermé et représentation de Riesz.

9. 213 - Espaces de Hilbert, bases hilbertiennes. Exemples et applications.

- ➢ Densité des polynômes orthogonaux.
- ➢ Projection sur un convexe fermé et représentation de Riesz.

10. 214 - Théorème d'inversion locale, théorème des fonctions implicites. Exemples et applications.

- ➢ Théorème de Brouwer.
- ➢ Surjectivité de l'exponentielle.
- ➢ Théorème des extrema liés.
- ➢ Lemme de Morse.

11. 215 - Applications différentiables définies sur un ouvert de $\mathbb{R}^n$. Exemples et applications.

- ➢ Théorème de Brouwer.
- ➢ Lemme de Morse.
- ➢ Théorème de Liapounov.
- ➢ ~~Différentielle d'une limite et application exponentielle.~~

12. 216 - Etude métrique des courbes. Exemples.

- ➢ Etude du folium de Descartes.
- ➢ Théorème des quatre sommets.

13. 217 - Sous-variétés de $\mathbb{R}^n$, exemples.

- ➢ Etude du folium de Descartes.
- ➢ Théorème des extrema liés.

14. 218 - Applications des formules de Taylor.

- ➢ Méthode de Newton.
- ➢ Lemme de Morse.

15. 219 - Problèmes d'extremum.

- ➢ Théorème des extrema liés.
- ➢ Ellipsoïde de John-Loewner.

16. 220 - Equations différentielles $X' = f(t, X)$, exemples d'études qualitatives des solutions.

- ➢ Théorème de Liapounov.
- ➢ Maximalité et globalité des solutions de $y' = f(t, y)$.

17. 221 - Equations différentielles linéaires. Systèmes d'équations différentielles linéaires. Exemples et applications.

- ➢ Théorème de Liapounov.
- ➢ Comportement des solutions d'une équation différentielle linéaire.
- ➢ ~~Sous-espaces de dimension finie de $\mathcal{C}(\mathbb{R}, \mathbb{C})$ stables par translations.~~

18. 223 - Convergence des suites numériques. Exemples et applications.

- ➢ Méthode de Newton.
- ➢ Preuve probabiliste de la formule de Stirling.

19. 224 - Comportement asymptotique des suites numériques. Rapidité de convergence. Exemples.

- ➢ Méthode de Newton.
- ➢ Preuve probabiliste de la formule de Stirling.

20. 226 - Comportement d'une suite réelle ou vectorielle définie par une itération $u_{n+1} = f(u_n)$. Exemples.

- ➢ Méthode de Newton.
- ➢ Méthode de Gauss-Seidel.

21. 228 - Continuité et dérivabilité des fonctions réelles d'une variable réelle. Exemples et contre-exemples.

- ➢ Densité des fonctions continues et nulle part dérivables.
- ➢ Théorème de Borel.

22. 229 - Fonctions monotones. Fonctions convexes. Exemples et applications.

- ➢ Méthode de Newton.
- ➢ Ellipsoïde de John-Loewner.

23. 230 - Séries de nombres réels ou complexes. Comportement des restes ou des sommes partielles des séries numériques. Exemples.

- ➢ Théorème de Hardy-Littlewood.
- ➢ Nombres de Bell.

24. 232 - Méthodes d'approximation des solutions d'une équation $F(X) = 0$. Exemples.

- ➢ Méthode de Newton.
- ➢ Méthode de Gauss-Seidel.

25. 234 - Espaces L^p.

- ➢ Complétude de L^p.

- ➢ Sous-espaces vectoriels fermés de L^p.

26. 235 - Suites et séries de fonctions intégrables. Exemples et applications.

- ➢ Calcul d'une intégrale.
- ➢ Complétude de L^p.
- ➢ Preuve probabiliste de la formule de Stirling.

27. 236 - Illustrer par des exemples quelques méthodes de calcul d'intégrales de fonctions d'une ou plusieurs variables réelles.

- ➢ Calcul d'une intégrale.
- ➢ La formule des compléments.

28. 239 - Fonctions définies par une intégrale dépendant d'un paramètre. Exemples et applications.

- ➢ Calcul d'une intégrale.
- ➢ Densité des polynômes orthogonaux.

29. 240 - Transformation de Fourier, produit de convolution. Applications.

- ➢ Densité des polynômes orthogonaux.
- ➢ Formule sommatoire de Poisson et application.

30. 241 - Suites et séries de fonctions. Exemples et contre-exemples.

- ➢ Théorème de Borel.
- ➢ Equation de la chaleur.
- ➢ Preuve probabiliste de la formule de Stirling.
- ➢ ~~Différentielle d'une limite et application exponentielle.~~

31. 243 - Convergence des séries entières, propriétés de la somme. Exemples et applications.

- ➢ Théorème de Hardy-Littlewood.
- ➢ Nombres de Bell.
- ➢ Dénombrement des solutions d'une équation diophantienne (à adapter avec les séries entières).

32. 245 - Fonctions holomorphes et méromorphes sur un ouvert de $\mathbb{C}$.

- ➢ La formule des compléments.
- ➢ Densité des polynômes orthogonaux.

33. 246 - Séries de Fourier. Exemples et applications.

- ➢ Equation de la chaleur.
- ➢ Formule sommatoire de Poisson et application.

34. 247 - Exemples de problèmes d'interversion de limites.

- ➢ Calcul d'une intégrale.
- ➢ Théorème de Hardy-Littlewood.
- ➢ ~~Différentielle d'une limite et application exponentielle.~~

35. 249 - Suites de variables de Bernoulli indépendantes.

- ➢ Théorème de Le Cam.
- ➢ Polynômes de Bernstein.

36. 250 - Loi des grands nombres, théorème de la limite centrale. Applications.

- ➢ Nombres normaux.
- ➢ Preuve probabiliste de la formule de Stirling.

37. 251 - Indépendance d'événements et de variables aléatoires. Exemples.

- ➢ Théorème de Le Cam.
- ➢ Nombres normaux.

38. 252 - Loi binomiale, loi de Poisson. Applications.

- ➢ Théorème de Le Cam.
- ➢ Polynômes de Bernstein.
- ➢ Preuve probabiliste de la formule de Stirling.

39. 253 - Utilisation de la notion de convexité en analyse.

- ➢ Méthode de Newton.
- ➢ Ellipsoïde de John-Loewner.

40. 254 - Espace de Schwartz et distributions tempérées. Transformation de Fourier dans $\mathcal{S}(\mathbb{R}^d)$ et $\mathcal{S}'(\mathbb{R}^d)$.

- ➢ Formule sommatoire de Poisson et application.
- ➢ ✗

41. 255 - Espaces de Schwartz. Distributions. Dérivation au sens des distributions.

- ➢ Formule sommatoire de Poisson et application.
- ➢ ✗

Chapitre 2

Développements d'algèbre

2.1 Algorithme pour le calcul des facteurs invariants

Théorème. *Soit A un anneau principal. Soit $U \in \mathcal{M}_{m\times n}(A)$. Alors il existe une famille $(d_1, \dots, d_s)$ d'éléments non nuls de A vérifiant $d_1 | \dots | d_s$ et telle que U soit équivalente à la matrice $D = \mathrm{Diag}(d_1, \dots, d_s, 0, \dots, 0) \in \mathcal{M}_{m\times n}(A)$, i.e. il existe $(P, Q) \in \mathrm{GL}_m(A) \times \mathrm{GL}_n(A)$ tel que $U = PDQ$. Les d_i sont uniques à des inversibles près, ils sont appelés facteurs invariants de la matrice U.*

Nous allons donner une preuve algorithmique de ce théorème dans le cas où A est un anneau euclidien. On note φ le stathme de A euclidien. Pour alléger les notations, on note constamment $U = (u_{i,j})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ la matrice modifiée à chaque étape. Sa i-ième ligne est notée L_i et sa j-ième colonne est notée C_j.

Etape 1. Si $U = 0$, fin de l'algorithme.

Etape 2. Sinon, soit (i_0, j_0) tel que $\varphi(u_{i_0,j_0}) = \min\{\varphi(u_{i,j}), u_{i,j} \neq 0\}$. Permuter les colonnes C_1 et C_{j_0} puis les lignes L_1 et L_{i_0} afin de placer u_{i_0,j_0} en haut à gauche de U.

Etape 3. Traitement de la première colonne. Soit $i = 2$.

Etape 3.a. Effectuer la division euclidienne de $u_{i,1}$ par $u_{1,1}$:

$$u_{i,1} = u_{1,1}q + r_i \text{ avec } r_i = 0 \text{ ou } \varphi(r_i) < \varphi(u_{1,1}).$$

Soustraire q fois la ligne L_1 à la ligne L_i pour obtenir $u_{i,1} = r_i$.

Etape 3.b. Si $r_i \neq 0$, échanger les lignes L_i et L_1 et retourner à **Etape 3.a.**

Etape 3.c. Si $r_i = 0$ et $i < m$, passer à la ligne suivante : $i := i + 1$ et aller à **Etape 3.a.**

Etape 3.d. Si $r_i = 0$ et $i = m$, aller à **Etape 4.**

Etape 4. Traitement de la première ligne. Soit $j = 2$.

Etape 4.a. Effectuer la division euclidienne de $u_{1,j}$ par $u_{1,1}$:

$$u_{1,j} = u_{1,1}q + s_j \text{ avec } s_j = 0 \text{ ou } \varphi(s_j) < \varphi(u_{1,1}).$$

Soustraire q fois la colonne C_1 à la colonne C_j pour obtenir $u_{1,j} = s_j$.

Etape 4.b. Si $s_j \neq 0$, échanger les colonnes C_j et C_1 et retourner à **Etape 3.**

Etape 4.c. Si $s_j = 0$ et $j < n$, passer à la colonne suivante : $j := j + 1$ et aller à **Etape 4.a.**

Etape 4.d. Si $s_j = 0$ et $j = n$, aller à **Etape 5.**

Etape 5. Divisibilité.

Etape 5.a. S'il existe $i_1 \geqslant 2$ et $j_1 \geqslant 2$ tels que $u_{1,1}$ ne divise pas u_{i_1,j_1}, ajouter la colonne C_{j_1} à la colonne C_1 et retourner à **Etape 3.**

Etape 5.b. Sinon, retourner à **Etape 1.** avec la matrice extraite $U = (u_{i,j})_{\substack{2\leqslant i\leqslant m\\ 2\leqslant j\leqslant n}}$.

Comme on revient souvent en arrière, il n'est pas garanti que l'agorithme termine. Cependant, il terminera car à chaque retour en arrière, $\varphi(u_{1,1})$ décroît d'au moins une unité, et comme φ est à valeur dans $\mathbb{N}$, l'algorithme s'arrêtera au bout d'un nombre fini d'étapes.

A chaque étape de l'algorithme, les opérations effectuées sont des opérations élémentaires sur les lignes et les colonnes, d'où le résultat après traduction en termes de matrices.

Unicité des d_i :

Pour $U \in \mathcal{M}_{m\times n}(A)$, on note

$$\Lambda_j(U) = \mathrm{PGCD}(\Delta_j,\ \Delta_j \text{ mineur de taille } j \text{ de } U).$$

Par convention, on pose $\Lambda_0(U) = 1$.

On note $d_1, \dots, d_s$ les facteurs invariants obtenus avec l'algorithme et on va montrer qu'ils sont uniques à des inversibles près. On a donc $U = PDQ$ avec $D = \mathrm{Diag}(d_1, \dots, d_s, 0, \dots, 0)$. De manière évidente, on a $\Lambda_j(D) = d_1 \dots d_j$. Par la proposition suivante, pour tout $j \leqslant s$, il existe $a_j \in A^*$ tel que $\Lambda_j(D) = d_1 \dots d_j = a_j\Lambda_j(U)$. Ainsi

$$d_j = \frac{a_j\Lambda_j(U)}{a_{j-1}\Lambda_{j-1}(U)} = \underbrace{\left(\frac{a_j}{a_{j-1}}\right)}_{\in A^*} \frac{\Lambda_j(U)}{\Lambda_{j-1}(U)},$$

ce qui prouve l'unicité à inversibles près.

Proposition. *Soient $U, U' \in \mathcal{M}_{m\times n}(A)$ deux matrices équivalentes. Alors pour tout $j \in [\![1, \min(m,n)]\!]$, on a*

$$\langle \Lambda_j(U) \rangle = \langle \Lambda_j(U') \rangle.$$

Démonstration. Supposons d'abord que $U = PU'$ avec $P \in \mathrm{GL}_m(A)$. Alors les lignes de U sont combinaisons linéaires des lignes de U'. Par multilinéarité du déterminant, un mineur de taille j de U est combinaison linéaire de mineurs de taille j de U', et ainsi $\langle \Lambda_j(U) \rangle \subset \langle \Lambda_j(U') \rangle$. Comme on peut aussi écrire $U' = P^{-1}U$, on obtient de même $\langle \Lambda_j(U') \rangle \subset \langle \Lambda_j(U) \rangle$, d'où égalité.

De la même manière, si $U = U'Q$ avec $Q \in \mathrm{GL}_n(A)$, on a $\langle \Lambda_j(U') \rangle = \langle \Lambda_j(U) \rangle$ (cette fois les colonnes de U sont combinaison linéaires des colonnes de U').

En rassemblant, si $U = PU'Q$, alors $\langle \Lambda_j(U) \rangle = \langle \Lambda_j(U') \rangle$. □

Remarque. Les corps sont des anneaux euclidiens. En effet, la division euclidienne dans un corps K s'écrit : pour tout $(x, y) \in K \times (K\backslash\{0\})$, $x = (xy^{-1})y + 0$. Le reste est toujours nul, donc on peut choisir n'importe quelle stathme φ.

Ainsi, lorsque A est un corps, l'algorithme aboutit à la matrice J_r : en réalité, on a n'importe quel nombre non nul sur la diagonale pour $i \leqslant s$, et comme les d_i sont donnés à des inversibles près, on peut mettre 1. L'algorithme fournit ainsi une méthode de calcul du rang par opérations élémentaires.

Enfin, la proposition donne la caractérisation du rang par les mineurs extraits : le rang d'une matrice est égal au plus grand ordre d'un mineur non nul de cette matrice.

Corollaire (Théorème de la base adaptée)**.** *Soient A un anneau principal et M un A-module libre de rang n. Si N est un sous-module de M, il existe une base $(e_1, \dots, e_n)$ de M et des scalaires non nuls $d_1, \dots, d_s$, uniques à inversibles près, vérifiant $d_1 | \dots | d_s$ et tels que la famille $(d_1 e_1, \dots, d_s e_s)$ soit une base de N.*

Démonstration. On admet qu'un sous-module d'un module libre de rang fini est également libre de rang fini : ça n'est pas immédiat, voir page 291 de *Objectif agrégation* (Beck, Malick, Peyré).

Soient alors $(v_1, \dots, v_s)$ une base de N et $(u_1, \dots, u_n)$ une base de M. On note U la matrice dans les bases (v_i) et (u_i) de l'injection canonique de N dans M.

D'après le théorème des facteurs invariants, U est équivalente à

$$U' = \mathrm{Diag}(d_1, \dots, d_r, 0, \dots, 0)$$

avec $d_1 | \dots | d_r$. Par conséquent, il existe une base $(e_1, \dots, e_n)$ de M et une base $(f_1, \dots, f_s)$ de N telles que $\mathrm{Id}(f_i) = d_i e_i$ pour tout $i \leqslant r$ et $\mathrm{Id}(f_i) = 0$ pour tout $i > r$. Comme $\mathrm{Id}(f_i) \neq 0$, on a nécessairement $r = s$. □

Corollaire (Théorème de structure)**.** *Soient A un anneau principal et M un A-module de type fini. Alors il existe un unique couple (r, s) d'entiers et $d_1 | \dots | d_s$ des éléments de A, uniques à inversibles près, tels que*

$$M \simeq A^r \oplus \left(\bigoplus_{i=1}^{s} A/(d_i) \right).$$

Les d_i sont appelés facteurs invariants du module M.

Démonstration. Comme M est de type fini, on a un morphisme surjectif $\varphi : A^n \to M$ avec $n \in \mathbb{N}$. D'après le corollaire précédent, il existe une base $(e_1, \dots, e_n)$ de A^n et $d_1 | \dots | d_s$ des élements de A tels que $(d_1 e_1, \dots, d_s e_s)$ soit une base de $\mathrm{Ker}(\varphi)$. Comme $M = \mathrm{Im}(\varphi) \simeq A^n / \mathrm{Ker}(\varphi)$, on obtient

$$M \simeq \left(\bigoplus_{i=1}^{n} Ae_i \right) / \left(\bigoplus_{i=1}^{s} Ad_i e_i \right) \simeq \underbrace{\left(\bigoplus_{i=1}^{s} Ae_i / Ad_i e_i \right)}_{\simeq \left(\bigoplus_{i=1}^{s} A/(d_i) \right)} \oplus \underbrace{\left(\bigoplus_{i=s+1}^{n} Ae_i \right)}_{\simeq A^{n-s}}.$$

□

Enonçons un théorème donnant une méthode de calcul des invariants de similitude d'une matrice :

Théorème. *Soient E un K-espace vectoriel de dimension finie n, $u \in \mathcal{L}(E)$, $\mathcal{B}$ une base de E et U la matrice de u dans la base $\mathcal{B}$. Les invariants de similitude de u sont les facteurs invariants non inversibles de la matrice $U - XI_n \in \mathcal{M}_n(K[X])$.*

Application. *Calculer les invariants de similitude de la matrice*

$$V = \begin{pmatrix} 3 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \in \mathcal{M}_3(\mathbb{Q}).$$

Par le théorème précédent, il nous suffit de calculer les facteurs invariants de la matrice $U = V - XI_3$. Pour cela, d'après ce qui précède, on a deux méthodes : appliquer l'algorithme ou calculer les PGCD des mineurs de taille $1, 2$ et 3 de U. Appliquer l'algorithme est très long et fastidieux, mais à l'issue, on aura également les matrices P et Q : par exemple P est le produit des matrices des opérations élémentaires qu'on aura effectué sur les lignes.

On va se contenter de calculer les PGCD des mineurs. On a directement $\Lambda_1(U) = 1$. Le calcul plutôt simple des 9 mineurs de taille 2 donne $\Lambda_2(U) = X - 1$. Enfin, le calcul du déterminant de U donne $\Lambda_3(U) = (X-1)^3$. Finalement, on obtient

$$\begin{cases} d_1 = \Lambda_1(U) = 1 \\ d_2 = \frac{\Lambda_2(U)}{\Lambda_1(U)} = X - 1 \\ d_3 = \frac{\Lambda_3(U)}{\Lambda_2(U)} = (X-1)^2, \end{cases}$$

en se rappelant que les d_i sont donnés à des inversibles près.

Finalement, les invariants de similitude de V sont $P_1 = X - 1$ et $P_2 = (X-1)^2$.

Références :

- Beck, Malick, Peyré - *Objectif agrégation* - Page 285 (algorithme), page 291 (théorème de la base adaptée), page 278 (théorème de structure), page 301 (théorème pour le calcul des invariants de similitude), page 319 (exercice de mise en pratique).

2.2 Comptage de racines par les formes quadratiques

Soit $P \in \mathbb{R}[X]$ un polynôme de degré n dont on note $\alpha_1, \dots, \alpha_n$ les racines complexes (comptées avec multiplicité). Pour $i \in \mathbb{N}$, on note

$$s_i = \sum_{k=1}^{n} \alpha_k^i.$$

Les s_i sont réels : en effet, on peut les calculer avec les formules de Newton qui donneront une expression en fonction des polynômes symétriques élémentaires, eux-mêmes fonction des coefficients de P.

Théorème. *Soit Q la forme quadratique définie pour $x = (x_0, \dots, x_{n-1}) \in \mathbb{R}^n$ par*

$$Q(x) = \sum_{0 \leqslant i,j \leqslant n-1} s_{i+j} x_i x_j.$$

On note (s,t) la signature de Q. Alors le nombre de racines réelles distinctes de P vaut $s-t$ et son nombre de racines complexes distinctes $s+t$ (i.e. le rang de Q).

Démonstration. On a :

$$Q(x) = \sum_{0 \leqslant i,j \leqslant n-1} s_{i+j} x_i x_j = \sum_{k=1}^{n} \sum_{0 \leqslant i,j \leqslant n-1} \alpha_k^{i+j} x_i x_j.$$

En notant $l_k(x) = \sum_{i=0}^{n-1} \alpha_k^i x_i$, on obtient :

$$Q(x) = \sum_{k=1}^{n} \left(\left(\sum_{i=0}^{n-1} \alpha_k^i x_i \right) \left(\sum_{j=0}^{n-1} \alpha_k^j x_j \right) \right) = \sum_{k=1}^{n} l_k(x)^2.$$

Quitte à réordonner, on peut supposer que $\alpha_1, \dots, \alpha_r$ sont les racines complexes distinctes de P dont on note $m_1, \dots, m_r$ les multiplicités. Ainsi :

$$Q = \sum_{k=1}^{r} m_k l_k^2.$$

Les formes linéaires l_k pour $k \in [\![1, r]\!]$ sont linéairement indépendantes sur $\mathbb{C}$. En effet, on écrit le déterminant de leurs r premières coordonnées dans la base duale $(e_0^*, \dots, e_{n-1}^*)$ de la base canonique de $\mathbb{C}^n$, sachant que $l_k = \sum_{i=0}^{n-1} \alpha_k^i e_i^*$:

$$\det \begin{pmatrix} 1 & \cdots & 1 \\ \alpha_1 & \cdots & \alpha_r \\ \vdots & \cdots & \vdots \\ \alpha_1^{r-1} & \cdots & \alpha_r^{r-1} \end{pmatrix}.$$

On reconnaît le déterminant de Vandermonde associé aux scalaires $\alpha_1, \dots, \alpha_r$, qui est non nul car les α_i sont deux à deux distincts.

On remarque maintenant que si α_k n'est pas réel, alors $\overline{\alpha_k}$ fait aussi partie des racines de P et est différent de α_k. Quitte à réordonner, on suppose que $\alpha_1, \dots, \alpha_p$ sont les racines réelles distinctes de P, et $\alpha_{p+1}, \overline{\alpha_{p+1}}, \dots, \alpha_{p+c}, \overline{\alpha_{p+c}}$ ses racines complexes distinctes (on a alors $r = p + 2c$). Il vient :

$$Q = \sum_{k=1}^{r} m_k l_k^2 = \sum_{k=1}^{p} m_k l_k^2 + \sum_{k=p+1}^{p+c} m_k \left(l_k^2 + \overline{l_k}^2 \right).$$

On note $v_k = \frac{l_k + \overline{l_k}}{2}$ et $w_k = \frac{l_k - \overline{l_k}}{2i}$ de sorte que les coefficients de v_k et w_k soient réels. On a alors :

$$v_k^2 - w_k^2 = \frac{l_k^2 + \overline{l_k}^2}{2}.$$

On en déduit une décomposition de Q sur $\mathbb{R}$ en somme de carrés de formes linéaires réelles :

$$Q = \sum_{k=1}^{p} m_k l_k^2 + \sum_{k=p+1}^{p+c} 2m_k \left(v_k^2 - w_k^2 \right).$$

Les l_k, v_i, w_j de cette décomposition sont linéairement indépendantes sur $\mathbb{R}$. En effet, si elles ne l'étaient pas, on aurait

$$a_1 l_1 + \dots + a_p l_p + b_{p+1} v_{p+1} + \dots + b_{p+c} v_{p+c} + c_{p+1} w_{p+1} + \dots + c_{p+c} w_{p+c} = 0,$$

et donc

$$\begin{aligned} a_1 l_1 + \dots + a_p l_p + \left(\frac{b_{p+1}}{2} + \frac{c_{p+1}}{2i} \right) l_{p+1} + \dots + \left(\frac{b_{p+c}}{2} + \frac{c_{p+c}}{2i} \right) l_{p+c} \\ + \left(\frac{b_{p+1}}{2} - \frac{c_{p+1}}{2i} \right) \overline{l_{p+1}} + \dots + \left(\frac{b_{p+c}}{2} - \frac{c_{p+c}}{2i} \right) \overline{l_{p+c}} \;\; = \;\; 0. \end{aligned}$$

En se rappelant que $\overline{l_{p+i}} = l_{p+c+i}$ (si on classe les l_k dans le bon ordre), on a donc obtenu une combinaison linéaire complexe non triviale des l_k, $1 \leqslant k \leqslant r$, ce qui est absurde car elles sont linéairement indépendantes sur $\mathbb{C}$ comme démontré ci-dessus.

Cette décomposition indique que la signature de Q est $(p + c, c)$, qui vaut aussi (s, t). Donc $p = s - c = s - t$ est le nombre de racines réelles distinctes de P. Enfin, $r = p + 2c = s - t + 2t = s + t$ est le nombre de racines complexes distinctes de P. □

Alternative pour montrer que $r = s + t$ juste après avoir montré l'indépendance des l_k, $1 \leqslant k \leqslant r$:

On a montré que les $(l_k)_{1\leqslant k\leqslant r}$ sont indépendantes sur $\mathbb{C}$, donc le rang de Q sur $\mathbb{C}$ (*i.e.* en tant que forme quadratique sur $\mathbb{C}$) vaut r : en effet, Q s'écrit comme somme de carrés de r formes linéaires indépendantes. Si on note $B(x,y)=\sum_{0\leqslant i,j\leqslant n} s_{i+j}x_iy_j$ la forme polaire associée à Q (en tant que forme quadratique sur $\mathbb{C}$), alors la matrice de Q est $(B(e_i,e_j))_{0\leqslant i,j\leqslant n-1}$ et est de rang r. Comme le rang d'une matrice ne dépend pas du corps de base (car il correspond à l'annulation d'un déterminant extrait), le rang de Q sur $\mathbb{R}$ vaut aussi r (en effet, $(B(e_i,e_j))_{0\leqslant i,j\leqslant n-1}$ est également la matrice de Q en tant que forme quadratique sur $\mathbb{R}$). Enfin, le rang de Q valant $s+t$, on obtient $r=s+t$, d'où la deuxième affirmation du théorème.

Définition. Soit $P\in A[X_1,\dots,X_n]$. On dit que P est symétrique si pour toute permutation $\sigma\in\mathcal{S}_n$, on a :

$$P(X_{\sigma(1)},\dots,X_{\sigma(n)})=P(X_1,\dots,X_n).$$

Définition. Les n polynômes

$$\Sigma_k=\sum_{1\leqslant i_1<\dots<i_k\leqslant n} X_{i_1}\dots X_{i_k}$$

pour $1\leqslant k\leqslant n$ sont symétriques. On les appelle polynômes symétriques élémentaires de $A[X_1,\dots,X_n]$.

Théorème (Relations coefficients-racines)**.** *Soit $P=a_nX^n+\dots+a_0\in K[X]$ avec $a_n\neq 0$. On suppose P scindé sur K et on note $x_1,\dots,x_n$ ses racines. Alors pour tout $k\in[\![1,n]\!]$, on a :*

$$\Sigma_k(x_1,\dots,x_n)=(-1)^k\frac{a_{n-k}}{a_n}.$$

Théorème (Formules de Newton)**.** *Pour $k\in\mathbb{N}$, on note*

$$S_k=\sum_{i=1}^{n}X_i^k\in A[X_1,\dots,X_n],$$

appelés sommes de Newton. On a les relations suivantes entre les sommes de Newton et les polynômes symétriques élémentaires :

$$S_k-\Sigma_1S_{k-1}+\dots+(-1)^{k-1}\Sigma_{k-1}S_1+(-1)^k\underbrace{k}_{\substack{ne\ pas\\ oublier}}\Sigma_k=0$$

pour tout $k\in[\![1,n-1]\!]$, et

$$S_k-\Sigma_1S_{k-1}+\dots+(-1)^{n-1}\Sigma_{n-1}S_{k-n-1}+(-1)^n\Sigma_nS_{k-n}=0$$

pour tout $k\geqslant n$.

Démonstration. Soient $x_1, \dots, x_n \in A$. On note $\sigma_i = \Sigma_i(x_1, \dots, x_n)$. On considère le polynôme $P = \prod_{i=1}^{n}(X - x_i) \in A[X]$. Il s'écrit aussi

$$P = X^n + \sum_{i=1}^{n}(-1)^i \sigma_i X^{n-i}.$$

Pour $i \in [\![1, n]\!]$, on a $P(x_i) = x_i^n - \sigma_1 x_i^{n-1} + \cdots + (-1)^n \sigma_n = 0$. On suppose alors $k \geqslant n$ et on multiplie par x_i^{k-n} :

$$x_i^k - \sigma_1 x_i^{k-1} + \cdots + (-1)^n \sigma_n x_i^{k-n} = 0.$$

On additionne enfin pour $1 \leqslant i \leqslant n$ et on obtient l'égalité :

$$S_k - \Sigma_1 S_{k-1} + \cdots + (-1)^{n-1}\Sigma_{n-1} S_{k-n-1} + (-1)^n \Sigma_n S_{k-n} = 0.$$

Soit maintenant $k \in [\![1, n-1]\!]$. Exprimons :

$$\sigma_1 S_{k-1} = \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{j=1}^{n} x_j^{k-1}\right) = \sum_{i,j} x_i x_j^{k-1} = S_k + \sum_{i \neq j} x_i x_j^{k-1}.$$

Ensuite,

$$\sigma_2 S_{k-2} = \left(\sum_{i<j} x_i x_j\right)\left(\sum_{l=1}^{n} x_l^{k-2}\right) = \sum_{i \neq j} x_i x_j^{k-1} + \sum_{\substack{i<j \\ l \neq i,j}} x_i x_j x_l^{k-2}.$$

On pose alors pour $p \in [\![0, k-1]\!]$,

$$A_p = \sum_{\substack{1 \leqslant i_1 < \cdots < i_p \leqslant n \\ i_{p+1} \neq i_1, \dots, i_p}} x_{i_1} \dots x_{i_p} x_{i_{p+1}}^{k-p}.$$

Pour $p \in [\![1, k-1]\!]$, on a

$$\sigma_p S_{k-p} = A_{p-1} + A_p,$$

en ayant remarqué que $A_0 = S_k$ et

$$A_{k-1} = \sum_{\substack{1 \leqslant i_1 < \cdots < i_{k-1} \leqslant n \\ i_k \neq i_1, \dots, i_{k-1}}} x_{i_1} \dots x_{i_k} = k\sigma_k$$

car on peut intercaler i_k dans les k emplacements entre 1 et i_1, entre i_1 et i_2, etc... En multipliant la première relation par -1, la seconde par $(-1)^2$,..., et la $(p-1)$-ième par $(-1)^{p-1}$, puis en sommant, on obtient

$$\sum_{k=1}^{p-1}(-1)^k \sigma_k S_{p-k} = -A_0 + (-1)^{p-1} A_{p-1} = -S_p + (-1)^{p-1} p\sigma_p.$$

On a donc le résultat voulu en faisant tout passer dans le membre de gauche. □

Remarque. Sur un corps de caractéristique nulle (il faut pouvoir diviser par les entiers), on peut inverser facilement le système triangulaire donné par les n premières égalités. On peut ainsi exprimer les sommes de Newton en fonction des polynômes symétriques élémentaires, et donc aussi en fonction des coefficients de P.

Définition. Soit Q une forme quadratique sur E, de forme polaire B (l'unique forme bilinéaire symétrique associée). On appelle rang de Q le rang de l'application linéaire

$$\begin{array}{rccc} \varphi : & E & \longrightarrow & E^* \\ & x & \longmapsto & (y \mapsto B(x,y)) \,. \end{array}$$

C'est aussi le rang de la matrice de Q par rapport à une base quelconque. En effet, si $\mathcal{B} = (e_1, \ldots, e_n)$ est une base de E, alors

$$\mathrm{Mat}_{\mathcal{B}}(\varphi) = \begin{pmatrix} B(e_1,e_1) & \ldots & B(e_n,e_1) \\ \vdots & \ddots & \vdots \\ B(e_1,e_n) & \ldots & B(e_n,e_n) \end{pmatrix} = \mathrm{Mat}_{\mathcal{B}}(Q)$$

Définition. Soit Q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E de dimension finie. Soit α la plus grande des dimensions des sous-espaces de E sur lesquels Q est définie positive. Soit β la plus grande des dimensions des sous-espaces de E sur lesquels Q est définie négative. Le couple (α, β) est appelé la signature de Q.

Théorème. *Soit Q une forme quadratique sur E de dimension finie. Il existe des formes linéaires $l_1, \ldots, l_k$ sur E, linéairement indépendantes, telles que :*

$$Q = l_1^2 + \cdots + l_m^2 - l_{m+1}^2 - \cdots - l_k^2.$$

Démonstration. La forme quadratique Q est un polynôme homogène de degré 2 en les x_i (coordonnées d'un point $x \in E$). On raisonne alors par récurrence sur le nombre de variables x_i apparaissant dans Q. S'il n'y en a qu'une, par exemple $Q(x) = ax_1^2$, c'est bon. S'il y en a au moins deux, il faut distinguer deux sous-cas : Q contient un terme carré, disons ax_1^2, et on pourra y incorporer les termes de la forme bx_1x_i. Donc on pourra écrire $Q = a(x_1 + A)^2 + Q_2$ avec Q_2 homogène du second degré en $x_2, \ldots, x_n$ et A polynôme homogone de degré 1 en $x_2, \ldots, x_n$. La forme linéaire $x_1 + A$ sera linéairement indépendante de celles de la décomposition de Q_2 (décomposition par l'hypothèse de récurrence) car la variable x_1 n'apparaît pas dans Q_2.

Le deuxième sous-cas est celui où Q ne contient que des termes rectangles de la forme ax_ix_j. On peut supposer qu'on a

$$Q = ax_1x_2 + \sum_{j=3}^{n} b_{1,j}x_1x_j + \sum_{j=3}^{n} b_{2,j}x_2x_j + Q_2$$

où Q_2 est homogène de degré 2 en $x_3, \dots, x_n$. On pose alors $r_1 = x_1 + \sum_{j=3}^{n} \frac{b_{1,j}}{a} x_j$ et $r_2 = x_2 + \sum_{j=3}^{n} \frac{b_{2,j}}{a} x_j$. On obtient

$$Q = a r_1 r_2 + Q_3$$

avec Q_3 homogène de degré 2 en $x_3, \dots, x_n$ et $(r_1, r_2, x_3, \dots, x_n)$ est libre. On pose ensuite $u_1 = \frac{r_1+r_2}{2}$ et $u_2 = \frac{r_1-r_2}{2}$ et on note que $u_1^2 - u_2^2 = r_1 r_2$. On obtient alors un polynôme en $u_1, u_2, x_3, \dots, x_n$ homogène de degré 2 contenant un terme carré $a u_1^2$. On peut alors appliquer le cas précédent. □

Conséquence : si $Q = l_1^2 + \dots + l_m^2 - l_{m+1}^2 - \dots - l_k^2$, alors sa forme polaire est $B(x,y) = l_1(x)l_1(y) + \dots + l_m(x)l_m(y) - l_{m+1}(x)l_{m+1}(y) - \dots - l_k(x)l_k(y)$. Comme les formes linéaires l_i sont indépendantes, on peut les compléter en une base de E^* puis définir sa base anté-duale $(f_1, \dots, f_n)$ dans E. Alors $B(f_i, f_j) = 0$ si $i \neq j$, 1 si $i = j \leqslant m$, -1 si $m+1 \leqslant i = j \leqslant k$, et 0 si $i = j \geqslant k+1$. La matrice de Q dans la base $(f_1, \dots, f_n)$ est donc diagonale et on obtient que le rang de Q vaut k.

Corollaire. *Il existe une base de E dans laquelle la matrice de Q est diagonale avec des 1, -1 et 0 sur la diagonale. On dit que la matrice de Q dans cette base est diagonale normalisée.*

Théorème. *Soit Q une forme quadratique sur E de dimension finie. Soit $\mathcal{B}$ une base dans laquelle la matrice A de Q soit diagonale normalisée. Soient a le nombre de coefficients $+1$ dans A et b le nombre de coefficients -1. Alors (a, b) est la signature de Q. En particulier, a et b ne dépendent pas de la base choisie.*

Démonstration. Soit (α, β) la signature de Q. On sait que Q est définie positive sur le sous-espace de E engendré par les vecteurs e de la base pour lesquels $Q(e) = +1$. Ce sous-espace est de dimension a. On a donc $\alpha \geqslant a$. D'autre part, Q est négative sur le sous-espace F engendré par les autres vecteurs de la base (les vecteurs e pour lesquels $Q(e) = -1$ ou 0). Soit G un sous-espace de E de dimension α sur lequel Q est définie positive. Si $\alpha > a$, alors $\dim(G) + \dim(F) > \dim(E)$. Par la formule de Grassmann, cela implique que $\dim(G \cap F) > 0$. Il existe donc $x \in F \cap G$, $x \neq 0$. Mais comme $x \in F$, $Q(x) \leqslant 0$, et comme $x \in G$ et $x \neq 0$, $Q(x) > 0$, ce qui est impossible. On a donc $\alpha = a$. En appliquant ce résultat à la forme quadratique $-Q$, on obtient aussi $\beta = b$. □

Références :

– Gantmacher - *Théorie des matrices, tome 2* - Page 199.

2.3 Décomposition de Dunford

Proposition. *Soient $f \in \mathcal{L}(E)$ et $P \in K[X]$ un polynôme annulateur de f. Soit $P = \beta M_1^{\alpha_1} \dots M_r^{\alpha_r}$ la décomposition de P en facteurs irréductibles de $K[X]$. On note $F_i = \mathrm{Ker}(M_i^{\alpha_i}(f))$. Alors $E = F_1 \oplus \dots \oplus F_r$ et pour tout i, la projection sur F_i parallèlement à $\bigoplus_{j \neq i} F_j$ est un polynôme en f.*

Démonstration. Le fait que $E = F_1 \oplus \dots \oplus F_r$ résulte directement du lemme des noyaux.

Pour tout i, on note $Q_i = \prod_{j \neq i} M_j^{\alpha_j}$. Aucun facteur n'est commun à tous les Q_i, ils sont donc premiers entre eux dans leur ensemble. Par l'identité de Bézout, il existe $U_1, \dots, U_r \in K[X]$ tels que $U_1 Q_1 + \dots + U_r Q_r = 1$, ce qui implique

$$U_1(f)Q_1(f) + \dots + U_r(f)Q_r(f) = \mathrm{Id}.$$

Pour tout i, on note $P_i = U_i Q_i$ et $p_i = P_i(f)$. On a donc

$$p_1 + \dots + p_r = \mathrm{Id}. \quad (*)$$

D'autre part, si $j \neq i$, alors P divise $Q_i Q_j$, donc $p_i \circ p_j = (U_i U_j)(f) \circ (Q_i Q_j)(f) = 0$ (puisque $P(f) = 0$, $(Q_i Q_j)(f) = 0$). En composant $(*)$ par p_i, on obtient alors $p_i = \sum_{i=1}^{r} p_i \circ p_j = p_i^2$. Les p_i sont donc des projecteurs.

Montrons que pour tout i, $\mathrm{Im}(p_i) = F_i$. Soit $y = p_i(x) \in \mathrm{Im}(p_i)$. Alors

$$M_i^{\alpha_i}(f)(y) = (M_i^{\alpha_i} U_i Q_i)(f)(x) = U_i(f) \circ P(f)(x) = 0.$$

Donc $\mathrm{Im}(p_i) \subset F_i$. Réciproquement, soit $x \in F_i$. D'après $(*)$, $x = p_1(x) + \dots + p_r(x)$. Or pour tout $j \neq i$, $p_j(x) = U_j(f) \circ Q_j(f)(x) = 0$ car $M_i^{\alpha_i}$ divise Q_j. Donc $x = p_i(x) \in \mathrm{Im}(p_i)$, et finalement, $\mathrm{Im}(p_i) = F_i$.

Il ne reste plus qu'à montrer que pour tout i, $\mathrm{Ker}(p_i) = \bigoplus_{j \neq i} F_j$. Pour tout $j \neq i$, on a $F_j \subset \mathrm{Ker}(p_i)$ car $M_j^{\alpha_j}$ divise $P_i = U_i Q_i$. Donc $\bigoplus_{j \neq i} F_j \subset \mathrm{Ker}(p_i)$. Réciproquement, soit $x \in \mathrm{Ker}(p_i)$. D'après $(*)$, $x = \sum_{j \neq i} p_j(x)$, donc $x \in \bigoplus_{j \neq i} F_j$. Finalement, $\mathrm{Ker}(p_i) = \bigoplus_{j \neq i} F_j$.

□

Théorème (Décomposition de Dunford)**.** *Soit $f \in \mathcal{L}(E)$ tel que son polynôme caractéristique P_f soit scindé sur K. Alors il existe un unique couple $(d, n) \in K[f]^2$ avec d diagonalisable et n nilpotent tels que $f = d + n$ et $d \circ n = n \circ d$.*

Démonstration. On écrit $P_f = (X - \lambda_1)^{m_1} \dots (X - \lambda_r)^{m_r}$ et on note $F_i = \mathrm{Ker}((f - \lambda_i \mathrm{Id})^{m_i})$. Alors d'après la proposition précédente, $E = F_1 \oplus \dots \oplus F_r$. On définit d et n sur les F_i en posant $d_{|F_i} = \lambda_i \mathrm{Id}_{F_i}$ et $n_{|F_i} = f - \lambda_i \mathrm{Id}_{F_i}$. Alors $n_{|F_i}$ est nilpotent

d'indice m_i et n est nilpotent d'indice $m = \max_i(m_i)$. Sur chaque F_i, d_i est une homothétie donc $d \circ n = n \circ d$. Si on note p_i le projecteur sur F_i parallèlement à $F_1 \oplus \cdots \oplus F_{i-1} \oplus F_{i+1} \oplus \cdots \oplus F_r$, alors $d = \lambda_1 p_1 + \cdots + \lambda_r p_r$ et d est un polynôme en f car les p_i le sont. Enfin, $n = f - d$ est aussi un polynôme en f. On a prouvé l'existence du couple (d, n) demandé.

Soit (d', n') un autre couple vérifiant les conditions. Comme $f = d' + n'$ et d' commute avec n', on a $f \circ d' = (d' + n') \circ d' = d' \circ (d' + n') = d' \circ f$. Donc F_i est stable par d' pour tout i. Comme d est une homothétie sur F_i, on en déduit que $d \circ d' = d' \circ d$ sur F_i pour tout i, donc sur E. De plus, d et d' sont diagonalisables, on peut donc les diagonaliser dans une même base, ce qui prouve que $d' - d$ est diagonalisable.

Comme $n = f - d$, $n' = f - d'$ et $d \circ d' = d' \circ d$, alors n et n' commutent. Si p et q sont tels que $n^p = n'^q = 0$, alors

$$(n - n')^{p+q} = \sum_{k=0}^{p+q} \binom{p+q}{k} n^k (-1)^{p+q-k} n'^{p+q-k} = 0.$$

Donc $n - n' = d' - d$ est nilpotent. Or $d' - d$ est diagonalisable, donc $d' - d = 0$, ce qui prouve que $d' = d$ et $n' = n$. On a donc unicité du couple (d, n) vérifiant les hypothèses du théorème. □

Calcul pratique de la décomposition de Dunford :

Il s'agit de calculer les projecteurs p_i car alors, $d = \sum_{i=1}^r \lambda_i p_i$ et $n = f - d$.

On remarque qu'on aurait pu remplacer P_f par n'importe quel polynôme P annulant f dans la démonstration de la décomposition de Dunford. Soit $P = \prod_{i=1}^r (X - \lambda_i)^{r_i}$ un polynôme annulateur de f. On décompose $\frac{1}{P}$ en éléments simples dans $K(X)$:

$$\frac{1}{P} = \sum_{i=1}^{r} \sum_{j=1}^{r_i} \frac{x_{i,j}}{(X - \lambda_i)^j}.$$

Pour tout i, on pose ensuite $U_i = \sum_{j=1}^{r_i} x_{i,j} (X - \lambda_i)^{r_i - j}$. Alors

$$\frac{1}{P} = \sum_{i=1}^{r} \frac{U_i}{(X - \lambda_i)^{r_i}}.$$

En multipliant par P, on obtient $\sum_{i=1}^r U_i Q_i = 1$ avec $Q_i = \prod_{j \neq i} (X - \lambda_i)^{r_i}$. Donc d'après la démonstration faite, en notant $P_i = U_i Q_i$, les projecteurs p_i sont donnés par $p_i = P_i(f)$.

Application au calcul d'exponentielle :

Si on connaît une base de diagonalisation de d, on a directement e^d. Comme n est nilpotente d'indice q, $e^n = \sum_{p=0}^{q-1} \frac{n^p}{p!}$. Enfin $e^f = e^d e^n$ car d et n commutent.

Sinon on utilise les projecteurs p_i. On a $d = \sum_{i=1}^{r} \lambda_i p_i$ et $n = \sum_{i=1}^{r} (f - \lambda_i \mathrm{Id}) p_i$. Les relations sur les p_i (*i.e.* $p_i^2 = p_i$ et $p_i \circ p_j = 0$) entraînent $d^p = \sum_{i=1}^{r} \lambda_i^p p_i$. Donc

$$e^d = \sum_{p=0}^{\infty} \frac{d^p}{p!} = \sum_{i=1}^{r} \left(\sum_{p=0}^{\infty} \frac{\lambda_i^p}{p!} \right) p_i = \sum_{i=1}^{r} e^{\lambda_i} p_i.$$

D'autre part,

$$e^n = \sum_{p=0}^{\infty} \frac{n^p}{p!} = \sum_{i=1}^{r} \left(\sum_{p=0}^{m_i - 1} \frac{(f - \lambda_i \mathrm{Id})^p}{p!} \right) p_i.$$

On en déduit, toujours grâce aux relations sur les p_i,

$$e^f = e^d e^n = \sum_{i=1}^{r} e^{\lambda_i} \left(\sum_{p=0}^{m_i - 1} \frac{(f - \lambda_i \mathrm{Id})^p}{p!} \right) p_i.$$

Application. *Calculer l'exponentielle de la matrice*

$$M = \begin{pmatrix} 1 & 4 & -2 \\ 0 & 6 & -3 \\ -1 & 4 & 0 \end{pmatrix}.$$

Démonstration. Le polynôme caractéristique est $P_M = (X-2)^2(X-3)$. On note $\lambda_1 = 2$, $\lambda_2 = 3$, et $Q_1 = (X-3)$, $Q_2 = (X-2)^2$. On décompose en éléments simples :

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

Donc $(X-2)^2 - (X-1)(X-3) = Q_2 - (X-1)Q_1 = 1$. On pose $U_1 = -(X-1)$ et $U_2 = 1$. Alors

$$p_1 = (U_1 Q_1)(M) = -(M - I_3)(M - 3I_3) = \begin{pmatrix} -2 & -4 & 6 \\ -3 & -3 & 6 \\ -3 & -4 & 7 \end{pmatrix},$$

et

$$p_2 = (M - 2I_3)^2 = \begin{pmatrix} 3 & 4 & -6 \\ 3 & 4 & -6 \\ 3 & 4 & -6 \end{pmatrix}.$$

Ainsi

$$e^M = e^{\lambda_1} \left(I_3 + \frac{1}{1!}(M - \lambda_1 I_3) \right) p_1 + e^{\lambda_2} p_2 = e^2 (M - I_3) p_1 + e^3 p_2.$$

□

Théorème (Lemme des noyaux)**.** *Soient $f \in \mathcal{L}(E)$ et $P = P_1 \dots P_r \in K[X]$ avec les P_i polynômes premiers entre eux deux à deux. Alors*

$$\mathrm{Ker}(P(f)) = \mathrm{Ker}(P_1(f)) \oplus \cdots \oplus \mathrm{Ker}(P_r(f)).$$

Démonstration. La preuve se fait bien par récurrence sur $k \geqslant 2$. □

Question : Peut-on calculer d et n de la décomposition de Dunford de f sans connaître les valeurs propres de f ?
Réponse : Oui, grâce à une démonstration effective (*i.e.* transformable en un algorithme) de la décomposition de Dunford.

Théorème. *Soient K un corps parfait et f un endomorphisme d'un K-espace vectoriel E de dimension finie. Alors il existe $d, n \in K[f]$ tels que $f = d + n$ avec d semi-simple et n nilpotent. De plus, cette décomposition est effective lorsque K est de caractéristique* 0*.*

Démonstration. Nous donnons seulement les idées de la preuve. Soit P un polynôme annulateur de f, $P = P_1^{\alpha_1} \dots P_r^{\alpha_r}$. On note $Q = P_1 \dots P_r$. Si K est de caractéristique 0, on peut calculer Q de manière effective grâce à l'égalité $P = \mathrm{PGCD}(P, P')Q$ (on a l'algorithme d'Euclide pour calculer un PGCD).

Soient $A = K[X]/(P)$ et $x = \overline{X}$. Comme $P(f) = 0$, on a un morphisme $\varphi : A \to K[f]$, $x \mapsto f$. Il s'agit donc de trouver une décomposition $x = u + v$ avec $Q(u) = 0$ et v nilpotent dans A. Alors $d = \varphi(u)$ et $n = \varphi(v)$ conviendront : d est semi-simple car annulé par Q qui est sans facteur carré.

Nous obtenons u grâce à la méthode de Newton : $x_0 = x$ et $x_{n+1} = x_n - \frac{Q(x_n)}{Q'(x_n)}$. On montre par récurrence que $Q'(x_n)$ est inversible dans A (la suite est alors bien définie) et $Q(x_n) \in (Q(x_n)^{2^n})$. Comme P divise $Q^{\mathrm{PPCM}(\alpha_i)}$, $Q(x)$ est nilpotent dans A et la suite stationne rapidement. Soient u sa limite, et $v = x - u$. Alors $Q(u) = 0$ et $v = \sum_{n=0}^{\infty}(x_{n+1} - x_n)$ (la somme est finie) est nilpotent comme somme de nilpotents. □

Références :

– Gourdon - *Algèbre* - Page 192.

2.4 Décomposition polaire dans $\mathrm{GL}_n(\mathbb{C})$

Théorème. *Soit $A \in \mathrm{GL}_n(\mathbb{C})$. Il existe un unique couple de matrices $(U, H) \in \mathcal{M}_n(\mathbb{C})^2$, avec U unitaire et H hermitienne définie positive, tel que $A = UH$. De plus, l'application*

$$\begin{array}{rccc} \Phi : & \mathcal{U}_n \times \mathcal{H}_n^{++} & \longrightarrow & \mathrm{GL}_n(\mathbb{C}) \\ & (U, H) & \longmapsto & UH \end{array}$$

est un homéomorphisme.

Démonstration. Si $A = UH$, alors ${}^t\overline{A} = {}^t\overline{H}\,{}^t\overline{U} = HU^{-1}$. Donc ${}^t\overline{A}A = H^2$. Nous allons par conséquent commencer par chercher une matrice hermitienne H vérifiant ${}^t\overline{A}A = H^2$.

On remarque que la matrice ${}^t\overline{A}A$ est hermitienne :

$${}^t\overline{\left({}^t\overline{A}A\right)} = {}^t\overline{A}\,{}^t\overline{\left({}^t\overline{A}\right)} = {}^t\overline{A}A.$$

De plus, ${}^t\overline{A}A$ est positive : en effet, pour tout vecteur colonne X, on a :

$${}^t\overline{X}\left({}^t\overline{A}A\right)X = {}^t\overline{AX}AX = \|AX\|^2 \geqslant 0,$$

où $\|\cdot\|$ désigne la norme hermitienne standard sur $\mathbb{C}^n$.

Or lorsqu'une matrice M est hermitienne positive, il existe une unique matrice R hermitienne positive telle que $M = R^2$ (ce qu'on démontre après). Donc il existe une unique matrice hermitienne H positive telle que ${}^t\overline{A}A = H^2$.

Si on suppose A inversible, alors H est inversible, donc H est définie positive (elle ne peut pas avoir de valeur propre nulle). On pose alors $U = AH^{-1}$ et on vérifie que U est unitaire :

$${}^t\overline{U}U = {}^t\overline{H^{-1}}\,{}^t\overline{A}AH^{-1} = \left({}^t\overline{H}\right)^{-1}H^2H^{-1} = H^{-1}H^2H^{-1} = I_n.$$

On vient donc de prouver l'existence et l'unicité du couple (U, H) lorsque A est inversible (l'unicité de U découle de celle de H). □

Proposition. *Soit $M \in \mathcal{M}_n(\mathbb{C})$ une matrice hermitienne positive. Alors il existe une unique matrice R hermitienne positive telle que $M = R^2$.*

Démonstration. La matrice M étant hermitienne, il existe une matrice unitaire C telle que

$${}^t\overline{C}MC = D$$

avec $D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ matrice diagonale réelle. Comme M est positive, tous les λ_i sont positifs. On pose $D' = \mathrm{Diag}\left(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}\right)$. Alors $R = CD'\,{}^t\overline{C}$ est hermitienne positive et vérifie $R^2 = M$.

Pour l'unicité, on considère le polynôme d'interpolation de Lagrange φ défini par $\varphi(\lambda_i) = \sqrt{\lambda_i}$ pour tout $i \leqslant n$. Alors $\varphi(M) = C\varphi(D)^t\overline{C} = CD'\,{}^t\overline{C} = R$. Donc R est polynomiale en M. Par conséquent, si S est matrice hermitienne positive telle que $S^2 = M$, alors S commute avec M (car $SM = S^3 = MS$), et puisque R est polynomiale en M, S commute avec R. Les matrices R et S commutent et sont diagonalisables (car hermitiennes), elles sont donc simultanément diagonalisables : $R = QD_1Q^{-1}$ et $S = QD_2Q^{-1}$ avec D_1 et D_2 diagonales à coefficients diagonaux positifs. Sachant que $M = R^2 = S^2$, on obtient $D_1^2 = D_2^2$, et donc $D_1 = D_2$. Finalement $R = S$ et l'unicité est prouvée. □

Il nous reste maintenant à démontrer que Φ est un homéomorphisme. La continuité de Φ est évidente et on vient de montrer que Φ est bijective. Il reste donc à montrer que Φ^{-1} est continue, *i.e.* que si (M_p) est une suite de $\mathrm{GL}_n(\mathbb{C})$, avec $M_p = U_pH_p$, qui converge vers $M = UH$, alors (U_p) converge vers U et (H_p) converge vers H.

Montrons que le groupe des matrices unitaires est compact. En effet, c'est un fermé borné de $\mathcal{M}_n(\mathbb{C})$ et $\mathcal{M}_n(\mathbb{C})$ est de dimension finie. Il est fermé car c'est l'image réciproque de $\{I_n\}$ par l'application continue $U \mapsto {}^t\overline{U}U$. Si $P = (p_{ij})$ est unitaire, alors

$$ {}^t\overline{P}P = \left(\sum_{k=1}^{n} \overline{p_{ki}} p_{kj} \right) = I_n. $$

En particulier, les coefficients diagonaux de ${}^t\overline{P}P$ sont égaux à 1, *i.e.* pour tout k,

$$ \sum_{k=1}^{n} |p_{ki}|^2 = 1. $$

Donc pour tout i, j, $|p_{ij}| \leqslant 1$, et donc $\mathcal{U}_n$ est borné par 1 au sens de la norme définie par $\|P\| = \sup |p_{ij}|$ (toutes les normes sont équivalentes sur $\mathcal{M}_n(\mathbb{C})$ de dimension finie).

Par compacité de $\mathcal{U}_n$, il existe une sous-suite $(U_{\varphi(p)})$ de (U_p) qui converge vers $U' \in \mathcal{U}_n$. Par conséquent :

$$ H_{\varphi(p)} = {}^t\overline{U_{\varphi(p)}} M_{\varphi(p)} \longrightarrow H' = {}^t\overline{U'}M. $$

L'expression de H' montre que H' est inversible, et de plus, H' est hermitienne positive (limite des matrices hermitiennes positives $H_{\varphi(p)}$), donc $H' \in \mathcal{H}_n^{++}$. Par unicité de la décomposition polaire, on a donc $U' = U$ et $H' = H$. Cela prouve que la suite (U_p) du compact $\mathcal{U}_n$ n'a qu'une seule valeur d'adhérence, donc qu'elle est convergente et $U_k \to U$. Il vient ensuite $H_k = {}^t\overline{U_k}M_k \to {}^t\overline{U}M = H$.

Enoncé du théorème adapté pour la leçon *Exemples d'actions de groupes sur les espaces de matrices* :

Théorème. *On considère l'action de groupe suivante sur $\mathcal{M}_n(\mathbb{C})$:*

$$\begin{array}{rccc} \varphi : & \mathcal{U}_n \times \mathcal{M}_n(\mathbb{C}) & \longrightarrow & \mathcal{M}_n(\mathbb{C}) \\ & (U, M) & \longmapsto & UM. \end{array}$$

Alors l'orbite de toute matrice M contient une matrice hermitienne positive. De plus, si M est inversible, l'orbite de $M \in \mathrm{GL}_n(\mathbb{C})$ contient une unique matrice hermitienne positive, qui est par conséquent définie positive.

Rappelons que l'orbite d'un élément $x \in E$ sous l'action du groupe G est

$$O_x = \{g.x,\, g \in G\}.$$

Donc ici, pour $M \in \mathcal{M}_n(\mathbb{C})$, on a $O_M = \{UM,\, U \in \mathcal{U}_n\}$. On vérifie donc bien que si $H \in \mathcal{H}_n \cap O_M$, alors $M = {}^t\overline{U}H$ est le produit d'une matrice unitaire par une matrice hermitienne positive.

Si on a du temps, on peut présenter une application facile de la décomposition polaire :

Proposition. *Le groupe $\mathcal{U}_n$ est un sous-groupe compact maximal de $\mathrm{GL}_n(\mathbb{C})$.*

Démonstration. On a déjà montré que $\mathcal{U}_n$ est compact. Soit G un sous-groupe compact de $\mathrm{GL}_n(\mathbb{C})$ avec $\mathcal{U}_n \subset G$. On va montrer que nécessairement $G = \mathcal{U}_n$ (on ne peut pas avoir $G = \mathrm{GL}_n(\mathbb{C})$ qui n'est pas compact).

Soit $M \in G$. On écrit la décomposition polaire de M : $M = UH$. Or $U \in \mathcal{U}_n \subset G$, donc $H = U^{-1}M \in G$. Le sous-groupe G étant compact, la suite $(H^k)_{k\in\mathbb{N}}$ admet une valeur d'adhérence dans G. Mais $H \in \mathcal{H}_n^{++}$ est diagonalisable en une matrice diagonale dont les éléments diagonaux sont des réels strictement positifs. La convergence d'une sous-suite de (H^k) implique que les valeurs propres de H sont $\leqslant 1$. Mais s'il existe une valeur propre de H strictement plus petite que 1, alors la limite de la sous-suite est non inversible, donc n'est pas dans G. Finalement, toutes les valeurs propres de H valent 1, donc $H = I_n$. Ainsi $M = U \in \mathcal{U}_n$, ce qui montre que $G \subset \mathcal{U}_n$. □

Remarque. L'existence de la décomposition polaire est encore vraie pour $M \in \mathcal{M}_n(\mathbb{C})$ (avec bien sûr H positive et non définie positive). Par contre on perd l'unicité. Pour montrer l'existence, on se sert de la densité de $\mathrm{GL}_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$ et de la compacité de $\mathcal{U}_n$.

On a aussi une décomposition polaire sur $\mathcal{M}_n(\mathbb{R})$ en tout point similaire à celle sur $\mathcal{M}_n(\mathbb{C})$, en remplaçant $\mathcal{U}_n$ par $\mathcal{O}_n(\mathbb{R})$ (les matrices orthogonales) et $\mathcal{H}_n$ par $\mathcal{S}_n(\mathbb{R})$ (la matrices symétriques). La démonstration est en tout point identique (tout se transpose à l'identique : $\mathcal{O}_n(\mathbb{R})$ est aussi compact, etc...).

Remarque. On peut exprimer facilement l'inverse de Φ grâce à la construction donnée dans la démonstration. On note :

$$\begin{array}{rccc} \varphi : & \mathcal{H}_n^{++} & \longrightarrow & \mathcal{H}_n^{++} \\ & H & \longmapsto & \sqrt{H} \end{array} .$$

Alors si $M \in \mathrm{GL}_n(\mathbb{C})$, on avait $M = UH$ avec $H = \varphi\left({}^t\overline{M}M\right)$. D'où :

$$\Phi^{-1}(M) = \left(M\varphi\left({}^t\overline{M}M\right)^{-1}, \varphi\left({}^t\overline{M}M\right)\right).$$

Définition. On appelle matrice *unitaire* toute matrice $U \in \mathcal{M}_n(\mathbb{C})$ vérifiant ${}^t\overline{U}U = I_n$.

Définition. On appelle matrice *hermitienne* toute matrice $H \in \mathcal{M}_n(\mathbb{C})$ vérifiant ${}^t\overline{H} = H$.

Proposition. *Si A est symétrique réelle, toutes ses valeurs propres sont réelles.*

Démonstration. Soit $\lambda \in \mathbb{C}$ une racine complexe du polynôme caractéristique de A. Il existe un vecteur colonne $X \neq 0$ tel que $AX = \lambda X$. On a d'une part :

$${}^tXA\overline{X} = {}^tX\overline{AX} = {}^tX\overline{\lambda X} = \overline{\lambda}\,{}^tX\overline{X},$$

car A est réelle, donc $\overline{A} = A$. D'autre part :

$${}^tXA\overline{X} = {}^tX\,{}^tA\overline{X} = {}^t(AX)\overline{X} = \lambda{}^tX\overline{X},$$

car A est symétrique. Mais si on écrit ${}^tX = (z_1, \ldots, z_n)$, alors ${}^tX\overline{X} = \sum_{i=1}^n |z_i|^2 > 0$ car $X \neq 0$. Donc $\overline{\lambda} = \lambda$, *i.e.* toutes les valeurs propres de A sont réelles. □

C'est pareil pour les matrices hermitiennes :

Proposition. *Si $A \in \mathcal{M}_n(\mathbb{C})$ est hermitienne, toutes ses valeurs propres sont réelles.*

Démonstration. On fait pareil mais cette fois en calculant de deux façons différentes ${}^tX\overline{A}\,\overline{X}$. □

Théorème. *Tout endormorphisme symétrique d'un espace euclidien est diagonalisable dans une base orthonormée, en une matrice diagonale réelle.*

Démonstration. On raisonne par récurrence sur la dimension n de l'espace. D'après la proposition précédente, il existe une valeur propre réelle λ, dont on note e_1 le vecteur propre associé. On prend $F = (\mathbb{R}e_1)^\perp$ et on montre que F est stable par l'endomorphisme u : si $x \in F$,

$$(u(x)|e_1) = (x|u(e_1)) = (x|\lambda e_1) = \lambda(x|e_1) = 0,$$

car u est symétrique pour la première égalité. Donc $u(x) \in F$. On applique ensuite l'hypothèse de récurrence à $u_{|F}$ qui est toujours symétrique. □

C'est pareil pour les endomorphismes hermitiens :

Théorème. *Tout endormorphisme hermitien d'un espace hermitien est diagonalisable dans une base orthonormée, en une matrice diagonale réelle.*

Démonstration. Pareil que le cas des endomorphismes symétriques. □

En termes de matrices, cela se traduit par :

Théorème. *Toute matrice hermitienne est diagonalisable en une matrice diagonale réelle au moyen d'une matrice unitaire.*

Démonstration. Soit H une matrice hermitienne. D'après le théorème d'avant, il existe une base orthonormée telle que l'endomorphisme associé à H soit diagonalisable. On note P la matrice de passage de la base canonique $(e_1, \dots, e_n)$ à cette base orthonormée $(\varepsilon_1, \dots, \varepsilon_n)$. Alors :

$$H = PDP^{-1}$$

avec D diagonale réelle. On note $P = (p_{ij})$, *i.e.* $\varepsilon_i = \sum_{k=1}^n p_{ki} e_k$ pour tout i. Alors la famille $(\varepsilon_1, \dots, \varepsilon_n)$ est orthonormée équivaut à

$$(\varepsilon_i|\varepsilon_j) = \left(\sum_{k=1}^n p_{ki} e_k \middle| \sum_{k=1}^n p_{kj} e_k \right) = \sum_{k=1}^n \overline{p_{ki}} p_{kj} = \delta_{ij},$$

qui équivaut à ${}^t\overline{P}P = I_n$, *i.e.* P est unitaire. □

Références :

- Gourdon - *Algèbre* - Page 249.
- Mneimné et Testard - *Introduction à la théorie des groupes de Lie classiques* - Page 19 (pour l'homéomorphisme).
- Serre - *Les matrices* (pour l'homéomorphisme et l'application).

2.5 Décomposition QR

On note $\mathcal{T}_n^{++}$ le groupe des matrices triangulaires supérieures de $\mathcal{M}_n(\mathbb{C})$ dont tous les coefficients diagonaux sont strictement positifs. On note aussi $\mathcal{U}_n$ le groupe des matrices de $\mathcal{M}_n(\mathbb{C})$ unitaires, *i.e.* les M telles que ${}^t\overline{M}M = I_n$.

Théorème. *Pour tout $A \in \mathrm{GL}_n(\mathbb{C})$, il existe un unique couple $(Q,R) \in \mathcal{U}_n \times \mathcal{T}_n^{++}$ tel que $A = QR$. De plus, cette factorisation permet de résoudre le système $Ax = b$.*

Démonstration. Nous ne montrerons que l'existence. Voir les suppléments plus bas pour l'unicité.

Soit $A \in \mathrm{GL}_n(\mathbb{C})$ dont on note $a_1, \dots, a_n$ ses vecteurs colonnes qui forment une base de $\mathbb{C}^n$. Appliquons le procédé d'orthonormalisation de Gram-Schmidt. On pose $q_1 = \frac{a_1}{\|a_1\|}$. On construit les vecteurs q_j pour $j \geqslant 2$ par la formule de récurrence de Gram-Schmidt :

$$q_j = \frac{w_j}{\|w_j\|} \text{ avec } w_j = a_j - \sum_{k=1}^{j-1} (q_k, a_j) q_k,$$

où $(\cdot,\cdot)$ désigne le produit scalaire hermitien canonique de $\mathbb{C}^n$. La base $(q_1, \dots, q_n)$ ainsi construite est orthonormée et donc si on note $Q = (q_1, \dots, q_n)$ la matrice formée par les q_i, alors $Q \in \mathcal{U}_n$.

Par ailleurs, pour tout j, $a_j = \|w_j\| q_j + \sum_{k=1}^{j-1} (q_k, a_j) q_k$. En notant $r_{j,j} = \|w_j\|$ pour tout j, et $r_{k,j} = (q_k, a_j)$ pour tout $1 \leqslant k \leqslant j-1$, on obtient $a_j = \sum_{k=1}^{j} r_{k,j} q_k$. On note $(q_{i,j})_{1 \leqslant i \leqslant n}$ les composantes du vecteur q_j et $(a_{i,j})_{1 \leqslant i \leqslant n}$ celles du vecteur a_j. Alors

$$\underbrace{\begin{pmatrix} q_{1,1} & \dots & q_{1,n} \\ \vdots & & \vdots \\ q_{n,1} & \dots & q_{n,n} \end{pmatrix}}_{=Q} \underbrace{\begin{pmatrix} r_{1,1} & \dots & r_{1,n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & r_{n,n} \end{pmatrix}}_{=R} = \left(\sum_{k=1}^{j} q_{i,k} r_{k,j} \right)_{i,j} = (a_{ij})_{i,j} = A,$$

et on a bien $R \in \mathcal{T}_n^{++}$.

On souhaite maintenant résoudre le système $Ax = b$. On décompose A en $A = QR$. Le système devient alors $QRx = b$, et en multipliant à gauche par ${}^t\overline{Q}$, le système se transforme en $Rx = {}^t\overline{Q}b$. Ce système se résout facilement par un algorithme de remontée puisque R est triangulaire supérieure. □

Algorithme 1: Procédé d'orthonormalisation de Gram-Schmidt

Données : $A = (a_1, \ldots, a_n)$
Résultat : $Q = (q_1, \ldots, q_n)$ et $R = (r_{ij})$
pour $j = 1$ *à* n **faire**
  $w_j = a_j$
  pour $k = 1$ *à* $j - 1$ **faire**
    $r_{kj} = (q_k, a_j)$
    $w_j = w_j - r_{kj} q_k$
  fin
  $r_{jj} = \|w_j\|$
  $q_j = w_j / r_{jj}$
fin

Comptons le nombre d'opérations que cet algorithme nécessite. Pour un produit scalaire, on doit effectuer $n - 1$ additions et n multiplications. Pour ajouter deux vecteurs, on doit effectuer n additions. Pour calculer une norme, on doit effectuer un produit scalaire et une extraction de racine carrée. Nous avons donc le nombre suivant d'additions dans l'algorithme :

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{j-1} (n - 1 + n) + (n - 1) \right) = \sum_{j=1}^{n} \big((2n-1)(j-1) + (n-1) \big) = \frac{n(n-1)(2n+1)}{2}.$$

De la même façon, on trouve n^3 multiplications, n^2 divisions et n extractions de racines carrées. Au final, on doit effectuer de l'ordre de $2n^3$ opérations.

Cependant, sur ordinateur, la propagation d'erreurs d'arrondis fait que les vecteurs q_i calculés ne sont pas forcément linéairement indépendants, donc en particulier pas orthogonaux. Cela empêche donc la matrice Q d'être exactement orthogonale. Ces instabilités numériques sont dues au fait que la procédure d'orthonormalisation produit des valeurs très petites, ce qui pose problème en arithmétique à virgule flottante (*i.e.* l'arithmétique des nombres réels sur ordinateur où tout réel est représenté par un nombre fini de bits). Il convient alors de recourir à une version plus stable de l'algorithme, appelée *procédé de Gram-Schmidt modifié.*

La modification consiste tout simplement en un réordonnancement des calculs de façon à ce que dès qu'un vecteur de la base orthonormée est obtenu, tous les vecteurs restants à orthonormaliser lui soient rendus orthogonaux. Les deux versions de Gram-Schmidt sont mathématiquement équivalentes, mais la version modifiée est préférable à la première lorsque les calculs sont effectués sur ordinateur.

Algorithme 2: Procédé d'orthonormalisation de Gram-Schmidt modifié

Données : $A = (a_1, \ldots, a_n)$
Résultat : $Q = (q_1, \ldots, q_n)$ et $R = (r_{ij})$
pour $i = 1$ *à* n **faire** $w_i = a_i$
pour $k = 1$ *à* n **faire**
  $r_{kk} = \|w_k\|$
  $q_k = w_k / r_{kk}$
  pour $j = k + 1$ *à* n **faire**
    $r_{kj} = (q_k, w_j)$
    $w_j = w_j - r_{kj} q_k$
  fin
fin

Après calculs, on trouve que les nombres d'additions, de multiplications, de divisions et d'extractions de racines carrées sont les mêmes que pour l'algorithme non modifié. On a donc encore de l'ordre de $2n^3$ opérations au total.

Enfin, pour résoudre un système $Ax = b$ où A est inversible et triangulaire inférieure, on utilise naturellement une méthode dite de descente :

$$\begin{cases} x_1 = \frac{b_1}{a_{11}} \\ x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j \right), \ i = 2, \ldots, n. \end{cases}$$

Cet algorithme effectue $\frac{n(n-1)}{2}$ additions, $\frac{n(n-1)}{2}$ multiplications et n divisions, soit un nombre d'opérations global de l'ordre de n^2.

Comparaison avec d'autres méthodes :

Méthodes	$+$	$*$	$/$	$\sqrt{\cdot}$
Résolution par la règle de Cramer	$(n+1)!$	$(n+2)!$	n	0
Décomposition LU	$\frac{n^3}{3}$	$\frac{n^3}{3}$	$\frac{n^2}{2}$	0
Décomposition QR	n^3	n^3	n^2	n

On a indiqué les ordres de grandeurs, en ne comptant que les opérations de la décomposition (pas la résolution finale du système avec l'algorithme de remontée par exemple, sauf pour la règle de Cramer).

Sur la décomposition QR par la méthode de Householder :

En pratique, on préfère calculer la factorisation QR d'une matrice par la méthode de Householder dont le principe est de multiplier A par une suite de matrices de transformations très simples dites de Householder pour l'amener progressivement sous forme triangulaire supérieure.

Définition. Soit v un vecteur non nul de $\mathbb{R}^n$. On appelle *matrice de Householder* associée au *vecteur de Householder* v la matrice définie par :

$$H(v) = I_n - 2\frac{v^t v}{{}^t vv}.$$

On pose de plus $H(0) = I_n$, ce qui permet de considérer la matrice identité comme une matrice de Householder.

Les matrices de Householder possèdent des propriétés intéressantes :

Propriété. *Soit v un vecteur non nul de $\mathbb{R}^n$ et $H(v)$ la matrice de Householder qui lui est associée. Alors $H(v)$ est symétrique et orthogonale. De plus, si $x \in \mathbb{R}^n$ et e est un vecteur unitaire tel que $x \neq \pm\|x\|_2 e$, on a*

$$H(x \pm \|x\|_2 e)x = \mp\|x\|_2 e.$$

La matrice de Householder $H(v)$ est la matrice de la symétrie orthogonale par rapport à l'hyperplan orthogonal à v. Les matrices de Householder peuvent par conséquent être utilisées pour annuler certaines composantes d'un vecteur x de $\mathbb{R}^n$ donné.

La méthode de Householder est sensiblement plus coûteuse que la méthode de Gram-Schmidt, mais son succès est sa grande stabilité numérique : elle ne modifie pas le conditionnement du problème (la norme $\|\cdot\|_2$ est invariante par transformation unitaire, et le conditionnement est calculé par $\text{cond}(A) = \|A\|\|A^{-1}\|$). De plus, les colonnes de Q sont numériquement orthonormale, cela ne dépend pas du degré d'indépendance des colonnes de la matrice A comme cela était le cas pour le procédé de Gram-Schmidt.

Cependant, si la connaissance de Q n'est pas requise et qu'on souhaite juste la solution de $Ax = b$, un raccourci existe : la structure des matrices de Householder fait qu'il n'est pas nécessaire d'assembler la matrice de Householder pour en effectuer le produit avec une autre matrice. Cela réduit le nombre d'opérations à un ordre de grandeur de $\frac{4}{3}n^3$, plus avantageux que Gram-Schmidt, mais encore le double de LU (par l'élimination de Gauss).

Méthode de Cholesky pour les matrices symétriques définies positives :
Elle nécessite un nombre d'opérations total de l'ordre de $\frac{n^3}{3}$.

Montrons l'unicité de la décomposition $A = QR$ annoncée dans le théorème.

Démonstration. Commençons par montrer que $\mathcal{T}_n^{++} \cap \mathcal{U}_n = \{I_n\}$. Il est clair que $I_n \in \mathcal{T}_n^{++} \cap \mathcal{U}_n$. Réciproquement, soit $M \in \mathcal{T}_n^{++} \cap \mathcal{U}_n$. Alors $M = (m_{ij})$ est triangulaire supérieure (*i.e.* $m_{ij} = 0$ si $i > j$) avec $m_{ii} > 0$ et $M^t\overline{M} = I_n$. On a donc :

$$M^t\overline{M} = \left(\sum_{k=\max(i,j)}^{n} m_{ik}\overline{m_{jk}} \right)_{i,j} = I_n.$$

En prenant $j = n$, on trouve $\sum_{k=n}^{n} m_{ik}\overline{m_{nk}} = m_{in}\overline{m_{nn}} = 0$ pour tout $i < n$, ce qui implique $m_{in} = 0$ pour tout $i < n$. En prenant ensuite $j = n-1$, on trouve

$$\sum_{k=n-1}^{n} m_{ik}\overline{m_{n-1,k}} = m_{i,n-1}\overline{m_{n-1,n-1}} + \underbrace{m_{i,n}}_{=0}\underbrace{\overline{m_{n-1,n}}}_{=0} = m_{i,n-1}\underbrace{\overline{m_{n-1,n-1}}}_{>0} = 0$$

pour tout $i < n-1$, ce qui implique $m_{i,n-1} = 0$ pour tout $i < n-1$. On recommence jusqu'à $j = 1$ et on obtient finalement que tous les coefficients non diagonaux sont nuls. Comme les coefficients diagonaux sont des réels > 0 tels que $|m_{ii}|^2 = 1$, on en déduit $M = I_n$.

Soient maintenant (Q_1, R_1) et $(Q_2, R_2) \in \mathcal{U}_n \times \mathcal{T}_n^{++}$ tels que $A = Q_1R_1 = Q_2R_2$. Alors $Q_2^{-1}Q_1 = R_2R_1^{-1}$. Or $Q_2^{-1}Q_1 \in \mathcal{U}_n$ et $R_2R_1^{-1}$ est triangulaire supérieure avec ses coefficients diagonaux strictements positifs. Donc

$$Q_2^{-1}Q_1 \in \mathcal{U}_n \cap \mathcal{T}_n^{++} \text{ et } R_2R_1^{-1} \in \mathcal{U}_n \cap \mathcal{T}_n^{++}.$$

Comme $\mathcal{U}_n \cap \mathcal{T}_n^{++} = \{I_n\}$, il vient $Q_2 = Q_1$ et $R_2 = R_1$. □

Références :

- Filbet - *Analyse numérique, algorithme et étude mathématique* - Exercice 1.5 page 49, solution page 55.
- Legendre - *Méthodes numériques, Introduction à l'analyse numérique et au calcul scientifique* - Cours PDF daté de 2009/2010 (pour Gram-Schmidt modifié, les algorithmes et les complexités).
- Ciarlet - *Introduction à l'analyse numérique matricielle et à l'optimisation* - Les nombres d'opérations y sont aussi.

2.6 Dénombrement des polynômes irréductibles sur un corps fini

Soient p un nombre premier, $r \in \mathbb{N}^*$ et $q = p^r$. On s'intéresse à l'ensemble $I_q(n)$ des polynômes irréductibles unitaires de degré n sur $\mathbb{F}_q$ dont on va calculer le cardinal, qu'on note $N_q(n)$.

Proposition. *Pour $n \geqslant 1$, on a*

$$X^{q^n} - X = \prod_{d|n} \prod_{P \in I_q(d)} P$$

et

$$N_q(n) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d,$$

μ désignant la fonction de Möbius.

Démonstration. Soient d un diviseur de n et $P \in I_q(d)$. On veut montrer que P divise $X^{q^n} - X$. Soit alors α une racine de P dans un corps de rupture de P sur $\mathbb{F}_q$. Ainsi $\mathbb{F}_q(\alpha)$ est un corps de rupture de P et $[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = \deg(P) = d$ car P est irréductible, c'est donc le polynôme minimal de α sur $\mathbb{F}_q$. Par conséquent, $\mathbb{F}_q(\alpha) \cong \mathbb{F}_{q^d}$. Or tout élément de $\mathbb{F}_{q^d}$ est racine du polynôme $X^{q^d} - X$, donc $\alpha^{q^d} = \alpha$. Comme d divise n, il existe un entier k tel que $n = kd$, et alors :

$$\alpha^{q^n} = \underbrace{\left(\left(\alpha^{q^d}\right)^{q^d} \dots\right)^{q^d}}_{k \text{ fois}} = \alpha.$$

Ainsi α est racine de $X^{q^n} - X$ et P divise ce polynôme car ses racines sont simples (c'est un polynôme irréductible sur un corps fini, et tout corps fini est parfait, donc P est séparable).

Si on prend un autre polynôme $Q \in I_q(d')$ avec d' un diviseur de n, alors Q divise aussi $X^{q^n} - X$, et les deux polynômes P et Q étant irréductibles, leur produit PQ divise $X^{q^n} - X$.

Finalement :

$$\left(\prod_{d|n} \prod_{P \in I_q(d)} P\right) \text{ divise } \left(X^{q^n} - X\right).$$

Réciproquement, soit $P \in \mathbb{F}_q[X]$ un facteur irréductible de $X^{q^n} - X$ de degré $d \geqslant 1$. Soit $\alpha \in \mathbb{F}_{q^n}$ une racine de P ($X^{q^n} - X$ est scindé sur $\mathbb{F}_{q^n}$, donc P aussi). Alors

$$[\mathbb{F}_{q^n} : \mathbb{F}_q(\alpha)][\mathbb{F}_q(\alpha) : \mathbb{F}_q] = n.$$

Or $[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = \deg(P) = d$ car P est irréductible sur $\mathbb{F}_q$, c'est donc le polynôme minimal de α sur $\mathbb{F}_q$. Ceci prouve que d divise n.

On vient donc de montrer que tout facteur irréductible de $X^{q^n} - X$ est un polynôme de $I_q(d)$ avec d un diviseur de n. Il existe donc des polynômes deux à deux distincts Q_i dans $I_q(d_i)$ avec d_i un diviseur de n et des entiers $\alpha_i \geqslant 1$ tels que

$$X^{q^n} - X = \prod_i Q_i^{\alpha_i}.$$

Montrons maintenant que $X^{q^n} - X$ est sans facteur carré dans $\mathbb{F}_q[X]$. Supposons par l'absurde que ce polynôme admet un facteur carré : on écrit $X^{q^n} - X = Q^2 P$ avec $P, Q \in \mathbb{F}_q[X]$. En dérivant, on obtient $q^n X^{q^n-1} - 1 = 2QQ'P + Q^2P'$. Comme $q^n = 0$ dans $\mathbb{F}_q$, on en déduit que Q divise le polynôme constant -1. Donc Q est constant, ce qui est absurde. Ceci montre que pour tout i, $\alpha_i = 1$, et enfin :

$$\left(X^{q^n} - X = \prod_i Q_i\right) \text{ divise } \left(\prod_{d|n} \prod_{P \in I_q(d)} P\right).$$

On a donc montré que

$$X^{q^n} - X = \prod_{d|n} \prod_{P \in I_q(d)} P.$$

Pour calculer $N_q(n)$, on passe au degré dans cette dernière égalité polynomiale :

$$q^n = \sum_{d|n} d N_q(d).$$

On introduit alors les fonctions arithmétiques f et g définies par $f(n) = nN_q(n)$ et $g(n) = q^n$ pour $n \in \mathbb{N}^*$. La formule précédente s'écrit :

$$g(n) = \sum_{d|n} f(d) = f * \mathbb{1}(n).$$

La formule d'inversion de Möbius nous donne alors :

$$f = g * \mu,$$

soit encore

$$N_q(n) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d.$$

□

Remarque. Si on pose $r_n = \sum_{\substack{d|n \\ d\neq n}} \mu\left(\frac{n}{d}\right) q^d$, alors :

$$|r_n| \leqslant \sum_{d=1}^{\left\lfloor \frac{n}{2} \right\rfloor} q^d = \frac{q^{\left\lfloor \frac{n}{2} \right\rfloor + 1} - q}{q-1}.$$

On en déduit :

$$N_q(n) = \frac{q^n + r_n}{n} \underset{n\to\infty}{\sim} \frac{q^n}{n} = \frac{q^n}{\log_q(q^n)}.$$

On peut faire l'analogie avec le nombre $\pi(n)$ de nombres premiers inférieurs à n :

$$\pi(n) \underset{n\to\infty}{\sim} \frac{n}{\ln(n)}.$$

Remarque. On obtient une autre démonstration de l'existence d'un corps à p^n éléments avec $n \in \mathbb{N}^*$: avec $r = 1$, le dénombrement des polynômes irréductibles sur $\mathbb{F}_q$ pour $q = p^r = p$ montre que

$$nN_p(n) = p^n + r_n \geqslant p^n - |r_n|.$$

avec r_n défini comme dans la remarque précédente (la fonction μ peut être négative). Or $|r_n| < p^n$, donc $N_p(n) \neq 0$. Il existe donc P irréductible de degré n dans $\mathbb{F}_p[X]$. Alors $K = \mathbb{F}_p[X]/(P)$ est un corps (car P est irréductible), de cardinal p^n car K est un $\mathbb{F}_p$-espace vectoriel de dimension n.

Remarque. Le corps $\mathbb{F}_p$ est un sous-corps de $\mathbb{F}_q$ (c'est son sous-corps premier, *i.e.* plus petit sous-corps inclu dans $\mathbb{F}_q$). En effet, la caractéristique d'un corps fini est nécessairement un nombre premier (plus petit entier n tel que $n1 = 0$), donc en notant $\varphi : \mathbb{Z} \to \mathbb{F}_q, n \mapsto n1$, alors $\mathrm{Ker}(\varphi) = p\mathbb{Z}$, si bien que $\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p \cong \mathrm{Im}(\varphi) \subset \mathbb{F}_q$, d'où $\mathbb{F}_p \subset \mathbb{F}_q$. Ensuite $\mathbb{F}_p$ est un corps premier : si k est un sous-corps de $\mathbb{F}_p$, en particulier k est un sous-groupe additif de $\mathbb{F}_p$, donc le cardinal de k divise celui de $\mathbb{F}_p$ (théorème de Lagrange), qui vaut p, donc $k = \{0\}$ ou $\mathbb{F}_p$. Comme $1 \in K$, on a donc $k = \mathbb{F}_p$.

Si P est un polynôme de $\mathbb{F}_p[X]$ irréductible dans $F_q[X]$, il est irréductible dans $\mathbb{F}_p[X]$. En effet, si $P = QR$ avec $Q, R \in \mathbb{F}_p[X]$, alors $P, Q, R \in \mathbb{F}_q[X]$, et donc Q ou R est de degré 0.

Définition (Fonction de Möbius)**.** On note μ la fonction de Möbius définie par

$$\begin{array}{rccl} \mu : & \mathbb{N}^* & \longrightarrow & \{-1, 0, 1\} \\ & & & \\ & n & \longmapsto & \begin{cases} 1 \text{ si } n = 1, \\ 0 \text{ si } n \text{ a un facteur carré}, \\ (-1)^r \text{ sinon, où } r \text{ est le nombre de facteurs premiers de } n. \end{cases} \end{array}$$

Proposition (Formule d'inversion de Möbius)**.** *Soient $f, g : \mathbb{N} \to \mathbb{C}$ telles que pour tout $n \geqslant 1$,*

$$g(n) = \sum_{d|n} f(d).$$

Alors pour tout $n \geqslant 1$,

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d).$$

Démonstration. Si f et g sont multiplicatives, alors $f * g$ aussi. Donc $\mu * \mathbb{1}$ est multiplicative et pour la connaître, il suffit de calculer $\mu * \mathbb{1}(p^\alpha)$:

$$\mu * \mathbb{1}(p^\alpha) = \mu(1) + \mu(p) + \underbrace{\mu(p^2)}_{=0} + \cdots + \underbrace{\mu(p^\alpha)}_{=0} = 1 - 1 = 0.$$

Donc immédiatement, on obtient $\mu * \mathbb{1} = \delta$, d'où le résultat. □

Références :

– Francinou et Gianella - *Exercices de mathématiques pour l'agrégation, Algèbre 1* - Page 189.

2.7 Dénombrement des solutions d'une équation diophantienne

Proposition. *Soient $\alpha_1, \dots, \alpha_r \in \mathbb{N}^*$ des entiers naturels non nuls premiers entre eux dans leur ensemble. Pour $n \in \mathbb{N}$, on considère l'équation diophantienne (E_n) : $\alpha_1 n_1 + \cdots + \alpha_r n_r = n$ d'inconnues $n_1, \dots, n_r \in \mathbb{N}$.*

Alors l'équation (E_n) admet un nombre fini de solutions s_n que l'on peut calculer de manière explicite. De plus

$$s_n \underset{n\to\infty}{\sim} \frac{1}{\alpha_1 \dots \alpha_r} \frac{n^{r-1}}{(r-1)!}.$$

Démonstration. Soit $F(X) = \prod_{i=1}^r \frac{1}{1-X^{\alpha_i}}$. Dans $\mathbb{C}[[X]]$, on a

$$\begin{aligned} F(X) &= \prod_{i=1}^r \left(\sum_{n=0}^\infty X^{n\alpha_i} \right) = \prod_{i=1}^r \left(\sum_{n=0}^\infty X^n \mathbb{1}_{\{n \in \alpha_i \mathbb{N}\}} \right) \\ &= \sum_{n=0}^\infty \left(\sum_{n_1 + \cdots + n_r = n} X^{n_1} \mathbb{1}_{\{n_1 \in \alpha_1 \mathbb{N}\}} \dots X^{n_r} \mathbb{1}_{\{n_r \in \alpha_r \mathbb{N}\}} \right) \\ &= \sum_{n=0}^\infty \left(\sum_{n_1\alpha_1 + \cdots + n_r\alpha_r = n} X^{n_1\alpha_1} \dots X^{n_r\alpha_r} \right) \\ &= \sum_{n=0}^\infty \left(\sum_{n_1\alpha_1 + \cdots + n_r\alpha_r = n} 1 \right) X^n = \sum_{n=0}^\infty s_n X^n. \end{aligned}$$

On note U_n le groupe des racines n-ième de l'unité dans $\mathbb{C}$. On décompose F en éléments simples dans $\mathbb{C}(X)$:

$$F(X) = \sum_{\omega \in \bigcup_{i=1}^r U_{\alpha_i}} \left(\frac{a_{\omega,1}}{\omega - X} + \cdots + \frac{a_{\omega,m_\omega}}{(\omega - X)^{m_\omega}} \right),$$

où m_ω est la multiplicité du pôle ω de F (attention, m_ω n'est pas nécessairement égal à 1 car ω peut être racine de $1 - X^{\alpha_i}$ et de $1 - X^{\alpha_j}$ pour $i \neq j$).

Ensuite, on a

$$\begin{aligned} \frac{1}{(\omega - X)^k} &= \frac{1}{(k-1)!} \left(\frac{1}{\omega - X} \right)^{(k-1)} = \frac{1}{(k-1)!} \left(\sum_{n=0}^\infty \frac{1}{\omega} \left(\frac{X}{\omega} \right)^n \right)^{(k-1)} \\ &= \frac{1}{(k-1)!} \sum_{n=k-1}^\infty n(n-1)\dots(n-(k-2)) \frac{1}{\omega^{n+1}} X^{n-(k-1)} \\ &= \frac{1}{(k-1)!} \sum_{m=0}^\infty (m+k-1)\dots(m+1) \frac{1}{\omega^{m+k}} X^m. \end{aligned}$$

En reportant dans l'expression de $F(X)$ et en identifiant les coefficients, on en déduit une expression de s_n qui est donc calculable explicitement.

Comme $F(X) = \prod_{i=1}^{r} \frac{1}{1-X^{\alpha_i}}$, 1 est un pôle de F d'ordre r. Soit ω un pôle de F. Puisque chaque polynôme $1 - X^{\alpha_i}$ est à racines simples, $m_\omega \leqslant r$. Si $m_\omega = r$, alors ω est racine de tous les $1 - X^{\alpha_i}$. Comme les α_i sont premiers entre eux dans leur ensemble, d'après l'identité de Bézout, il existe $u_1, \dots, u_r \in \mathbb{Z}$ tels que $u_1\alpha_1 + \cdots + u_r\alpha_r = 1$. Ainsi, comme $\omega^{\alpha_i} = 1$ pour tout i, on a $\omega = \omega^{u_1\alpha_1+\cdots+u_r\alpha_r} = 1$. Finalement, pour tout $\omega \neq 1$ pôle de F, on a $m_\omega < r$.

Le terme général de la série $\frac{1}{(\omega-X)^k}$ vaut, d'après le calcul ci-dessus,

$$\frac{1}{(k-1)!}(n+1)\dots(n+k-1)\frac{1}{\omega^{n+k}} = O\left(n^{k-1}\right)$$

car $|\omega| = 1$. Donc pour tout $k < r$, c'est un $O(n^{r-2})$, soit encore un $o(n^{r-1})$. On en déduit que dans l'expression de s_n, les contributions des $\frac{1}{(\omega-X)^k}$ sont négligeables devant celle de $\frac{1}{(1-X)^r}$ quand $n \to \infty$. Ainsi,

$$s_n \underset{n\to\infty}{\sim} a_{1,r}\frac{n^{r-1}}{(r-1)!}.$$

Il ne reste plus qu'à calculer $a_{1,r}$. On a

$$(1-X)^r F(X) = \prod_{i=1}^{r} \frac{1-X}{1-X^{\alpha_i}} = \prod_{i=1}^{r} \frac{1}{1+X+\cdots+X^{\alpha_i-1}}.$$

Comme $a_{1,r} = [(1-z)^r F(z)](1)$, on obtient $a_{1,r} = \frac{1}{\alpha_1 \dots \alpha_r}$, d'où le résultat. □

Application. *Le nombre de solutions $(x, y, z) \in \mathbb{N}^3$ de l'équation $x + 2y + 3z = n$ est*

$$s_n = \frac{(n+1)(n+5)}{12} + \frac{17}{72} + \frac{(-1)^n}{8} + \frac{2}{9}\cos\left(\frac{2k\pi}{3}\right).$$

Théorème (Décomposition en éléments simples)**.** *Soit $\frac{P}{Q} \in K(X)$. On écrit la décomposition de Q en produit de polynômes unitaires irréductibles : $Q = \lambda Q_1^{\alpha_1} \dots Q_r^{\alpha_r}$. Alors on peut écrire*

$$\frac{P}{Q} = R + \frac{A_{1,1}}{Q_1} + \cdots + \frac{A_{1,\alpha_1}}{Q_1^{\alpha_1}} + \cdots + \frac{A_{k,1}}{Q_k} + \cdots + \frac{A_{k,\alpha_k}}{Q_k^{\alpha_k}}$$

où R et les $A_{i,j}$ sont des polynômes de $K[X]$ tels que $\deg(A_{i,j}) < \deg(Q_i)$ pour tout i, j.

Démonstration. Les polynômes $T_i = \lambda Q_1^{\alpha_1} \dots \widehat{Q_i^{\alpha_i}} \dots Q_r^{\alpha_r}$ sont premiers entre eux dans leur ensemble, donc d'après l'identité de Bézout, on a

$$S_1 T_1 + \cdots + S_r T_r = 1.$$

On en déduit l'écriture, en notant $B_i = PS_i$:

$$\frac{P}{Q} = \frac{PS_1T_1 + \cdots + PS_rT_r}{Q} = \frac{B_1}{Q_1^{\alpha_1}} + \cdots + \frac{B_r}{Q_r^{\alpha_r}}.$$

A présent, on effectue la division euclidienne de B_1 par Q_1 : $B_1 = B_{1,\alpha_1}Q_1 + A_{1,\alpha_1}$ avec $\deg(A_{1,\alpha_1}) < \deg(Q_1)$. On obtient alors

$$\frac{B_1}{Q_1^{\alpha_1}} = \frac{B_{1,\alpha_1}}{Q_1^{\alpha_1 - 1}} + \frac{A_{1,\alpha_1}}{Q_1^{\alpha_1}}.$$

On effectue une nouvelle division euclidienne sur B_{1,α_1}, par Q_1, et en recommençant un nombre suffisant de fois, on trouve

$$\frac{B_1}{Q_1^{\alpha_1}} = B_{1,1} + \frac{A_{1,1}}{Q_1} + \cdots + \frac{A_{1,\alpha_1}}{Q_1^{\alpha_1}},$$

d'où la décomposition en éléments simples de $\frac{P}{Q}$. □

Références :

- Chambert-Loir - *Analyse 1.*
- Francinou, Gianella, Nicolas - *Oraux X-ENS, Analyse 2* - Page 197.
- Gourdon - *Analyse.*

2.8 Ellipse de Steiner et application

Soit ABC un triangle du plan affine euclidien $\mathcal{P}$, qu'on suppose non applati et non équilatéral. On prend pour origine le point O centre de gravité du triangle ABC et on rapporte le plan à un repère orthonormé direct $(O, \overrightarrow{u}, \overrightarrow{v})$. On note respectivement a, b, c les affixes de A, B, C et I, J, K les points d'affixes respectives $1, j, j^2$.

Lemme. *Il existe $\alpha, \beta \in \mathbb{C}^*$ tels que l'application $f : \mathcal{P} \to \mathcal{P}$ associée à la transformation $\varphi : z \mapsto \alpha z + \beta \overline{z}$ envoie respectivement les points I, J, K sur A, B, C.*

Démonstration. On cherche α, β tels que $\alpha + \beta = a$, $\alpha j + \beta j^2 = b$ et $\alpha j^2 + \beta j = c$. Il s'agit d'un système de 3 équations à 2 inconnues, et pour avoir une solution, il est nécessaire que ces équations soient liées. On vérifie :

$$\det \begin{pmatrix} 1 & 1 & a \\ j & j^2 & b \\ j^2 & j & c \end{pmatrix} = (a+b+c)(j^2 - j) = 0$$

car O étant le centre de gravité de ABC, on a $a + b + c = 0$. D'autre part, comme $\det \begin{pmatrix} 1 & 1 \\ j & j^2 \end{pmatrix} \neq 0$, ce système est de rang 2. Il existe donc un unique couple $(\alpha, \beta) \in \mathbb{C}^2$ vérifiant les 3 équations. Si $\alpha = 0$, alors le triangle ABC est l'image du triangle IJK par la similitude indirecte $z \mapsto \beta \overline{z}$ qui conserve les rapports de distances, il est donc équilatéral comme IJK, ce qui est exclu. Donc $\alpha \neq 0$, et de même $\beta \neq 0$. □

Proposition. *Les affixes des foyers de l'ellipse de Steiner du triangle ABC sont les deux racines carrées du nombre complexe $\alpha\beta$.*

Démonstration. Soit C le cercle de Steiner du triangle équilatéral IJK. Comme f est une transformation affine bijective (car par exemple φ envoie la base $(\overrightarrow{IJ}, \overrightarrow{IK})$ sur la base $(\overrightarrow{AB}, \overrightarrow{AC})$), elle envoie C sur une conique $C' = f(C)$ de même nature, *i.e.* une ellipse. Comme les barycentres sont conservés par les applications affines, les milieux des côtés du triangle IJK sont envoyés sur les milieux des côtés du triangle ABC. En admettant que C' est tangente à chacun des côtés de ABC en leurs milieux (voir plus bas la démonstration de l'existence de l'ellipse de Steiner), par unicité de l'ellipse de Steiner, C' est l'ellipse de Steiner du triangle ABC.

Comme C est le cercle inscrit au triangle équilatéral IJK, c'est le cercle de centre O et de rayon $\frac{1}{2}$ (il passe par le milieu de JK qui est en $z = -\frac{1}{2}$). Donc $M \in C$ si et seulement s'il existe t tel que l'affixe de M dans $(O, \overrightarrow{u}, \overrightarrow{v})$ soit $z = \frac{1}{2}e^{it}$. Donc $M' = f(M) \in C'$ si et seulement s'il existe t tel que l'affixe de M' dans $(O, \overrightarrow{u}, \overrightarrow{v})$ soit $z' = \varphi(z) = \frac{1}{2}(\alpha e^{it} + \beta e^{-it})$.

On note $\alpha = |\alpha|e^{i\varphi}$, $\beta = |\beta|e^{i\psi}$, $\overrightarrow{u_1}$ le vecteur d'affixe $e^{i\frac{\varphi+\psi}{2}}$ et $\overrightarrow{v_1}$ le vecteur d'affixe $ie^{i\frac{\varphi+\psi}{2}}$ (on a fait tourner le repère d'un angle $\frac{\varphi+\psi}{2}$). Alors $M' \in C'$ si et seulement s'il existe t tel que l'affixe de M' dans $(O, \overrightarrow{u_1}, \overrightarrow{v_1})$ soit $z'_1 = z'e^{-i\frac{\varphi+\psi}{2}}$. On a alors

$$\begin{aligned} z'_1 = \frac{1}{2}(\alpha e^{it} + \beta e^{-it})e^{-i\frac{\varphi+\psi}{2}} &= \frac{1}{2}(|\alpha|e^{i(\varphi+t)} + |\beta|e^{i(\psi-t)})e^{-i\frac{\varphi+\psi}{2}} \\ &= \frac{1}{2}(|\alpha|e^{i\frac{\varphi-\psi}{2}+it} + |\beta|e^{-i\frac{\varphi-\psi}{2}-it}). \end{aligned}$$

On en déduit une représentation paramétrique de C' dans le repère $(O, \overrightarrow{u_1}, \overrightarrow{v_1})$:

$$\begin{cases} x'_1 = \frac{1}{2}(|\alpha| + |\beta|)\cos\left(\frac{\varphi-\psi}{2} + t\right) \\ y'_1 = \frac{1}{2}(|\alpha| - |\beta|)\sin\left(\frac{\varphi-\psi}{2} + t\right). \end{cases}$$

Une équation de C' dans ce repère est donc

$$\frac{x'^2_1}{\left(\frac{|\alpha|+|\beta|}{2}\right)^2} + \frac{y'^2_1}{\left(\frac{|\alpha|-|\beta|}{2}\right)^2} = 1.$$

Les foyers F et F' de C' appartiennent au grand axe qui est dirigé par $\overrightarrow{u_1}$ et O est leur centre. Dans $(O, \overrightarrow{u}, \overrightarrow{v})$, un argument de F est donc $\frac{\varphi+\psi}{2}$ et un argument de F' est $\frac{\varphi+\psi}{2} + \pi$. Ensuite on note $c = OF$. Si N désigne un point de C sur la médiatrice de $[FF']$ (*i.e.* N est un sommet de l'ellipse sur son petit axe), alors on a $ON = \frac{|\alpha|-|\beta|}{2}$, $OF = c$ et $NF = \frac{|\alpha|+|\beta|}{2}$ (le demi-grand axe, car par la définition bifocale, $NF + NF' = 2\frac{|\alpha|+|\beta|}{2}$ et comme N est sur la médiatrice, $NF = NF'$). Par le théorème de Pythagore, on a donc

$$c^2 = \left(\frac{|\alpha| + |\beta|}{2}\right)^2 - \left(\frac{|\alpha| - |\beta|}{2}\right)^2 = |\alpha\beta|.$$

Donc les affixes de F et F' dans $(O, \overrightarrow{u}, \overrightarrow{v})$ sont $\pm\sqrt{|\alpha\beta|}e^{i\frac{\varphi+\psi}{2}}$, *i.e.* les deux racines carrées de $\alpha\beta$. □

Application. *Si $Q = (X-a)(X-b)(X-c)$, alors les racines de Q' sont les affixes des foyers de l'ellipse de Steiner du triangle ABC.*

Démonstration. On a

$$Q = X^3 - (a+b+c)X^2 + (ab+bc+ca)X - abc.$$

On rappelle que $a + b + c = 0$ puisque O est le centre de gravité du triangle ABC. On calcule ensuite, en se rappelant que α et β sont tels que $\alpha + \beta = a$, $\alpha j + \beta j^2 = b$ et $\alpha j^2 + \beta j = c$:

$$ab + bc + ca = (\alpha+\beta)(\alpha j + \beta j^2) + (\alpha j + \beta j^2)(\alpha j^2 + \beta j) + (\alpha j^2 + \beta j)(\alpha + \beta)$$

$$\begin{aligned} &= (1+j+j^2)\alpha^2 + (1+j+j^2)\beta^2 + 3\alpha\beta(j+j^2) \\ &= -3\alpha\beta, \end{aligned}$$

car $1+j+j^2 = 0$. Donc

$$Q = X^3 - 3\alpha\beta X - abc.$$

On obtient ainsi que $Q' = 3X^2 - 3\alpha\beta$. Les racines de Q' sont donc les racines de $X^2 - \alpha\beta$, *i.e.* les deux racines carrées de $\alpha\beta$, qui sont les affixes des foyers de l'ellipse de Steiner de ABC d'après la proposition précédente. □

Théorème. *L'image d'une conique par une application affine bijective est une conique de même nature.*

Démonstration. Soit f une application affine bijective associée à une application linéaire bijective φ. Soit $(O, \overrightarrow{u}, \overrightarrow{v})$ un repère. Alors les coordonnées du point M dans $(O, \overrightarrow{u}, \overrightarrow{v})$ sont les mêmes que les coordonnées de $f(M)$ dans $(f(O), \varphi(\overrightarrow{u}), \varphi(\overrightarrow{v}))$ (qui est bien un repère puisque φ est bijective et donc conserve les bases). En effet, par bijectivité de φ, on a l'équivalence

$$\overrightarrow{OM} = x\overrightarrow{u} + y\overrightarrow{v} \iff \varphi\left(\overrightarrow{OM}\right) = \overrightarrow{f(O)f(M)} = x\varphi(\overrightarrow{u}) + y\varphi(\overrightarrow{v}),$$

où l'on a utilisé que $\varphi\left(\overrightarrow{MN}\right) = \overrightarrow{f(M)f(N)}$.

Ainsi, si C est une conique d'équation (E) dans $(O, \overrightarrow{u}, \overrightarrow{v})$, alors son image $C' = f(C)$ est définie par la même équation dans $(f(O), \varphi(\overrightarrow{u}), \varphi(\overrightarrow{v}))$. Donc C' est une conique de même nature que C. □

Théorème. *Pour tout triangle du plan, il existe une unique ellipse tangente à chacun des côtés en leur milieu, appelée* ellipse de Steiner.

Démonstration. Nous ne montrerons que l'existence, l'unicité est plus délicate.

Soit ABC un triangle non applati du plan affine. On note A', B', C' les milieux respectifs des segments $[BC], [AC], [AB]$.

Comme les points A, B, C ne sont pas alignés, ils définissent un repère du plan affine. Si le triangle ABC est équilatéral, l'existence d'une ellipse répondant au problème est évidente : il s'agit du cercle inscrit au triangle. L'idée est donc de se ramener à un triangle équilatéral.

Soit $A_0B_0C_0$ un triangle équilatéral. Il existe une unique bijection affine telle que

$$g(A_0) = A,\ g(B_0) = B \text{ et } g(C_0) = C.$$

En effet, g est définie par $g(A_0) = A$, $\overrightarrow{g}(\overrightarrow{A_0B_0}) = \overrightarrow{AB}$ et $\overrightarrow{g}(\overrightarrow{A_0C_0}) = \overrightarrow{AC}$, et on note que $\overrightarrow{g}$ est uniquement déterminée par l'envoi d'une base sur une base. Ainsi par g, le triangle équilatéral $A_0B_0C_0$ est envoyé sur le triangle ABC.

Comme les barycentres sont conservés par les applications affines, les milieux des côtés du triangle $A_0B_0C_0$ sont envoyés sur les milieux des côtés du triangle ABC.

De plus, l'image d'une conique par une application affine bijective reste une conique de même nature. Donc g envoie le cercle inscrit C de $A_0B_0C_0$ sur une ellipse E.

Enfin, une application affine est différentiable, de différentielle sa partie linéaire. En effet, si M et N sont deux points, on a

$$g(M) = g(N) + \overrightarrow{g}(\overrightarrow{NM}) = g(N) + Dg(N).\overrightarrow{NM} + o\left(\|\overrightarrow{NM}\|\right),$$

et donc $Dg(N) = \overrightarrow{g}$. Par conséquent, si $\overrightarrow{u}$ est un vecteur directeur de la tangente en un point M de C, alors $\overrightarrow{g}(\overrightarrow{u})$ dirige la tangente à l'ellipse E au point $g(M)$.

Ainsi, comme la droite (A_0B_0) est tangente à C en C_0' le milieu de $[A_0B_0]$, on obtient que $\overrightarrow{g}(\overrightarrow{A_0B_0}) = \overrightarrow{g(A_0)g(B_0)} = \overrightarrow{AB}$ dirige la tangente à E en $g(C_0') = C'$. Donc la droite (AB) est tangente à E en C'. Le raisonnement est le même pour les autres tangences. L'ellipse E est donc tangente au triangle ABC aux points A', B', C'. □

Proposition. *On se place dans un repère $\mathcal{R}$. Alors l'ellipse d'équation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ avec $a \geqslant b$ est exactement l'ensemble des points $M = (x, y)$ vérifiant $MF + MF' = 2a$ où $F = (c, 0)$ et $F' = (-c, 0)$ avec $c = \sqrt{a^2 - b^2}$. Les points F et F' sont appelés les foyers de l'ellipse.*

Démonstration.

$$\begin{aligned}
MF + MF' = 2a &\iff (MF + MF')^2 = 4a^2 \\
&\iff (x-c)^2 + y^2 + (x+c)^2 + y^2 + 2MF.MF' = 4a^2 \\
&\iff 2x^2 + 2y^2 + 2c^2 + 2MF.MF' = 4a^2 \\
&\iff MF.MF' = 2a^2 - c^2 - x^2 - y^2 \\
&\implies (MF.MF')^2 = (a^2 + b^2 - x^2 - y^2)^2 \\
&\iff (x^2 - 2cx + c^2 + y^2)(x^2 + 2cx + c^2 + y^2) \\
&\quad = (a^2 + b^2 - x^2 - y^2)^2 \\
&\iff (x^2 + y^2 + c^2)^2 - 4c^2x^2 = (a^2 + b^2 - x^2 - y^2)^2 \\
&\iff (x^2 + y^2 + c^2)^2 - (a^2 + b^2 - x^2 - y^2)^2 = 4c^2x^2 \\
&\iff (a^2 + b^2 + c^2)(2x^2 + 2y^2 + c^2 - a^2 - b^2) = 4c^2x^2 \\
&\iff 2a^2.2(x^2 + y^2 - b^2) = 4c^2x^2
\end{aligned}$$

$$\begin{aligned}
&\iff a^2x^2 + a^2y^2 - a^2b^2 = c^2x^2 \\
&\iff (a^2 - c^2)x^2 + a^2y^2 = a^2b^2 \\
&\iff b^2x^2 + a^2y^2 = a^2b^2 \\
&\iff \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.
\end{aligned}$$

Il reste à prouver que l'implication au milieu est en fait une équivalence. A la fin, on obtient que nécessairement, $x^2 \leqslant a^2$ et $y^2 \leqslant b^2$, donc $a^2 + b^2 - x^2 - y^2 \geqslant 0$ et on a bien l'équivalence. □

Références :

- Deuxième composition du CAPES externe de 1990.

2.9 Irréductibilité des polynômes cyclotomiques sur $\mathbb{Q}$

Théorème. *Le polynôme cyclotomique Φ_n est irréductible sur $\mathbb{Q}$.*

Démonstration. Soit K un corps de décomposition de Φ_n sur $\mathbb{Q}$. On note ω une racine primitive n-ième de 1, et on fixe un nombre premier p ne divisant pas n.

On remarque que ω^p est une autre racine primitive de 1. En effet, les racines primitives de 1 sont les ω^m avec m et n premiers entre eux : si $\omega = e^{\frac{2ik\pi}{n}}$, alors $(k,n) = 1$ (k et n sont premiers entre eux car ω est une racine primitive de 1), et donc $(kp,n) = 1$ comme p ne divise pas n.

Soit $P \in \mathbb{Q}[X]$ le polynôme minimal de ω sur $\mathbb{Q}$ et $Q \in \mathbb{Q}[X]$ celui de ω^p. Nous allons montrer que $P \in \mathbb{Z}[X]$ et $P = \Phi_n$.

Comme $\mathbb{Z}[X]$ est factoriel, on peut écrire $\Phi_n = R_1^{\alpha_1} \dots R_r^{\alpha_r}$ avec les $R_i \in \mathbb{Z}[X]$ irréductibles. Puisque Φ_n est unitaire, quitte à multiplier les R_i par -1, on peut les supposer unitaires. Enfin, ω est racine de Φ_n, donc de l'un des R_i. Or R_i est unitaire et irréductible sur $\mathbb{Z}$, donc sur $\mathbb{Q}$ (un polynôme primitif de $\mathbb{Z}[X]$ de degré $\geqslant 1$ non irréductible sur $\mathbb{Q}$ ne peut pas être irréductible sur $\mathbb{Z}$, voir les irréductibles d'un anneau de polynômes), et donc R_i est le polynôme minimal de ω sur $\mathbb{Q}$, *i.e.* $P = R_i \in \mathbb{Z}[X]$. De même, $Q \in \mathbb{Z}[X]$.

On note au passage que P et Q divisent Φ_n dans $\mathbb{Z}[X]$.

Montrons maintenant que $P = Q$. On suppose $P \neq Q$. Comme P et Q sont irréductibles et distincts, leur produit PQ divise Φ_n dans $\mathbb{Q}[X]$. D'autre part, ω^p est racine de Q, donc ω est racine de $Q(X^p)$, et donc P divise $Q(X^p)$ dans $\mathbb{Q}[X]$ (car P est le polynôme minimal de ω sur $\mathbb{Q}$) : il existe $R \in \mathbb{Q}[X]$ tel que

$$Q(X^p) = P(X)R(X).$$

En écrivant $R = \frac{a}{b}R_2$ avec $R_2 \in \mathbb{Z}[X]$ primitif, $a, b \in \mathbb{Z}$, a et b premiers entre eux, et en prenant les contenus, on obtient :

$$c(bQ(X^p)) = b\,c(Q(X^p)) = c(aPR_2) = a\,c(P)c(R_2).$$

Or P et $Q(X^p)$ sont unitaires, leurs contenus valent donc 1. D'où $b = \pm a$ (le contenu est défini modulo $\mathbb{Z}^*$), et donc $R \in \mathbb{Z}[X]$.

On projette l'égalité $Q(X^p) = P(X)R(X)$ dans $\mathbb{F}_p$. On écrit $Q = a_r X^r + \dots + a_0$ avec $a_i \in \mathbb{Z}$. Alors comme $\overline{a_i} = \overline{a_i}^p$ pour tout i, on a :

$$\overline{Q}(X^p) \;=\; \overline{a_r}X^{pr} + \dots + \overline{a_0}$$

$$\begin{aligned} &= \overline{a_r}^p X^{pr} + \cdots + \overline{a_0}^p \\ &= (\overline{a_r} X^r + \cdots + \overline{a_0})^p \\ &= \overline{Q}(X)^p. \end{aligned}$$

Soit $\overline{S}$ un facteur irréductible de $\overline{P}$ sur $\mathbb{F}_p$. Puisque $\overline{Q}^p = \overline{PR}$, on obtient $\overline{S}$ divise $\overline{Q}$. Comme PQ divise Φ_n sur $\mathbb{Q}$ (on est toujours sous l'hypothèse $P \neq Q$ faite plus haut), il divise Φ_n sur $\mathbb{Z}$ (grâce aux contenus, de la même façon que P divise $Q(X^p)$ sur $\mathbb{Z}$), alors $\overline{PQ}$ divise $\overline{\Phi_n}$ sur $\mathbb{F}_p$, et donc $\overline{S}^2$ divise $\overline{\Phi_n}$ sur $\mathbb{F}_p$. On en déduit que Φ_n a une racine double (dans un corps de rupture de $\overline{S}$ sur $\mathbb{F}_p$). Mais $\overline{\Phi_n}$ n'a pas de racine double : en effet, sinon, comme $\overline{\Phi_n}$ divise $\overline{X^n - 1}$, ce polynôme aurait une racine double. Ce n'est pas le cas puisque sa dérivée vaut $\overline{n} X^{n-1}$ ayant 0 pour seule racine (car p et n sont premiers entre eux), et 0 n'est pas racine de $\overline{X^n - 1}$.

On a donc une absurdité qui nous permet de conclure que $P = Q$.

Soit ω' une racine primitive n-ième de 1. Alors il existe m premier avec n tel que $\omega' = \omega^m$. On écrit $m = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ avec p_i ne divisant pas n. Par une récurrence sur le nombre de facteurs premiers de m, on montre que ω' et ω ont même polynôme minimal sur $\mathbb{Q}$. En effet, ω et ω^{p_1} ont même polynôme minimal d'après ce qui précède, puis ω^{p_1} et $(\omega^{p_1})^{p_1} = \omega^{p_1^2}$ ont même polynôme minimal, etc... Donc $P(\omega') = 0$ pour toute racine primitive n-ième de 1, et par conséquent,

$$\Phi_n \text{ divise } P \text{ sur } \mathbb{Q}.$$

Comme on avait déjà P divise Φ_n et que les polynômes P et Φ_n sont unitaires, on en déduit $\Phi_n = P$ et Φ_n est irréductible sur $\mathbb{Q}$. □

Remarque. Comme Φ_n est irréductible sur $\mathbb{Q}$, que $\Phi_n \in \mathbb{Z}[X]$ et que son contenu vaut 1 (car il est unitaire), on en déduit que Φ_n est irréductible sur $\mathbb{Z}$.

Définition (Polynôme cyclotomique d'ordre n)**.** Pour $n \in \mathbb{N}^*$, on définit Φ_n le n-ième polynôme cyclotomique sur $\mathbb{C}$ par

$$\Phi_n = \prod_{\omega \in U_n^*} (X - \omega),$$

où $U_n = \left\{ e^{\frac{2ik\pi}{n}}, k \in \mathbb{Z} \right\}$ et U_n^* est l'ensemble des racines primitives n-ième de l'unité.

Proposition. *On a :*

$$X^n - 1 = \prod_{d|n} \Phi_d.$$

Démonstration. Soit $\omega = e^{\frac{2i\pi}{n}}$. Soient $k \in [\![0, n-1]\!]$ et d l'ordre de ω^k dans U_n. On a donc que d divise n. D'autre part, $(\omega^k)^d = 1$, donc $\omega^k \in U_d$, et étant d'ordre d, on a même $\omega^k \in U_d^*$. Ainsi $(X - \omega^k)$ divise Φ_d, donc divise $\prod_{d|n} \Phi_d$.

Les ω^k pour $k \in [\![0, n-1]\!]$ étant deux à deux distincts, on en déduit que :

$$X^n - 1 = \prod_{k=0}^{n-1} (X - \omega^k) \text{ divise } \prod_{d|n} \Phi_d.$$

Or

$$\deg\left(\prod_{d|n} \Phi_d\right) = \sum_{d|n} \deg(\Phi_d) = \sum_{d|n} \varphi(d) = n = \deg(X^n - 1).$$

Les deux polynômes $X^n - 1$ et $\prod_{d|n} \Phi_d$ sont unitaires et de même degré, avec le premier qui divise le deuxième, ils sont donc égaux. □

Proposition. *Pour tout $n \in \mathbb{N}^*$, $\Phi_n \in \mathbb{Z}[X]$.*

Démonstration. On le montre par récurrence :

Pour $n = 1$, $\Phi_1 = X - 1 \in \mathbb{Z}[X]$.

On suppose maintenant que pour tout $k \leqslant n-1$, $\Phi_k \in \mathbb{Z}[X]$. Alors le polynôme $P = \prod_{\substack{d|n \\ d \neq n}} \Phi_d \in \mathbb{Z}[X]$ et on a $X^n - 1 = P\Phi_n$ d'après la proposition précédente.

Comme P est unitaire, son coefficient dominant est inversible dans $\mathbb{Z}$ et on peut effectuer la division euclidienne de $X^n - 1$ par P dans $\mathbb{Z}[X]$: il existe $Q, R \in \mathbb{Z}[X]$ tels que

$$X^n - 1 = PQ + R$$

avec $\deg(R) < \deg(P)$. Comme on a d'autre part $X^n - 1 = P\Phi_n$ dans $\mathbb{C}[X]$ (a priori, on a seulement $\Phi_n \in \mathbb{C}[X]$), en faisant la différence, il vient $P(\Phi_n - Q) = R$. Or $\deg(R) < \deg(P)$, donc nécessairement $\Phi_n = Q \in \mathbb{Z}[X]$. □

Définition (Contenu d'un polynôme)**.** Soit A un anneau. Pour $P \in A[X]$, $P = a_n X^n + \cdots + a_0 \neq 0$, on définit le contenu de P, noté $c(P)$, par :

$$c(P) = \mathrm{PGCD}(a_0, \ldots, a_n).$$

Le contenu est bien sûr défini modulo A^*.

Proposition (Lemme de Gauss)**.** *Soit A un anneau factoriel. Pour $P, Q \in A[X]$ non nuls, on a :*

$$c(PQ) = c(P)c(Q) \mod A^*.$$

Démonstration. On suppose d'abord P et Q primitifs, *i.e.* $c(P) = c(Q) = 1$. Si $c(PQ) \neq 1$, comme A est factoriel, il existe $p \in A$ irréductible qui divise $c(PQ)$.

On écrit $P = a_n X^n + \cdots + a_0$ et $Q = b_m X^m + \cdots + b_0$. Comme $c(P) = 1$, p ne divise pas tous les a_i. Soit i_0 le plus petit i tel que p ne divise pas a_i. De même, soit j_0 le plus petit j tel que p ne divise pas b_j. On a $PQ = c_{n+m}X^{n+m} + \cdots + c_0$. En particulier,

$$c_{i_0+j_0} = \sum_{i+j=i_0+j_0} a_i b_j = a_{i_0}b_{j_0} + \sum_{\substack{i+j=i_0+j_0 \\ i<i_0 \text{ ou } j<j_0}} a_i b_j.$$

Alors p divise $c_{i_0+j_0}$ (il divise tous les c_k) et p divise $\sum_{\substack{i+j=i_0+j_0 \\ i<i_0 \text{ ou } j<j_0}} a_i b_j$. Donc p divise $a_{i_0}b_{j_0}$, ce qui est absurde, car p étant irréductible, cela implique qu'il divise a_{i_0} ou b_{j_0}. Finalement

$$c(PQ) = 1.$$

Remarque : Dans le cas $A = \mathbb{Z}$, on peut réduire modulo p : $\overline{PQ} = 0$ dans $\mathbb{F}_p[X]$. Et comme $\mathbb{F}_p[X]$ est intègre, cela implique $\overline{P} = 0$ ou $\overline{Q} = 0$, ce qui est absurde puisque $c(P) = c(Q) = 1$.

On suppose maintenant P et Q quelconques, et on note $a = c(P)$, $b = c(Q)$. Alors $P_2 = \frac{1}{a}P$ et $Q_2 = \frac{1}{b}Q$ sont primitifs, donc $c(P_2Q_2) = 1$ d'après le cas précédent. Par conséquent, $PQ = abP_2Q_2$ et le PGCD des coefficient de PQ vaut clairement $ab = c(P)c(Q)$. □

Proposition (Irréductibles d'un anneau de polynômes)**.** *Soit A un anneau et K son corps des fractions. Les polynômes irréductibles dans $A[X]$ sont :*

(i) Les constantes $p \in A$ irréductibles dans A.

(ii) Les polynômes P de degré $\geqslant 1$ primitifs et irréductibles dans $K[X]$.

Démonstration. On vérifie que ces éléments sont bien irréductibles dans $A[X]$:

(*i*) Si $p \in A$ et si $p = P(X)Q(X)$, alors $\deg(P) = \deg(Q) = 0$, donc $P, Q \in A$. En supposant p irréductible, on a donc $P \in A^*$ ou $Q \in A^*$, donc $P \in A[X]^*$ ou $Q \in A[X]^*$, *i.e.* p est irréductible dans $A[X]$.

(*ii*) Soit P de degré $\geqslant 1$ primitif et irréductible dans $K[X]$. Si $P = QR$ avec $Q, R \in A[X]$, alors comme Q et R sont aussi dans $K[X]$, on a Q ou R dans $K[X]^*$. Supposons par exemple que $Q \in K[X]^*$. Alors $\deg(Q) = 0$ avec $Q \neq 0$, donc $Q = a \in A$. On en déduit que $P = aR$, donc a divise $c(P)$, et comme $c(P) = 1$, $a \in A^*$ et P est irréductible dans $A[X]$.

Montrons maintenant que ces éléments sont les seuls irréductibles dans $A[X]$. Soit $P \in A[X]$ irréductible dans $A[X]$.

(*i*) Si $\deg(P) = 0$, $P = p \in A$ est irréductible dans $A[X]$, donc clairement dans A.

(*ii*) Si $\deg(P) > 0$, on a nécessairement $c(P) = 1$ (on rappelle que le contenu est défini modulo A^*). En effet, on aurait sinon $P = c(P)Q$ avec $c(P) \notin A^*$ et $Q \in A[X]$. Comme P est irréductible et que $c(P) \notin A[X]^*$ (car $A[X]^* = A^*$), on a $Q \in A[X]^*$. Donc $Q \in A^*$ et alors $\deg(P) = 0$.

Il reste à montrer que P est irréductible dans $K[X]$. On écrit $P = QR$ avec $Q, R \in K[X]$. On peut écrire $Q = \frac{a}{b}Q_2$ et $R = \frac{c}{d}R_2$ avec $Q_2, R_2 \in A[X]$ primitifs et $a, b, c, d \in A$. On a alors

$$bdP = acQ_2R_2.$$

En prenant les contenus, on obtient $c(bdP) = bd = c(acQ_2R_2) = ac$ (toujours modulo A^*). On a donc $P = \lambda Q_2R_2$ avec $\lambda \in A^*$. Comme P est irréductible dans $A[X]$, on a Q_2 ou R_2 dans $A[X]^* = A^*$, donc de degré 0. Par exemple $Q = \frac{a}{b}Q_2$ est de degré 0, donc $Q \in K^* = K[X]^*$, *i.e.* P est irréductible dans $K[X]$.

□

Proposition (Division euclidienne)**.** *Soit A un anneau et $P \in A[X]$, $P \neq 0$ de coefficient dominant inversible. Soit $F \in A[X]$. Alors il existe $Q, R \in A[X]$ tels que*

$$F = PQ + R$$

et $\deg(R) < \deg(P)$ *ou* $R = 0$.

Démonstration. On peut le faire par récurrence. Sinon, comme le coefficient dominant de P est inversible, on peut le supposer unitaire : $P = X^n + a_{n-1}X^{n-1} + \cdots + a_0$. Soient l'anneau quotient $B = A[X]/(P)$ et x l'image de X dans B. Il suffit de prouver que tout élément de B est combinaison linéaire à coefficients dans A de $1, x, \ldots, x^{n-1}$. En effet, on aurait alors $\overline{F} = \overline{R} + PA[X]$ où $\overline{R} \in B$, *i.e.* $R = 0$ ou $\deg(R) < n = \deg(P)$, et on pourrait alors écrire $F = R + PQ$ avec $R, Q \in A[X]$.

Par linéarité, il suffit même de montrer que x^i est combinaison linéaire à coefficients dans A de $1, x, \ldots, x^{n-1}$. Comme $\overline{P} = 0$ dans B, on a

$$x^n = -a_{n-1}x^{n-1} - \cdots - a_0.$$

On a donc le résultat par une récurrence immédiate. □

Remarque. On peut prouver l'unicité du couple (Q, R) : si (Q_1, R_1) et (Q_2, R_2) sont deux couples vérifiant les hypothèses, on a

$$(Q_2 - Q_1)P = (R_1 - R_2).$$

Comme le coefficient dominant de P est inversible dans A, si $Q_2 - Q_1 \neq 0$, on a $\deg((Q_2 - Q_1)P) \geqslant \deg(P)$. Or $\deg(R_1 - R_2) < \deg(P)$. Donc $Q_2 = Q_1$, puis $R_2 = R_1$.

Références :

- Perrin - *Cours d'algèbre* - Page 83.
- Gourdon - *Algèbre* - Page 91 (pour la définition des polynômes cyclotomiques et le fait qu'ils sont dans $\mathbb{Z}[X]$).

2.10 Méthode de Gauss-Seidel

Méthode (Méthodologie générale)**.** On considère $K = \mathbb{R}$ ou $\mathbb{C}$. Soient $A \in \mathrm{GL}_n(K)$ et $b \in K^n$. On cherche une approximation de $x \in K^n$ solution de $Ax = b$. Pour cela on pose $A = M - N$ où M est une matrice inversible "facile à inverser" (par exemple diagonale, triangulaire ou orthogonale). Pour x_0 donné, on définit alors la suite itérative

$$x_{k+1} = M^{-1}Nx_k + M^{-1}b.$$

Si la suite (x_k) converge, alors sa limite x doit satisfaire $x = M^{-1}Nx + M^{-1}b$, soit $Ax = b$. Il s'agit d'établir un critère sur les matrices M et N pour que la méthode converge. Par définition (on pose cette définition), l'algorithme converge si pour tout b et tout x_0, l'erreur $e_k = x_k - x$ vérifie $\|e_k\| \to 0$ quand $k \to \infty$ (pour une norme quelconque, elles sont toutes équivalentes). On a :

$$\begin{aligned} e_{k+1} &= x_{k+1} - x \\ &= (M^{-1}Nx_k + M^{-1}b) - (M^{-1}Nx + M^{-1}b) \\ &= M^{-1}Ne_k. \end{aligned}$$

En posant $G = M^{-1}N$, on obtient $e_k = G^k e_0$. Or $G^k e_0 \to 0$ pour tout e_0 est équivalent à $\rho(G) < 1$ où $\rho(G)$ désigne le rayon spectral de G.

Méthode (Gauss-Seidel)**.** On choisit pour M le triangle inférieur de A, et N le triangle supérieur strict. Le vecteur x_{k+1} est solution du système triangulaire inférieur $Mx_{k+1} = Nx_k + b$, d'où :

$$x_{k+1,i} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_{k+1,j} - \sum_{j=i+1}^{n} a_{ij} x_{k,j} \right),$$

pour tout $i \in [\![1, n]\!]$.

Théorème. *Si la matrice A est symétrique réelle définie positive, alors la méthode de Gauss-Seidel converge.*

Démonstration. Comme A est symétrique définie positive, ses termes diagonaux sont strictement positifs : on prend le vecteur X ayant un 1 en i-ème position et des 0 partout ailleurs, et après calcul, ${}^t\overline{X}SX = a_{ii} > 0$. Donc M est inversible et la méthode de Gauss-Seidel s'applique.

Soit $\lambda \in \mathbb{C}$ une valeur propre de $G = M^{-1}N$ de vecteur propre associé x, et on a $Nx = \lambda Mx$. On écrit $M = D + E$ où D est constituée par la diagonale de A et E le triangle inférieur strict de A, puis $N = -F$, de sorte que $A = D + E + F$. On note $(\cdot, \cdot)$ le produit scalaire hermitien canonique de $\mathbb{C}^n$. On a :

$$(x, Ax) = (x, Dx) + (x, Ex) + (x, Fx). \quad (1)$$

Comme $Nx = \lambda Mx$, on a $Fx = -\lambda(Dx + Ex)$, d'où

$$(x, Ax) = (1 - \lambda)\big((x, Dx) + (x, Ex)\big). \quad (2)$$

Or $(x, Ax) > 0$ car $x \neq 0$ et A définie positive, donc $\lambda \neq 1$. On prend ensuite la conjugaison complexe de cette égalité. Sachant que $\overline{(x, Ax)} = (x, Ax)$ (puisque c'est un réel), que $\overline{(x, Dx)} = (x, Dx)$ (c'est aussi un réel car D est symétrique réelle définie positive), et que $\overline{(x, Ex)} = (Ex, x) = {}^t\overline{Ex}x = {}^t\overline{x}Fx = (x, Fx)$ (car ${}^t\overline{E} = {}^tE = F$), on obtient

$$(x, Ax) = (1 - \overline{\lambda})\big((x, Dx) + (x, Fx)\big). \quad (3)$$

En faisant $\frac{1}{1-\lambda}(2) + \frac{1}{1-\overline{\lambda}}(3) - (1)$, on aboutit à

$$\left(\frac{1}{1-\lambda} + \frac{1}{1-\overline{\lambda}} - 1\right)(x, Ax) = (x, Dx),$$

soit

$$\frac{1 - |\lambda|^2}{|1-\lambda|^2}(x, Ax) = (x, Dx).$$

Or $(x, Ax) > 0$ et $(x, Dx) > 0$ (les matrices A et D sont définies positives), donc $|\lambda| < 1$. On en conclut que $\rho(G) < 1$ et que la méthode de Gauss-Seidel converge. □

Théorème. *Si la matrice A est à diagonale strictement dominante, alors la méthode de Gauss-Seidel converge.*

Démonstration. On note $A = (a_{ij})$ et on a par hypothèse $|a_{ii}| > \sum_{j\neq i} |a_{ij}|$, pour tout i. En particulier, A a tous ses termes diagonaux non nuls et on peut utiliser la méthode de Gauss-Seidel (car M sera bien inversible).

On écrit $A = M - N$ avec M le triangle inférieur de A, et $G = M^{-1}N$. Soit $\lambda \in \mathbb{C}$ une valeur propre de G de vecteur propre associé x. On a alors $Nx = \lambda Mx$, *i.e.*

$$-\sum_{j=i+1}^{n} a_{ij}x_j = \lambda \sum_{j=1}^{i} a_{ij}x_j$$

pour tout $i \in [\![1, n]\!]$. Soit i tel que $|x_i| = \max_k |x_k|$. On suppose $|\lambda| \geqslant 1$. Alors :

$$|\lambda||a_{ii}| = \frac{1}{|x_i|}\left|-\sum_{j=i+1}^{n} a_{ij}x_j - \lambda\sum_{j=1}^{i-1} a_{ij}x_j\right| \leqslant \sum_{j\neq i} |\lambda||a_{ij}|,$$

ce qui donne $|a_{ii}| \leqslant \sum_{j\neq i} |a_{ij}|$ et contredit le fait que A est à diagonale strictement dominante. Donc $\rho(G) < 1$ et la méthode de Gauss-Seidel converge. □

Sur le nombre d'itérations à effectuer :

On définit le vecteur résidu à l'itération k par $r_k = b - Ax_k$. Le vecteur x est solution de $Ax = b$ pour lequel le résidu $r = b - Ax$ est nul. On sera donc d'autant plus proche de la solution que le résidu sera petit. On poursuit les itérations jusqu'à ce que

$$\frac{\|r_k\|}{\|A\|} \leqslant \varepsilon.$$

Or $\|r_k\| = \|b - Ax_k\| = \|Ax - Ax_k\| \leqslant \|A\|\|e_k\|$. On poursuit donc jusqu'à ce que $\|e_k\| \leqslant \varepsilon$. Si on a $e_k = B^k e_0$ avec $\|B\| < 1$, alors $\|e_k\| \leqslant \|B\|^k \|e_0\|$. D'autre part

$$\|e_0\| \leqslant \|x_0 - x_1\| + \|e_1\| \leqslant \|x_0 - x_1\| + \|B\|\|e_0\|,$$

d'où

$$\|e_0\| \leqslant \frac{1}{1 - \|B\|}\|x_0 - x_1\|.$$

Il vient alors

$$\|e_k\| \leqslant \frac{\|B\|^k}{1 - \|B\|}\|x_1 - x_0\|$$

et il est facile de trouver un nombre d'itérations k de sorte que $\|e_k\| \leqslant \varepsilon$ pour le $\varepsilon > 0$ que l'on veut.

Sur le nombre d'opérations à effectuer :

Pour résoudre un système $Ax = b$ où A est inversible et triangulaire inférieure, on utilise naturellement une méthode dite de descente :

$$\begin{cases} x_1 = \frac{b_1}{a_{11}} \\ x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j \right), \ i = 2, \ldots, n. \end{cases}$$

Cet algorithme effectue $\frac{n(n-1)}{2}$ additions, $\frac{n(n-1)}{2}$ multiplications et n divisions, soit un nombre d'opérations global de l'ordre de n^2.

Comptons le nombre d'opérations pour passer de l'étape k à l'étape $k+1$ dans la méthode de Gauss-Seidel. On doit effectuer $\sum_{i=1}^{n} n = n^2$ additions, $\sum_{i=1}^{n}(n-1) = n(n-1)$ multiplications, et n divisions. Il faut ensuite calculer le vecteur résidu pour savoir si on s'arrête où non : $n + n(n-1) = n^2$ additions (pour ajouter deux vecteurs, on doit faire n additions, et pour multiplier une matrice par un vecteur, $n(n-1)$ additions), n^2 multiplications (dans la multiplication d'une matrice par un vecteur). Enfin, il faudra réaliser $n-1$ additions, n multiplications et une extraction de racine carrée (extraction dont on peut se passer si on raisonne avec la norme au carrée) pour calculer la norme du résidu. On obtient de l'ordre de $4n^2$ opérations au total.

On peut en réalité faire un peu mieux. Ci-dessus, on a calculé x_{k+1} en se servant de la relation $Mx_{k+1} = Nx_k + b$. Mais $N = M - A$ et la relation devient $x_{k+1} =$

$x_k + M^{-1}(b - Ax_k) = x_k + M^{-1}r_k$. On commence donc par calculer le résidu r_k, puis pour calculer $y = M^{-1}r_k$, on résout $My = r_k$ par un algorithme de descente, donc de l'ordre de n^2 opérations. Il ne reste plus qu'à faire les n additions de $x_k + y$ pour obtenir x_{k+1}. On obtient de l'ordre de $3n^2$ opérations au total.

Dans tous les cas, si le nombre d'itérations à effectuer reste petit devant n, la méthode de Gauss-Seidel est plus efficace que les méthodes directes qui nécessitent un $O(n^3)$ opérations (voir document "Décomposition QR").

La méthode de Gauss-Seidel est à mettre en comparaison avec celle de Jacobi :

Méthode (Jacobi)**.** On choisit pour M la matrice diagonale formée par la diagonale de A. Le vecteur x_{k+1} est solution du système triangulaire inférieur $Mx_{k+1} = Nx_k + b$, d'où :

$$x_{k+1,i} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_{k,j} \right),$$

pour tout $i \in [\![1, n]\!]$.

Comme M est diagonale, la résolution de $My = r_k$ est encore plus simple que pour la méthode de Gauss-Seidel : on n'a besoin de faire que n divisions. Cette méthode nécessite donc un nombre d'opérations de l'ordre de $2n^2$ au total.

Théorème. *Si A est à diagonale strictement dominante, alors la méthode de Jacobi converge.*

Remarque. La méthode de Jacobi nécessite de garder en mémoire le vecteur x_k entier pour le calcul des composantes de x_{k+1}, contrairement à la méthode de Gauss-Seidel. Pour Jacobi, on a donc besoin de $2n$ cases mémoires, contre n pour Gauss-Seidel. Cette remarque concerne a priori l'utilisation de $Mx_{k+1} = Nx_k + b$, et non la reformulation en $x_{k+1} = x_k + M^{-1}r_k$.

Remarque. De manière générale, quand les méthodes de Jacobi et Gauss-Seidel sont convergentes, il vaut mieux choisir celle de Gauss-Seidel. Mais il se peut que la méthode de Gauss-Seidel diverge alors que celle de Jacobi converge. Exemple :

$$A = \begin{pmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}.$$

Pour la méthode de Jacobi, toutes les valeurs propres sont nulles, alors que pour celle de Gauss-Seidel, on a 0, 0.35 et $5.64 > 1$.

Théorème (Critères de convergence)**.** *Soit $A \in \mathcal{M}_n(K)$. Les quatre propositions suivantes sont équivalentes :*

(i) Pour une norme subordonnée quelconque, on a $\lim_{k\to\infty} \|A^k\| = 0$.

(ii) Pour tout vecteur $v \in K^n$, on a $\lim_{k\to\infty} A^k v = 0$.

(iii) Le rayon spectral de A vérifie $\rho(A) < 1$.

(iv) Il existe une norme matricielle subordonnée (dont le choix dépend de A) telle que $\|A\| < 1$.

Démonstration. Noter qu'étant en dimension finie, toutes les normes sont équivalentes.

$(i) \Rightarrow (ii)$: résulte de $\|A^k x\| \leqslant \|A^k\|\|x\|$.

$(ii) \Rightarrow (iii)$: si $\rho(A) \geqslant 1$, soient $|\lambda| = \rho(A)$ et x un vecteur propre associé à λ, *i.e.* tel que $Ax = \lambda x$. Alors la suite de terme général $A^k x = \lambda^k x$ ne converge pas vers 0.

$(iii) \Rightarrow (iv)$: c'est l'implication la plus délicate et celle qui nous intéresse le plus. On va se servir du résultat suivant qu'on démontre dans la suite : pour tout $\varepsilon > 0$, il existe une norme subordonnée $\|\cdot\|_{A,\varepsilon}$ telle que $\|A\|_{A,\varepsilon} \leqslant \rho(A) + \varepsilon$. Dans le cas où $\rho(A) < 1$, on peut trouver $\varepsilon > 0$ tel que $\rho(A) + \varepsilon < 1$, ce qui donne le résultat.

$(iv) \Rightarrow (i)$: conséquence de l'inégalité $\|A^k\| \leqslant \|A\|^k$. □

Proposition. *Soit $A \in \mathcal{M}_n(K)$. On note $\|A\|_\infty = \sup_{x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty}$ la norme subordonnée de A à la norme $\|\cdot\|_\infty$ sur K^n. Alors*

$$\|A\|_\infty = \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}|.$$

Démonstration. On majore $\|Ax\|_\infty$ pour $x \neq 0$:

$$\|Ax\|_\infty = \max_{1\leqslant i\leqslant n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \leqslant \|x\|_\infty \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}|.$$

Donc $\|A\|_\infty \leqslant \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}|$.

Soit k un indice tel que $\sum_{j=1}^{n} |a_{kj}| = \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}|$. Soit $x \in K^n$ tel que $x_j = \frac{\overline{a_{kj}}}{|a_{kj}|}$ si $a_{kj} \neq 0$, $x_j = 1$ sinon. Ainsi $\|x\|_\infty = 1$ et

$$(Ax)_k = \sum_{j=1}^{n} a_{kj} x_j = \sum_{j=1}^{n} |a_{kj}|.$$

Donc

$$\|A\|_\infty \geqslant \frac{\|Ax\|_\infty}{\|x\|_\infty} = \|Ax\|_\infty \geqslant (Ax)_k \geqslant \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}|,$$

d'où le résultat. □

Théorème. *Soient $A \in \mathcal{M}_n(K)$ et $\varepsilon > 0$. Alors il existe une norme matricielle subordonnée $\|\cdot\|_{A,\varepsilon}$ telle que*

$$\|A\|_{A,\varepsilon} \leqslant \rho(A) + \varepsilon.$$

Démonstration. On note $\lambda_1, \dots, \lambda_n$ les valeurs propres complexes de A. On peut trigonaliser A dans $\mathbb{C}$: il existe $U \in \mathrm{GL}_n(\mathbb{C})$ telle que $U^{-1}AU = D + C$ avec $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ et

$$C = \begin{pmatrix} 0 & c_{1,2} & \dots & c_{1,n} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & c_{n-1,n} \\ 0 & \dots & \dots & 0 \end{pmatrix}.$$

Pour $\delta > 0$, on introduit la matrice diagonale $D_\delta = \mathrm{Diag}(1, \delta, \delta^2, \dots, \delta^{n-1})$ de sorte que

$$D_\delta^{-1}(D+C)D_\delta = D + D_\delta^{-1}CD_\delta = \begin{pmatrix} 0 & \delta c_{1,2} & \dots & \delta^{n-1}c_{1,n} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \delta c_{n-1,n} \\ 0 & \dots & \dots & 0 \end{pmatrix}.$$

Ainsi pour $\delta > 0$ assez petit, on a

$$\|D_\delta^{-1}CD_\delta\|_\infty = \max_{1 \leqslant i \leqslant n} \sum_{j=i+1}^{n} \delta^{j-i}|c_{ij}| \leqslant \varepsilon.$$

On obtient alors

$$\|(UD_\delta)^{-1}A(UD_\delta)\|_\infty \leqslant \|D\|_\infty + \|D_\delta^{-1}CD_\delta\|_\infty \leqslant \rho(A) + \varepsilon.$$

Soit la norme $\|B\|_{A,\varepsilon} = \|(UD_\delta)^{-1}B(UD_\delta)\|_\infty$ pour $B \in \mathrm{GL}_n(K)$. Cette norme vérifie bien $\|A\|_{A,\varepsilon} \leqslant \rho(A) + \varepsilon$. Il reste à montrer que c'est une norme subordonnée. On a :

$$\|B\|_{A,\varepsilon} = \sup_{x \neq 0} \frac{\|(UD_\delta)^{-1}B(UD_\delta)x\|_\infty}{\|x\|_\infty} = \sup_{y \neq 0} \frac{\|(UD_\delta)^{-1}By\|_\infty}{\|(UD_\delta)^{-1}y\|_\infty} = \sup_{y \neq 0} \frac{\|By\|_{A,\varepsilon}}{\|y\|_{A,\varepsilon}}$$

où $\|y\|_{A,\varepsilon} = \|(UD_\delta)^{-1}y\|_\infty$. □

Références :

- Rombaldi - *Analyse matricielle, cours et exercices résolus.*
- Filbet - *Analyse numérique, algorithme et étude mathématique* - Pages 38 à 47 (pour une autre méthode et des compléments).

2.11 Quelques propriétés des homographies

Cette section propose 4 développements possibles. Dans l'ordre :

1. Les homographies conservent les pseudo-cercles.
2. Les homographies conservent les angles orientés.
3. Existence et unicité d'une homographie envoyant trois points distincts z_1, z_2, z_3 sur trois points distincts z_1', z_2', z_3'.
4. Les homographies conservent le birapport et caractérisation de la cocyclicité (ou alignement) de quatre points avec le birapport.

Dans toute cette section, on se place dans le plan $\mathbb{R}^2$ que l'on identifie à $\mathbb{C}$.

On note H le groupe des homographies constitué des transformations de la forme $z \mapsto \frac{az+b}{cz+d}$ avec $ad - bc \neq 0$.

Enfin, on note $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.

Proposition. *Le groupe H est engendré par les similitudes directes $z \mapsto az+b$ (avec $a \neq 0$) et l'application $z \mapsto \frac{1}{z}$.*

Démonstration. Soit $h : z \mapsto \frac{az+b}{cz+d}$. Si $c = 0$, h est une similitude directe. Si $c \neq 0$, on décompose h en éléments simples :

$$h(z) = \frac{a}{c} + \frac{bc - ad}{c} \frac{1}{cz + d}.$$

On note $s_1(z) = cz + d$, $f(z) = \frac{1}{z}$ et $s_2(z) = \frac{a}{c} + \frac{bc-ad}{c} z$. Alors $h = s_2 \circ f \circ s_1$. □

Définition. On appelle *pseudo-cercle* de $\mathbb{C}$ tout cercle ou droite de $\mathbb{C}$, les droites étant vues comme des cercles de rayon infini. On appelle *pseudo-cercle* de $\hat{\mathbb{C}}$ tout cercle de $\mathbb{C}$ ou droite de $\mathbb{C}$ à laquelle on ajoute le point $\{\infty\}$.

Proposition. *Les pseudo-cercles de $\mathbb{C}$ sont exactement donnés par les équations de la forme $az\overline{z} + bz + \overline{b}\overline{z} + c = 0$ avec $a, c \in \mathbb{R}$ et $b \in \mathbb{C}$ tels que $|b|^2 > ac$.*

Démonstration. Une droite a pour équation $\alpha x + \beta y + \gamma = 0$ avec $(\alpha, \beta) \neq (0, 0)$. On pose $a = 0$, $b = \frac{\alpha - i\beta}{2}$ et $c = \gamma$. On obtient bien l'équation voulue, avec de plus $|b|^2 > ac = 0$.

Un cercle a pour équation $x^2+y^2+\alpha x+\beta y = R^2$ avec $\alpha, \beta, R \in \mathbb{R}$ et $(\alpha, \beta, R) \neq (0,0,0)$. On pose $z = x + iy$ et l'équation devient :

$$|z|^2 + \alpha\frac{z+\overline{z}}{2} + \beta\frac{z-\overline{z}}{2i} - R^2 = az\overline{z} + bz + \overline{b}\overline{z} + c = 0,$$

en ayant posé $a = 1$, $b = \frac{\alpha - i\beta}{2}$ et $c = -R^2$. De plus, on a $|b|^2 > ac = -R^2$ dès lors que $(b, R) \neq (0,0)$. Mais si $b = R = 0$, alors $\alpha = \beta = 0$ et l'équation devient $x^2 + y^2 = 0$ qui ne représente pas un cercle.

Réciproquement, considérons l'équation $az\overline{z} + bz + \overline{b}\overline{z} + c = 0$ avec $a, c \in \mathbb{R}$ et $|b|^2 > ac$. On écrit $z = x + iy$ et $b = \alpha + i\beta$ et on obtient

$$ax^2 + ay^2 + 2\alpha x - 2\beta y + c = 0.$$

Si $a = 0$, on a $(\alpha, \beta) \neq (0,0)$ car sinon $b = 0$ et on aurait une contradiction avec l'hypothèse $|b|^2 > ac$. Ainsi l'équation $2\alpha x - 2\beta y + c = 0$ est l'équation d'une droite du plan.

Si $a \neq 0$, l'équation peut se réécrire

$$\left(x + \frac{\alpha}{a}\right)^2 + \left(y - \frac{\beta}{a}\right)^2 = -\frac{c}{a} + \frac{\alpha^2}{a^2} + \frac{\beta^2}{a^2} = \frac{-ac + |b|^2}{a^2}.$$

On retrouve bien l'équation d'un cercle car $|b|^2 - ac > 0$. □

Théorème. *Les homographies conservent les pseudo-cercles de* $\hat{\mathbb{C}}$.

Démonstration. Grâce aux deux premières propositions, il suffit de montrer que les équations de la forme $az\overline{z} + bz + \overline{b}\overline{z} + c = 0$ sont préservées par les similitudes et par l'application $z \mapsto \frac{1}{z}$.

Les similitudes sont les composées de translations, homothéties et rotations, elles transforment donc les cercles en cercles et les droites en droites.

L'ensemble des $z \in \mathbb{C}^*$ tels que $az\overline{z} + bz + \overline{b}\overline{z} + c = 0$ est transformé par l'application $z \mapsto \frac{1}{z}$ en l'ensemble des $z' \in \mathbb{C}^*$ tels que

$$\frac{a}{z'\overline{z'}} + \frac{b}{z'} + \frac{\overline{b}}{\overline{z'}} + c = 0,$$

soit encore $a + b\overline{z'} + \overline{b}z' + cz'\overline{z'} = 0$ qui est un pseudo-cercle car $|b|^2 > ac$. De plus, si $z = 0$ appartient au pseudo-cercle d'équation $az\overline{z} + bz + \overline{b}\overline{z} + c = 0$, alors $c = 0$ et l'ensemble obtenu après transformation est une droite à laquelle on ajoute le point $\{\infty\}$. □

Proposition. *Les homographies conservent les angles orientés.*

Démonstration. Si on admet que les applications holomorphes conservent les angles, le résultat est immédiat.

Sinon on se sert du fait que les homographies sont engendrées par les similitudes directes qui préservent les angles orientés, et l'application $z \mapsto \frac{1}{z}$. Cette dernière est la composée de la réflexion $z \mapsto \overline{z}$ et de l'inversion $z \mapsto \frac{1}{\overline{z}}$ qui renversent toutes les deux les angles orientés, d'où le résultat. □

Définition. Soit A un point du plan et k un nombre réel non nul. On appelle *inversion* de pôle A et de puissance k la transformation $I_{A,k} : M \mapsto M'$ où M' est le point de la droite AM vérifiant l'égalité $\overline{AM}.\overline{AM'} = k$.

On peut exprimer la même chose en d'autres termes :

$$\overrightarrow{OM'} = \frac{k}{OM^2}\overrightarrow{OM}.$$

Si z_A, z, z' sont les affixes de A, M, M', c'est équivalent à

$$z' = I_{A,k}(z) = z_A + \frac{k}{|z - z_A|^2}(z - z_A) = z_A + \frac{k}{\overline{z} - \overline{z_A}} = \frac{z_A\overline{z} - |z_A|^2 + k}{\overline{z} - \overline{z_A}}.$$

Remarque. Il est clair que l'inversion $I_{A,k}$ est involutive ($I_{A,k} \circ I_{A,k} = \mathrm{Id}$. Si $k > 0$, elle admet des points fixes qui sont les points du cercle de centre O et de rayon $\sqrt{k}$, appelé *cercle d'inversion*. Une inversion échange intérieur et extérieur de son cercle d'inversion.

Proposition. *Les réflexions renversent les angles orientés.*

Démonstration. Soient D une droite vectorielle et u, v deux vecteurs unitaires. On considère la réflexion s_D par rapport à la droite D (la symétrie orthogonale par rapport à la droite D) et on veut montrer que $(s_D(u), s_D(v)) = (v, u)$.

Considérons la droite D' engendrée par $u + v$. Alors si $s_{D'}$ est la réflexion par rapport à D', on a $s_{D'}(u) = v$ et $s_{D'}(v) = u$. D'autre part, $s_D \circ s_{D'}$ est une isométrie (composée d'isométries), positive (la composée de deux isométries négatives et une isométrie positive), donc une rotation, et donc on a l'égalité des angles :

$$(v, u) = \big(s_D \circ s_{D'}(v), s_D \circ s_{D'}(u)\big).$$

Mais ce dernier angle vaut $(s_D(u), s_D(v))$ car $s_{D'}(u) = v$ et $s_{D'}(v) = u$. □

Proposition. *Les inversions renversent les angles orientés.*

Démonstration. On calcule la différentielle de l'inversion $\varphi = I_{O,k}$:

$$\overrightarrow{O\varphi(M+u)} = \frac{k}{\|OM + u\|^2}(\overrightarrow{OM} + u)$$

$$\begin{aligned} &= \frac{k}{OM^2}\left(1+2\frac{\overrightarrow{OM}.u}{OM^2}+\frac{\|u\|^2}{OM^2}\right)^{-1}(\overrightarrow{OM}+u) \\ &= \frac{k}{OM^2}\overrightarrow{OM}+\frac{k}{OM^2}\left(u-2\frac{\overrightarrow{OM}.u}{OM^2}\overrightarrow{OM}\right)+o\left(\|OM\|\right). \end{aligned}$$

On en déduit que

$$D\varphi(M).u = \frac{k}{OM^2}\left(u-2\frac{\overrightarrow{OM}.u}{OM^2}\overrightarrow{OM}\right).$$

On se rappelle maintenant (c'est facile à démontrer) que si H est un hyperplan d'un espace vectoriel euclidien E avec x_0 vecteur non nul de $H^\perp$, alors la réflexion s_H par rapport à l'hyperplan H est

$$s_H(x) = x - 2\frac{x.x_0}{\|x_0\|^2}x_0.$$

Donc $D\varphi(M)$ est la composée de la réflexion par rapport à la droite vectorielle orthogonale à OM et de l'homothétie de rapport $\frac{k}{OM^2}$. En particulier, $D\varphi(M)$ transforme les angles orientés en leurs opposés car on a montré dans la proposition précédente que les réflexions renversent les angles orientés.

A présent, soient deux courbes planes se coupant en un point A. L'angle de ces deux courbes est l'angle de leurs tangentes en A. Considérons les images de ces courbes par une inversion dont le pôle n'est pas en A. Ce sont deux nouvelles courbes qui se coupent en l'image A' de A. Les tangentes en A' à ces deux courbes sont les images des tangentes en A aux deux courbes d'origine par la différentielle de l'inversion utilisée. En particulier, l'angle des courbes images est l'opposé de l'angle des courbes de départ. □

Proposition. *L'action $H\times\hat{\mathbb{C}}\to\hat{\mathbb{C}}$ qui à (h,z) associe $h(z)$ est fidèle et simplement 3-transitive, i.e.*

(i) Seule l'identité fixe tous les points de $\hat{\mathbb{C}}$ (action fidèle).

(ii) Pour tout $z_1,z_2,z_3\in\hat{\mathbb{C}}$ deux à deux distincts, et tout $z_1',z_2',z_3'\in\hat{\mathbb{C}}$ deux à deux distincts, il existe une unique homographie h telle que $h(z_i)=z_i'$ pour $i\in\{1,2,3\}$ (action simplement 3-transitive).

Démonstration. La fidélité est facile : si $\frac{az+b}{cz+d}=z$ pour tout $z\in\hat{\mathbb{C}}$, on obtient $cz^2+(d-a)z-b=0$ pour tout $z\in\hat{\mathbb{C}}$, et donc ce polynôme du second degré en z est nul, *i.e.* $c=b=0$ et $d=a$, d'où $z\mapsto\frac{az+b}{cz+d}$ est l'identité.

Pour la simple 3-transitivité, il suffit de montrer la propriété pour $(z_1', z_2', z_3') = (0, 1, \infty)$, car alors il existera une unique homographie h_1 envoyant (z_1, z_2, z_3) sur $(0, 1, \infty)$ et une unique homographie h_2 envoyant (z_1', z_2', z_3') sur $(0, 1, \infty)$, donc $h_2^{-1} \circ h_1$ est une homographie envoyant (z_1, z_2, z_3) sur (z_1', z_2', z_3'). Elle est unique car si h_3 en est une autre, $h_3 \neq h_2^{-1} \circ h_1$, alors $h_2 \circ h_3$ est une homographie différente de h_1 et qui envoie (z_1, z_2, z_3) sur $(0, 1, \infty)$, ce qui est absurde.

Traitons les différents cas possibles :

1. Si $z_1, z_2, z_3 \in \mathbb{C}$:

$$\frac{az_1 + b}{cz_1 + d} = 0 \iff az_1 + b = 0 \iff b = -az_1.$$

$$\frac{az_3 + b}{cz_3 + d} = \infty \iff cz_3 + d = 0 \iff d = -cz_3.$$

$$\frac{az_2 + b}{cz_2 + d} = 1 \iff az_2 + b = cz_2 + d.$$

Si $z_2 = 0$,

$$a = c\frac{z_3}{z_1}, b = -cz_3, d = -cz_3.$$

Si $z_2 \neq 0$, $az_2 + b = cz_2 + d \Longrightarrow az_2 - az_1 = cz_2 - cz_3$, donc

$$a = c\frac{z_2 - z_3}{z_2 - z_1}, b = -c\frac{z_2 - z_3}{z_2 - z_1}z_1, d = -cz_3.$$

Dans tous les cas, a, b, d s'expriment en fonction de c, z_1, z_2, z_3 et dépendent multiplicativement de c, donc a, b, c, d définiront la même homographie quelle que soit la valeur de c (multiplier au numérateur et au dénominateur par un même nombre ne change pas l'homographie). Il faut encore vérifier que $c \neq 0$: si $c = 0$, alors comme $h(z_3) = \infty$, cela implique $z_3 = \infty \notin \mathbb{C}$, et on avait supposé $z_3 \in \mathbb{C}$. Enfin,

$$ad - bc = a(-cz_3) - (-az_1)c = ac(z_1 - z_3) \neq 0.$$

2. Si $z_1 = \infty$:

$$\frac{az_1 + b}{cz_1 + d} = 0 \iff a = 0.$$

Il vient ensuite

$$\frac{b}{cz_3 + d} = \infty \iff d = -cz_3 \ ; \ \frac{b}{cz_2 + d} = 1 \iff b = c(z_2 - z_3).$$

3. Si $z_2 = \infty$:

$$\frac{az_2 + b}{cz_2 + d} = 1 \iff a = c.$$

Il vient ensuite

$$\frac{az_1 + b}{cz_1 + d} = 0 \iff b = -az_1 = -cz_1 \ ; \ \frac{az_3 + b}{cz_3 + d} = \infty \iff d = -cz_3.$$

4. Si $z_3 = \infty$:

$$\frac{az_3+b}{cz_3+d} = \infty \iff \frac{a}{c} = \infty \iff c = 0.$$

Il vient ensuite

$$\frac{az_1+b}{d} = 0 \iff b = -az_1 \ ; \ \frac{az_2+b}{d} = 1 \iff d = a(z_2 - z_1).$$

□

Définition. Le *birapport* de quatre points $w_1, w_2, w_3, w_4 \in \hat{\mathbb{C}}$ avec w_1, w_2, w_3 deux à deux distincts est l'image de w_4 par l'unique homographie h qui envoie (w_1, w_2, w_3) sur $(\infty, 0, 1)$. On le note $[w_1, w_2, w_3, w_4] = h(w_4)$.

Proposition. *Soient $w_1, w_2, w_3, w_4 \in \hat{\mathbb{C}}$ avec w_1, w_2, w_3 deux à deux distincts. Alors*

$$[w_1, w_2, w_3, w_4] = \frac{w_4 - w_2}{w_4 - w_1} \div \frac{w_3 - w_2}{w_3 - w_1}.$$

Démonstration. Soit h l'unique homographie qui envoie (w_1, w_2, w_3) sur $(\infty, 0, 1)$. On vérifie facilement que

$$h(z) = \frac{z - w_2}{z - w_1} \div \frac{w_3 - w_2}{w_3 - w_1},$$

d'où l'expression de $[w_1, w_2, w_3, w_4] = h(w_4)$. □

Proposition. *Soient $w_1, \ldots, w_4, w_1', \ldots, w_4' \in \hat{\mathbb{C}}$ avec w_1, w_2, w_3 deux à deux distincts et w_1', w_2', w_3' deux à deux distincts. Alors*

$$[w_1, w_2, w_3, w_4] = [w_1', w_2', w_3', w_4']$$

si et seulement s'il existe une homographie h telle que $h(w_i) = w_i'$ pour tout $i \in \{1, 2, 3, 4\}$.

En particulier, les homographies conservent le birapport : si $h \in H$,

$$[w_1, w_2, w_3, w_4] = [h(w_1), h(w_2), h(w_3), h(w_4)].$$

Démonstration. On note k l'unique homographie qui envoie (w_1, w_2, w_3) sur $(\infty, 0, 1)$ et k' l'unique homographie qui envoie (w_1', w_2', w_3') sur $(\infty, 0, 1)$. Alors

$$[w_1, w_2, w_3, w_4] = k(w_4) \ \text{ et } \ [w_1', w_2', w_3', w_4'] = k'(w_4').$$

Supposons $[w_1, w_2, w_3, w_4] = [w_1', w_2', w_3', w_4']$, *i.e.* $k(w_4) = k'(w_4')$. Alors $h = k'^{-1} \circ k$ est une homographie telle que $h(w_i) = w_i'$ pour tout $i \in \{1, 2, 3\}$ et

$$h(w_4) = k'^{-1}(k(w_4)) = k'^{-1}(k'(w_4')) = w_4'.$$

Réciproquement, supposons qu'il existe une homographie h telle que $h(w_i) = w'_i$ pour tout $i \in \{1, 2, 3, 4\}$. Alors h est l'unique homographie qui envoie (w_1, w_2, w_3) sur (w'_1, w'_2, w'_3). Comme $k'^{-1} \circ k$ effectue ces trois mêmes transformations, par unicité, on a donc $h = k'^{-1} \circ k$. Donc $w'_4 = h(w_4) = k'^{-1}(k(w_4))$, d'où $k'(w'_4) = k(w_4)$, soit encore

$$[w_1, w_2, w_3, w_4] = [w'_1, w'_2, w'_3, w'_4].$$

□

Proposition. *Quatre points $w_1, w_2, w_3, w_4 \in \hat{\mathbb{C}}$ distincts sont cocycliques ou alignés si et seulement si $[w_1, w_2, w_3, w_4] \in \mathbb{R}$.*

Démonstration. Rappelons que les homographies conservent les pseudo-cercles de $\hat{\mathbb{C}}$.

Soit h l'unique homographie envoyant (w_1, w_2, w_3) sur $(\infty, 0, 1)$. Si w_4 appartient au pseudo-cercle contenant w_1, w_2, w_3 (une droite s'ils sont alignés, le cercle circonscrit au triangle $w_1w_2w_3$ sinon), alors $h(w_4)$ appartient au pseudo-cercle contenant $\infty, 0$ et 1, *i.e.* à la droite réelle. Par conséquent,

$$[w_1, w_2, w_3, w_4] = h(w_4) \in \mathbb{R}.$$

Réciproquement, supposons que $[w_1, w_2, w_3, w_4] \in \mathbb{R}$. Alors

$$h(w_4) = [w_1, w_2, w_3, w_4] \in \mathbb{R},$$

et par conséquent, $h(w_4)$ appartient à la droite réelle, qui est également le pseudo-cercle contenant $\infty, 0$ et 1. Comme h^{-1} est encore une homographie, on en déduit que w_4 appartient au pseudo-cercle contenant w_1, w_2 et w_3, *i.e.* $w_1, \ldots, w_4$ sont cocycliques ou alignés. □

Références :

- Issu d'un document ayant pour sources :
  - Samuel - *Géométrie projective.*
  - Cartan - *Fonctions d'une ou plusieurs variables complexes.*
  - Berger.
- Audin - *Géométrie* - Page 78 (les réflexions renversent les angles orientés), page 95 (les inversions renversent les angles orientés).

2.12 Réduction des isométries d'un espace euclidien

Théorème. *Soit f une isométrie d'un espace euclidien E. Alors il existe une base orthonormée de E dans laquelle la matrice de f s'écrit :*

$$\begin{pmatrix} D_1 & \dots & \dots & \dots & 0 \\ \vdots & D_{-1} & & & \vdots \\ \vdots & & C_1 & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \dots & \dots & \dots & C_p \end{pmatrix}$$

où D_1 est une matrice diagonale de 1 (éventuellement de taille nulle), D_{-1} une matrice diagonale de -1 (aussi éventuellement de taille nulle), $p \geqslant 0$ (éventuellement nul) et les C_i de la forme $\begin{pmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{pmatrix}$.

Démonstration. On va commencer par montrer que f possède une droite ou un plan stable. Pour cela, on va se servir de l'endomorphisme $f + f^{-1}$. Comme f est une isométrie, $f^* = f^{-1}$. On a donc :

$$(f + f^{-1})^* = f^* + (f^{-1})^* = f^{-1} + f,$$

i.e. $f + f^{-1}$ est symétrique. Par conséquent, $f + f^{-1}$ est diagonalisable dans une base orthonormée.

Soit λ une valeur propre réelle de $f + f^{-1}$. Alors il existe un vecteur $x \neq 0$ tel que $(f + f^{-1})(x) = \lambda x$. En composant par f, il vient $f^2(x) - \lambda f(x) + x = 0$. Ainsi le sous-espace $\mathrm{Vect}(x, f(x))$ est stable par f et est soit un plan (si x et $f(x)$ ne sont pas colinéaires), soit une droite (si x et $f(x)$ sont colinéaires).

Raisonnons par récurrence pour montrer que la matrice de f a la forme demandée dans une certaine base. D'après ce qui précède, f admet une droite ou un plan stable, qu'on note V.

Si V est une droite, soit e_1 un vecteur unitaire engendrant V. Alors soit $f(e_1) = e_1$, soit $f(e_1) = -e_1$. En effet, $f(e_1) = \lambda e_1$ et $\|f(e_1)\| = 1 = |\lambda|$.

Si V est un plan, on note (e_1, e_2) une base orthonormée de V. Comme f est une isométrie, $f_{|V}$ aussi (elle conserve toujours le produit scalaire). Démontrons que si $f_{|V}$ est une isométrie directe, il existe θ_1 tel que

$$\mathrm{Mat}_{(e_1,e_2)}(f_{|V}) = \begin{pmatrix} \cos(\theta_1) & -\sin(\theta_1) \\ \sin(\theta_1) & \cos(\theta_1) \end{pmatrix}.$$

Comme $\|f(e_1)\| = 1$, on peut écrire $f(e_1) = \cos(\theta_1)e_1 + \sin(\theta_1)e_2$. Ensuite il existe $a, b \in \mathbb{R}$ tels que $f(e_2) = ae_1 + be_2$. On exploite le fait que $\big(f(e_1)|f(e_2)\big) = 0$: $a\cos(\theta_1)+b\sin(\theta_1) = 0$. On a soit $\cos(\theta_1) \neq 0$, soit $\sin(\theta_1) \neq 0$ (car $\|f(e_1)\| = 1 \neq 0$). Supposons que $\sin(\theta_1) \neq 0$. Alors

$$b = \frac{-a\cos(\theta_1)}{\sin(\theta_1)}.$$

On a d'autre part que $\|f(e_2)\|^2 = a^2 + b^2 = 1$, donc $a^2 + a^2\frac{\cos^2(\theta_1)}{\sin^2(\theta_1)} = 1$, et après multiplication par $\sin^2(\theta_1)$, il vient $a^2 = \sin^2(\theta_1)$. Donc il existe $\varepsilon \in \{-1, 1\}$ tel que $a = \varepsilon\sin(\theta_1)$. Puis $b = -\varepsilon\cos(\theta_1)$. Comme $f_{|V}$ est directe, on trouve $\varepsilon = -1$.

Si $f_{|V}$ est une isométrie indirecte, on a

$$\mathrm{Mat}_{(e_1,e_2)}(f_{|V}) = \begin{pmatrix} \cos(\theta_1) & \sin(\theta_1) \\ \sin(\theta_1) & -\cos(\theta_1) \end{pmatrix},$$

qui est une matrice symétrique réelle. On obtient donc une valeur propre réelle pour f et donc f admet une droite stable : on est ramené au cas où V est une droite.

On écrit maintenant $E = V \oplus V^\perp$. Le sous-espace $V^\perp$ est stable par f : en effet, soient $v \in V$ et $w \in V^\perp$. Comme $f_{|V}$ est une isométrie de V, elle est bijective, donc il existe $v' \in V$ tel que $v = f(v')$. Ainsi,

$$(f(w), v) = (f(w), f(v')) = (w, v') = 0,$$

ce qui prouve que $f(w) \in V^\perp$. Comme $f_{|V^\perp}$ est une isométrie de $V^\perp$, on peut appliquer l'hypothèse de récurrence. Enfin, quitte à réordonner les vecteurs de la base trouvée, on obtient bien la matrice de f de la forme demandée. □

Théorème. *Soient E un espace euclidien et u un endomorphisme de E. Alors il existe un unique endomorphisme v tel que pour tout $x, y \in E$, $(u(x), y) = (x, v(y))$. On appelle v l'adjoint de u et on le note u^*. De plus, si $\mathcal{B}$ est une base orthonormée de E, on a $\mathrm{Mat}_{\mathcal{B}}(u^*) = {}^t\mathrm{Mat}_{\mathcal{B}}(u)$.*

Démonstration. On note $A = \mathrm{Mat}_{\mathcal{B}}(u)$. Alors $(u(x), y) = {}^t(AX)Y = {}^tX{}^tAY$. On note v l'endomorphisme associé à la matrice ${}^tA$ dans la base $\mathcal{B}$. L'unicité est facile à montrer. □

Proposition. *Si u est une isométrie, alors u est bijective et $u^* = u^{-1}$.*

Démonstration. Si u est une isométrie, u conserve la norme, et on prouve alors facilement que $\mathrm{Ker}(u) = \{0\}$. Donc u est bijective. Ensuite :

$$(u(x), y) = \big(u(x), u(u^{-1}(y))\big) = (x, u^{-1}(y))$$

car u conserve le produit scalaire. Par unicité de l'adjoint, on a donc $u^* = u^{-1}$. □

Références :

- Aucune.

2.13 Simplicité de A_n

Théorème. *Pour $n \neq 4$, le groupe alterné A_n est simple.*

Démonstration. On commence par les cas où $n \leqslant 4$: $A_1 = S_1 = \{\mathrm{Id}\}$ est simple, $A_2 = \{\mathrm{Id}\}$ est simple, $\mathrm{Card}(A_3) = \frac{3!}{2} = 3$ donc A_3 est simple (le cardinal d'un sous-groupe de A_3 divise 3, donc vaut 1 ou 3, c'est donc soit $\{\mathrm{Id}\}$, soit A_3), A_4 n'est pas simple car il contient $V_4 = \{\mathrm{Id}, (12)(34), (13)(24), (14)(23)\}$ comme sous-groupe distingué non trivial.

Soit maintenant $n \geqslant 5$. On commence par montrer que les 3-cycles sont conjugués dans A_n. Soient $c_1 = (i_1 i_2 i_3)$ et $c_2 = (j_1 j_2 j_3)$ deux 3-cycles. Si $\sigma \in S_n$, on vérifie facilement que

$$\sigma(i_1 i_2 i_3)\sigma^{-1} = (\sigma(i_1)\sigma(i_2)\sigma(i_3)).$$

On choisit alors σ telle que $\sigma(i_1) = j_1$, $\sigma(i_2) = j_2$, $\sigma(i_3) = j_3$ et envoyant bijectivement $[\![1, n]\!] \backslash \{i_1, i_2, i_3\}$ sur $[\![1, n]\!] \backslash \{j_1, j_2, j_3\}$. Si $\sigma \in A_n$, c'est terminé. Sinon, puisque $n \geqslant 5$, il existe $i_4, i_5 \notin \{i_1, i_2, i_3\}$ et alors en notant $\tau = (i_4 i_5)$ et $\sigma' = \sigma\tau$, on a

$$\sigma' c_1 \sigma'^{-1} = c_2$$

et $\sigma' \in A_n$ (en effet, $\varepsilon(\sigma') = \varepsilon(\sigma)\varepsilon(\tau) = 1$).

Soit $H \neq \{\mathrm{Id}\}$ un sous-groupe distingué de A_n. On va montrer que H contient un 3-cycle. Ainsi, H contiendra tous les 3-cycles car on vient de montrer qu'ils sont conjugués dans A_n, et comme H est distingué dans A_n, il contient tous les conjugués de ses éléments. Mais on sait que A_n est engendré par les 3-cycles, donc on aura $H = A_n$, *i.e.* A_n est simple.

Soit r le minimum des cardinaux des supports des éléments de $H \backslash \{\mathrm{Id}\}$. Si $r = 3$, on aura le résultat.

Soit $\sigma \in H$, $\sigma \neq \mathrm{Id}$. Il existe $a, b \in [\![1, n]\!]$ tels que $b = \sigma(a) \neq a$. Soient s le 3-cycle (abc) avec $c \notin \{a, b\}$ et $\tau = s\sigma s^{-1}\sigma^{-1}$. Comme $\sigma \in H$, $s \in A_n$ et H distingué dans A_n, on a : $\tau = \underbrace{s\sigma s^{-1}}_{\in H}\underbrace{\sigma^{-1}}_{\in H} \in H$. Comme $s^{-1} = (acb)$, on a :

$$\sigma s^{-1}\sigma^{-1} = (\sigma(a)\sigma(c)\sigma(b)) = (b\sigma(c)\sigma(b)) \quad \text{et} \quad \tau = (abc)(b\sigma(c)\sigma(b)).$$

Ainsi τ est une permutation dont le support est inclu dans $\{a, b, c, \sigma(c), \sigma(b)\}$. De plus, pour s'assurer que $\tau \neq \mathrm{Id}$, choisissons $c \neq \sigma(b)$. Ainsi $\tau(\sigma(b)) = c \neq \sigma(b)$. On a donc prouvé que $r \leqslant 5$.

Supposons par l'absurde que $r = 5$. Alors H contient un élément h dont le support est de cardinal 5 : c'est nécessairement un 5-cycle car le produit d'un 3-cycle et d'une transposition n'est pas dans A_n. On note $h = (i_1 i_2 i_3 i_4 i_5)$. Alors après

calculs :

$$\underbrace{h}_{\in H}\underbrace{\big((i_2i_5)(i_1i_4)\big)h^{-1}\big((i_2i_5)(i_1i_4)\big)^{-1}}_{\in H \text{ car } (i_2i_5)(i_1i_4)\in A_n} = (i_1i_4i_3) \in H,$$

ce qui est absurde car H contiendrait un élément dont le support est de cardinal $3 < r$. Donc $r \leqslant 4$.

Supposons maintenant par l'absurde que $r = 4$. Alors H contient un élément h dont le support est de cardinal 4 : c'est nécessairement un produit de deux transpositions à supports disjoints car les 4-cycles ne sont pas dans A_n. On note $h = (i_1i_2)(i_3i_4)$. Alors après calculs :

$$\underbrace{h}_{\in H}\underbrace{(i_3i_4i_5)h^{-1}(i_3i_4i_5)^{-1}}_{\in H \text{ car } (i_3i_4i_5)\in A_n} = (i_3i_4i_5) \in H,$$

ce qui est absurde car H contiendrait un élément dont le support est de cardinal $3 < r$. Donc $r \leqslant 3$.

Enfin, aucune transposition n'est dans H car les transpositions ne sont pas dans A_n. D'où $r \geqslant 3$, et donc $r = 3$, ce qui achève la démonstration. □

Applications possibles :

Application. *Pour $n \neq 4$, les seuls sous-groupes distingués de S_n sont* $\{\mathrm{Id}\}$, *A_n et S_n.*

Démonstration. Soit H un sous-groupe distingué de S_n. Alors $H \cap A_n$ est un sous-groupe distingué de A_n. Donc par le théorème précédent, on a soit $H \cap A_n = A_n$, soit $H \cap A_n = \{\mathrm{Id}\}$.

Si $H \cap A_n = A_n$, alors $A_n \subset H$ et donc $\mathrm{Card}(A_n) = \frac{n!}{2}$ divise $\mathrm{Card}(H)$ qui divise $\mathrm{Card}(S_n) = n!$. On écrit $\mathrm{Card}(H) = k\frac{n!}{2}$ et $n! = l\mathrm{Card}(H)$, ce qui donne $n! = \frac{n!}{2}kl$. Donc nécessairement, $kl = 2$. Si $k = 2$, on obtient $H = S_n$, et si $k = 1$, on obtient $H = A_n$.

Si $H \cap A_n = \{\mathrm{Id}\}$, alors $\mathrm{Ker}(\varepsilon_{|H}) = H \cap A_n = \{\mathrm{Id}\}$. Donc la signature ε induit un isomorphisme de H sur $\varepsilon(H) \subset \{-1, 1\}$. Donc $\mathrm{Card}(H) \leqslant 2$. Si $\mathrm{Card}(H) = 2$, il existe $\tau \in S_n$ tel que $H = \{\mathrm{Id}, \tau\}$. Comme on doit avoir $\tau^{-1} \in H$, nécessairement τ est une transposition. De plus, H étant distingué dans S_n, on a $\sigma\tau\sigma^{-1} \in H$ pour tout $\sigma \in S_n$. Mais si on note $\tau = (ij)$, alors pour tout $\sigma \in S_n$, $\sigma\tau\sigma^{-1} = (\sigma(i)\sigma(j)) \in H = \{\mathrm{Id}, (ij)\}$: c'est clairement impossible. On en déduit que $\mathrm{Card}(H) = 1$ et donc $H = \{\mathrm{Id}\}$. □

Application. *Pour $n \geqslant 5$, il n'existe pas de surjection de S_n dans S_{n-1}.*

Démonstration. Supposons qu'une telle surjection f existe. Alors $S_n/\mathrm{Ker}(f) \simeq \mathrm{Im}(f) = S_{n-1}$ par passage au quotient et surjectivité de f. Donc $[S_n : \mathrm{Ker}(f)] = \mathrm{Card}(S_{n-1}) = (n-1)!$, d'où $\mathrm{Card}(\mathrm{Ker}(f)) = n$. Or $\mathrm{Ker}(f)$ est un sous-groupe distingué de S_n (c'est toujours le cas pour un morphisme de groupes comme on peut le vérifier facilement). De plus, $\mathrm{Ker}(f)$ est non trivial : s'il était réduit au neutre, f serait bijective ce qui est impossible car S_n et S_{n-1} sont de cardinal différent, et s'il était égal à S_n, f serait identiquement nulle, donc non surjective. Le seul sous-groupe distingué de S_n non trivial est A_n d'après l'application précédente, donc $\mathrm{Ker}(f) = A_n$. En prenant les cardinaux, on trouve $n = \frac{n!}{2}$, ce qui est absurde lorsque $n \geqslant 5$. □

Définition. Le support d'une permutation σ est l'ensemble des i tels que $\sigma(i) \neq i$.

Définition. Un groupe G est dit *simple* s'il a exactement deux sous-groupes distingués : $\{e\}$ et lui-même.

Proposition. *On a* $\mathrm{Card}(A_n) = \frac{n!}{2}$.

Démonstration. Soit τ une transposition. Alors $f : \sigma \mapsto \sigma\tau$ est une bijection de S_n qui induit une bijection de A_n sur l'ensemble de permutations impaires. Ces deux ensembles ont donc même cardinal, et puisqu'ils forment une partition de S_n, on a $\mathrm{Card}(A_n) = \frac{\mathrm{Card}(S_n)}{2} = \frac{n!}{2}$. □

Proposition. *Pour $n \geqslant 3$, les cycles de longueur 3 engendrent A_n.*

Démonstration. Puisque tout élément de S_n peut s'écrire comme un produit de transpositions, A_n est aussi l'ensemble des produits pairs de transpositions. Notons A'_n le sous-groupe de S_n engendré par les cycles de longueur 3 et montrons que $A'_n = A_n$. Clairement, $A'_n \subset A_n$ (un cycle de longueur 3 est de signature 1). Pour montrer l'inclusion inverse, il suffit de prouver que le produit de deux transpositions est dans A'_n (car A_n est l'ensemble des produits pairs de transpositions).

Montrons donc que pour tout i, j, k, l, $(i,j)(k,l) \in A'_n$:

- Si i, j, k, l sont deux à deux distincts, alors $(i,j)(k,l) = (i,j,k)(j,k,l) \in A'_n$.
- Si i, j, k sont deux à deux distincts et $l = i$ (par exemple), alors $(i,j)(k,l) = (k,j,i) \in A'_n$.
- Si $i \neq j$, $k = i$ et $l = j$ (par exemple), alors $(i,j)(k,l) = \mathrm{Id} \in A'_n$.

□

Références :

- Risler et Boyer - *Algèbre pour la licence 3.*
- Perrin - *Cours d'algèbre* - Page 30 (pour la première application).

2.14 Sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$

Lemme. *Soient E un espace euclidien, K un convexe compact de E et H un sous-groupe compact de $\mathrm{GL}(E)$. Si K est stable par tous les éléments de H, alors il existe $a \in K$ qui est fixé par tous les éléments de H.*

Démonstration. Soit $\|\cdot\|$ une norme euclidienne sur E. Pour $x \in E$, on pose

$$N(x) = \sup_{u \in H} \|u(x)\| = \max_{u \in H} \|u(x)\|,$$

le borne supérieure étant atteinte sur le compact H (l'application $u \mapsto \|u(x)\|$ est continue grâce à l'inégalité triangulaire bis, et en prenant la norme triple sur $\mathrm{GL}(E)$). Alors N est une norme sur E. En effet :

(i) Si $N(x) = 0$, comme $\mathrm{Id}_E \in H$, on a en particulier $\|\mathrm{Id}_E(x)\| = \|x\| = 0$, *i.e.* $x = 0$.

(ii) On a clairement $N(\lambda x) = |\lambda| N(x)$ pour tout $\lambda \in \mathbb{R}$.

(iii) $\|u(x+y)\| = \|u(x) + u(y)\| \leqslant \|u(x)\| + \|u(y)\| \leqslant N(x) + N(y)$ pour tout $u \in H$ et tout $x, y \in E$. En passant au sup à gauche, on obtient $N(x+y) \leqslant N(x) + N(y)$ pour tout $x, y \in E$.

De plus, N vérifie : pour tout $v \in H$, $N(v(x)) = N(x)$, ce qui est clair car $u \mapsto u \circ v$ est une permutation de H.

On montre ensuite que N vérifie : si $N(x+y) = N(x) + N(y)$, alors x et y sont positivement liés. Soient $x, y \in E$ tels que $N(x+y) = N(x) + N(y)$. Soit $u_0 \in H$ tel que le sup définissant $N(x+y)$ soit atteint : $N(x+y) = \|u_0(x+y)\|$. Alors :

$$N(x+y) = \|u_0(x+y)\| \leqslant \|u_0(x)\| + \|u_0(y)\| \leqslant N(x) + N(y) = N(x+y).$$

Les inégalités précédentes sont donc des égalités, d'où $\|u_0(x+y)\| = \|u_0(x)\| + \|u_0(y)\|$, ce qui implique que $u_0(x)$ et $u_0(y)$ sont positivement liés (une norme euclidienne vérifie cette propriété). Comme u_0 est linéaire et inversible, cela entraîne immédiatement que x et y sont positivement liés. Le résultat est donc établi.

Par compacité de K et continuité de N sur K (une norme est toujours continue), il existe $a \in K$ tel que $N(x) \geqslant N(a)$ pour tout $x \in K$. Ce point a est unique : si $b \neq a$ vérifie $N(b) = N(a)$, comme K est convexe, $\frac{a+b}{2} \in K$. Si a et b sont positivement liés, alors $a = \lambda b$ avec $\lambda > 0$ (par exemple), et $N(a) = N(b)$ entraîne $\lambda = 1$, ce qui est absurde. Donc $\frac{a}{2}$ et $\frac{b}{2}$ ne sont pas positivement liés, d'où, d'après ce qui précède :

$$N\left(\frac{a+b}{2}\right) < \frac{1}{2}N(a) + \frac{1}{2}N(b) = N(a).$$

Ceci contredit la minimalité de N en a.

On suppose K stable par tous les éléments de H. Donc si $v \in H$, $v(a) \in K$, et d'autre part, $N(v(a)) = N(a)$. Par unicité de a, cela implique $v(a) = a$. Le point a est donc un point de K fixé par tous les éléments de H. □

Théorème. *Tout sous-groupe compact de* $\mathrm{GL}_n(\mathbb{R})$ *est conjugué à un sous-groupe de* $\mathcal{O}_n(\mathbb{R})$.

Démonstration. Soit G un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$. On munit G d'une nouvelle structure de groupe $(G, *)$ définie par $A * B = BA$ pour tout $A, B \in G$ (c'est bien une loi interne). On considère l'application :

$$\begin{array}{rccc} \rho : & (G,*) & \longrightarrow & \big(\mathrm{GL}(\mathcal{S}_n(\mathbb{R})), \circ\big) \\ & A & \longmapsto & \big(S \mapsto {}^tASA\big). \end{array}$$

On vérifie facilement que ρ est bien définie et qu'il s'agit d'un morphisme de groupes (grâce au changement de la loi sur G, on a bien $\rho(A * B) = \rho(BA) = \rho(A) \circ \rho(B)$). De plus, ρ est continue : pour le voir on peut écrire $\rho = b \circ f$ avec $b : (A, B) \mapsto (S \mapsto {}^tASB)$ application bilinéaire de $\mathcal{M}_n(\mathbb{R})^2$ dans $\mathcal{L}(\mathcal{S}_n(\mathbb{R}), \mathcal{M}_n(\mathbb{R}))$ continue car dimensions finies, et $f : A \mapsto (A, A)$ bien sûr continue.

On pose $J = \{{}^tMM, M \in G\}$ qui est un compact (non vide) du convexe $\mathcal{S}_n^{++}(\mathbb{R})$ comme image du compact G par l'application continue $M \mapsto {}^tMM$. Le corollaire du théorème de Carathéodory nous dit que l'enveloppe convexe $K = \mathrm{Cv}(J)$ reste compacte dans $\mathcal{S}_n^{++}(\mathbb{R})$. De plus, $H = \rho(G)$ est un sous-groupe (car ρ est un morphisme de groupes) compact (car ρ est continue et G est compact) de $\mathrm{GL}(\mathcal{S}_n(\mathbb{R}))$. Enfin, on montre que K est stable par tous les éléments de H : si $A \in G$ et $M \in G$, alors

$$\rho(A)({}^tMM) = {}^tA{}^tMMA = {}^t(MA)MA \in J,$$

car G étant un sous-groupe, $MA \in G$. Comme $\rho(A)$ est linéaire, il conserve les combinaisons convexes, donc $\rho(A)(K) \subset K$ (on se sert de la caractérisation : K est l'ensemble des combinaisons convexes d'éléments de J), *i.e.* K est stable par tous les éléments de H.

On peut donc appliquer le lemme : il existe $S \in K$ tel que pour tout $A \in G$, on ait $\rho(A)(S) = S$, *i.e.* ${}^tASA = S$. Comme $K \subset \mathcal{S}_n^{++}(\mathbb{R})$, on a $S \in \mathcal{S}_n^{++}(\mathbb{R})$ et on obtient $G \subset \mathcal{O}_n(q, \mathbb{R})$ où $q : x \mapsto {}^txSx$ est une forme quadratique définie positive. En réduisant la forme quadratique q, on obtient enfin que G est conjugué à un sous-groupe de $\mathcal{O}_n(\mathbb{R})$. □

Compléments de la démonstration :

1. Justifions que $J \subset \mathcal{S}_n^{++}(\mathbb{R})$: si $A = {}^tMM \in J$, alors ${}^tXAX = {}^t(MX)MX = \|MX\|^2 \geqslant 0$. Or $M \in G \subset \mathrm{GL}_n(R)$, donc si ${}^tXAX = 0$, alors $MX = 0$, et donc $X = 0$.
2. Justifions que $\mathcal{S}_n^{++}(\mathbb{R})$ est convexe : si $A, B \in \mathcal{S}_n^{++}(\mathbb{R})$, alors pour $t \in [0,1]$, $M = tA + (1-t)B \in \mathcal{S}_n^{++}(\mathbb{R})$. En effet, on a directement ${}^tM = M$, d'où $M \in \mathcal{S}_n(\mathbb{R})$. De plus, si $X \neq 0$,

$$ {}^tXMX = t\underbrace{{}^tXAX}_{>0} + (1-t)\underbrace{{}^tXBX}_{>0} > 0.$$

   Donc $M \in \mathcal{S}_n^{++}(\mathbb{R})$.
3. Justifions que $G \subset \mathcal{O}_n(q, \mathbb{R})$ implique que G est conjugué à un sous-groupe de $\mathcal{O}_n(\mathbb{R})$. Par définition,

$$\mathcal{O}_n(q,\mathbb{R}) = \left\{ M \in \mathcal{M}_n(\mathbb{R}) \,/\, {}^tMSM = S \right\}.$$

   Mais S est diagonalisable au moyen d'une matrice orthogonale pour le produit scalaire usuel : il existe $P \in \mathcal{O}_n(\mathbb{R})$ telle que $S = {}^tPDP$ avec D diagonale dont les éléments diagonaux sont strictement positifs. On écrit $D = C^2$ avec C diagonale dont les éléments diagonaux sont strictement positifs (en prenant les racines carrées des éléments diagonaux de D). Alors

$$\begin{aligned}
\mathcal{O}_n(q,\mathbb{R}) &= \left\{ M \in \mathcal{M}_n(\mathbb{R}) \,/\, {}^t(CPM{}^tPC^{-1})(CPM{}^tPC^{-1}) = I_n \right\} \\
&= \left\{ M \in \mathcal{M}_n(\mathbb{R}) \,/\, CPM{}^tPC^{-1} \in \mathcal{O}_n(\mathbb{R}) \right\} \\
&= \left\{ {}^tPC^{-1}XCP,\, X \in \mathcal{O}_n(\mathbb{R}) \right\} \\
&= {}^tPC^{-1}\mathcal{O}_n(\mathbb{R})CP \\
&= (CP)^{-1}\mathcal{O}_n(\mathbb{R})CP.
\end{aligned}$$

   Finalement, $G \subset (CP)^{-1}\mathcal{O}_n(\mathbb{R})CP$, soit encore $CPG(CP)^{-1} \subset \mathcal{O}_n(\mathbb{R})$. Mais $CPG(CP)^{-1}$ reste un sous-groupe, donc $CPG(CP)^{-1} = H$ avec H sous-groupe de $\mathcal{O}_n(\mathbb{R})$. D'où

$$G = (CP)^{-1}HCP.$$

Application. *Le groupe $\mathcal{O}_n(\mathbb{R})$ est un sous-groupe compact maximal de* $\mathrm{GL}_n(\mathbb{R})$.

Démonstration. Soit G un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ avec $\mathcal{O}_n(\mathbb{R}) \subset G \subset \mathrm{GL}_n(\mathbb{R})$. D'après ce qui précède, il existe un sous-groupe H de $\mathcal{O}_n(\mathbb{R})$, il existe C diagonale à coefficients diagonaux strictement positifs et $P \in \mathcal{O}_n(\mathbb{R})$ tels que

$$G = (CP)^{-1}HCP.$$

Le fait que $\mathcal{O}_n(\mathbb{R}) \subset G \subset (CP)^{-1}\mathcal{O}_n(\mathbb{R})CP$ entraîne que $CP\mathcal{O}_n(\mathbb{R})(CP)^{-1} \subset \mathcal{O}_n(\mathbb{R})$. Or $P\mathcal{O}_n(\mathbb{R})P^{-1} = \mathcal{O}_n(\mathbb{R})$ car $P \in \mathcal{O}_n(\mathbb{R})$. Donc $C\mathcal{O}_n(\mathbb{R})C^{-1} \subset \mathcal{O}_n(\mathbb{R})$. On note $C = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ et

$$J = \begin{pmatrix} 0 & \dots & 1 \\ \vdots & \iddots & \vdots \\ 1 & \dots & 0 \end{pmatrix}.$$

On vérifie facilement que $J \in \mathcal{O}_n(\mathbb{R})$. On calcule ensuite :

$$CJC^{-1} = \begin{pmatrix} 0 & \dots & \dots & 0 & \frac{\lambda_1}{\lambda_n} \\ 0 & \dots & \dots & \frac{\lambda_2}{\lambda_{n-1}} & 0 \\ \vdots & & \iddots & & \vdots \\ 0 & \frac{\lambda_{n-1}}{\lambda_2} & \dots & \dots & \vdots \\ \frac{\lambda_n}{\lambda_1} & 0 & \dots & \dots & 0 \end{pmatrix}$$

et cette matrice appartient à $\mathcal{O}_n(\mathbb{R})$. Donc ${}^t(CJC^{-1})(CJC^{-1}) = I_n$, *i.e.* :

$$\begin{pmatrix} 0 & \dots & \frac{\lambda_n}{\lambda_1} \\ \vdots & \iddots & \vdots \\ \frac{\lambda_1}{\lambda_n} & \dots & 0 \end{pmatrix} \begin{pmatrix} 0 & \dots & \frac{\lambda_1}{\lambda_n} \\ \vdots & \iddots & \vdots \\ \frac{\lambda_n}{\lambda_1} & \dots & 0 \end{pmatrix} = \begin{pmatrix} \left(\frac{\lambda_n}{\lambda_1}\right)^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \left(\frac{\lambda_1}{\lambda_n}\right)^2 \end{pmatrix} = I_n.$$

Comme les λ_i sont positifs, on en déduit $\lambda_n = \lambda_1$, $\lambda_{n-1} = \lambda_2$, etc... On recommence cela en faisant monter d'une ligne tous les 1 dans la matrice J :

$$J = \begin{pmatrix} 0 & \dots & 0 & 1 & 0 \\ \vdots & & \iddots & 0 & 0 \\ 0 & 1 & \iddots & & \vdots \\ 1 & 0 & & & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}$$

On répète cela jusqu'à retomber sur la première matrice J. Au final, on obtient $\lambda_i = \lambda_j$, pour tout $i \neq j$, *i.e.* $C = \lambda_1 I_n$. Donc C commute avec toutes les matrices, d'où :

$$\mathcal{O}_n(\mathbb{R}) \subset G \subset (CP)^{-1}\mathcal{O}_n(\mathbb{R})CP = P^{-1}\mathcal{O}_n(\mathbb{R})P = \mathcal{O}_n(\mathbb{R}),$$

la dernière égalité étant due au fait que $P \in \mathcal{O}_n(\mathbb{R})$. Soit enfin $G = \mathcal{O}_n(\mathbb{R})$. □

Remarque. On peut partir juste du fait que G est conjugué à un sous-groupe de $\mathcal{O}_n(\mathbb{R})$. Ainsi il existe $Q \in \mathrm{GL}_n(\mathbb{R})$ tel que $G \subset Q^{-1}\mathcal{O}_n(\mathbb{R})Q$, d'où encore

$$Q\mathcal{O}_n(\mathbb{R})Q^{-1} \subset \mathcal{O}_n(\mathbb{R}).$$

En écrivant $Q = SO$ avec S symétrique réelle définie positive et $O \in \mathcal{O}_n(\mathbb{R})$ par la décomposition polaire, on obtient $S\mathcal{O}_n(\mathbb{R})S^{-1} \subset \mathcal{O}_n(\mathbb{R})$. Il existe $P \in \mathcal{O}_n(\mathbb{R})$ telle que $S = PCP^{-1}$ avec C diagonale à coefficients diagonaux strictement positifs. On obtient alors

$$S\mathcal{O}_n(\mathbb{R})S^{-1} = PC\mathcal{O}_n(\mathbb{R})(PC)^{-1} \subset \mathcal{O}_n(\mathbb{R}),$$

et en multipliant à gauche par P^{-1} et à droite par P :

$$C\mathcal{O}_n(\mathbb{R})C^{-1} \subset P^{-1}\mathcal{O}_n(\mathbb{R})P = \mathcal{O}_n(\mathbb{R}).$$

On est donc ramené à la démonstration qui précède.

Soient E un espace affine euclidien de dimension finie non nulle et A une partie de E.

Définition. L'enveloppe convexe d'une partie A de E, notée $\mathrm{Cv}(A)$, est l'intersection de tous les convexes contenant A. C'est donc le plus petit convexe de E contenant A.

Proposition.

(i) $\mathrm{Cv}(A)$ *est l'ensemble des combinaisons convexes d'éléments de* A.

(ii) Si A *est convexe et compacte, on a* $A = \mathrm{Cv}(\mathrm{Fr}(A))$ *où* $\mathrm{Fr}(A)$ *désigne la frontière de* A.

Théorème (Carathéodory)**.** *Tout élément de* $\mathrm{Cv}(A)$ *s'écrit comme combinaison convexe de* k *points de* A *avec* $k \leqslant \dim(E) + 1$.

Démonstration. Soit $x \in \mathrm{Cv}(A)$. On écrit $x = t_1a_1 + \cdots + t_ka_k$ avec $k \in \mathbb{N}^*$, $t_i > 0$ et $\sum_{i=1}^k t_i = 1$. On suppose $k > \dim(E)+1$. Alors la famille $((a_2-a_1), \ldots, (a_k-a_1))$ est liée car de cardinal $> \dim(E)$. Donc il existe $\lambda_2, \ldots, \lambda_k$ non tous nuls tels que

$$\lambda_2(a_2 - a_1) + \cdots + \lambda_k(a_k - a_1) = 0.$$

On note $\mu_1 = \lambda_2 + \cdots + \lambda_k$ et $\mu_i = -\lambda_i$ pour tout $i \geqslant 2$. Alors pour tout $y \in E$,

$$\mu_1(a_1 - y) + \cdots + \mu_k(a_k - y) = 0.$$

En particulier, l'égalité est vraie pour x. On a d'autre part $\sum_{j=1}^k \mu_j = 0$, donc il existe j tel que $\mu_j > 0$. Soit

$$\lambda = \min\left(\frac{t_i}{\mu_i},\ \mu_i > 0\right) = \frac{t_{i_0}}{\mu_{i_0}}.$$

Soit ensuite $\nu_i = t_i - \lambda\mu_i$. Alors, sachant que $t_i \neq 0$ (car $t_i > 0$),

$$\nu_i = t_i \left(1 - \underbrace{\frac{t_{i_0}}{\mu_{i_0}} \frac{\mu_i}{t_i}}_{\leqslant 1} \right) \geqslant 0.$$

D'autre part,

$$\sum_{i=1}^{k} \nu_i = \underbrace{\sum_{i=1}^{k} t_i}_{=1} - \lambda \underbrace{\sum_{i=1}^{k} \mu_i}_{=0} = 1.$$

Et de plus, il existe q tel que $\nu_q = 0$. Donc

$$\begin{aligned} x &= t_1 a_1 + \cdots + t_k a_k \\ &= t_1 a_1 + \cdots + t_k a_k - \lambda\big(\mu_1(a_1 - x) + \cdots + \mu_k(a_k - x)\big) \\ &= (t_1 - \lambda\mu_1)a_1 + \cdots + (t_k - \lambda\mu_k)a_k - \lambda \underbrace{\left(\sum_{i=1}^{k} \mu_i \right)}_{=0} x \\ &= \sum_{i=1}^{k} \nu_i a_i = \sum_{i \neq q} \nu_i a_i. \end{aligned}$$

On a ainsi écrit x comme combinaison convexe de $k-1$ points de A. □

Corollaire. *On suppose A non vide. Si A est compacte, il en est de même de* $\mathrm{Cv}(A)$.

Démonstration. Soit $n = \dim(E)$. On note

$$K = \left\{ (t_1, \ldots, t_{n+1}) \in [0,1]^{n+1} \,/\, t_1 + \cdots + t_{n+1} = 1 \right\}.$$

L'ensemble K est compact car fermé borné, fermé comme image réciproque de 1 par la fonction continue $g : (t_1, \ldots, t_{n+1}) \mapsto t_1 + \cdots + t_{n+1}$. Soit

$$\begin{array}{rccc} f : & K \times E^{n+1} & \longrightarrow & E \\ & (t_1, \ldots, t_{n+1}, a_1, \ldots, a_{n+1}) & \longmapsto & t_1 a_1 + \cdots + t_{n+1} a_{n+1}. \end{array}$$

D'après le théorème de Carathéodory, $\mathrm{Cv}(A) = f(K \times A^{n+1})$. Comme f est continue et $K \times A^{n+1}$ est compact, on en déduit que $\mathrm{Cv}(A)$ est compact. □

Autre résultat intéressant :

Proposition. *On suppose A non vide. Si A est bornée, il en est de même de* $\mathrm{Cv}(A)$. *De plus, $\delta(A) = \delta(\mathrm{Cv}(A))$ où $\delta(A) = \sup_{x,y \in A} \|x - y\|$.*

Démonstration. Comme la partie A est bornée, elle est contenue dans une boule fermée $B = \overline{B}(a, r)$. Or, B étant convexe, il vient $\mathrm{Cv}(A) \subset B$. D'où $\mathrm{Cv}(A)$ est bornée.

Il est clair que $\delta(A) \leqslant \delta(\mathrm{Cv}(A))$.

Soit maintenant $x = t_1 a_1 + \cdots + t_k a_k \in \mathrm{Cv}(A)$ écrit comme combinaison convexe d'éléments de A. Si $a \in A$, alors

$$\begin{aligned} \|a - x\| &= \|t_1(a - a_1) + \cdots + t_k(a - a_k)\| \\ &\leqslant t_1\|a - a_1\| + \cdots + t_k\|a - a_k\| \\ &\leqslant \delta(A)(t_1 + \cdots + t_k) = \delta(A), \end{aligned}$$

sachant que $\|a - a_i\| \leqslant \delta(A)$ car $a, a_i \in A$. Soit $y \in \mathrm{Cv}(A)$. Alors :

$$\begin{aligned} \|y - x\| &= \|t_1(y - a_1) + \cdots + t_k(y - a_k)\| \\ &\leqslant t_1\|y - a_1\| + \cdots + t_k\|y - a_k\| \\ &\leqslant \delta(A)(t_1 + \cdots + t_k) = \delta(A), \end{aligned}$$

sachant que $\|y - a_i\| \leqslant \delta(A)$ d'après l'inégalité précédente avec $y \in \mathrm{Cv}(A)$ et $a_i \in A$.

On en déduit que $\delta(\mathrm{Cv}(A)) \leqslant \delta(A)$, d'où le résultat. □

Références :

- Alessandri - *Thèmes de géométrie* - Pages 59, 141 et 160.
- Tauvel - *Cours de géométrie* - Page 77 (pour le théorème de Carathéodory).

2.15 Surjectivité de l'exponentielle

Lemme. *Soient G un groupe topologique et H un sous-groupe de G contenant un voisinage du neutre e de G. Alors H est ouvert et fermé dans G.*

Démonstration. Montrons que H est ouvert. Par hypothèse, il existe un voisinage ouvert V de e dans G avec $V \subset H$. Alors si $h \in H$, $hV \subset H$. Si on note $\varphi : g \mapsto h^{-1}g$ le morphisme de multiplication à gauche par h^{-1} dans G qui est continu, alors $hV = \varphi^{-1}(V)$ est ouvert comme image réciproque d'un ouvert par une application continue. Ainsi H est un voisinage de chacun de ses points, donc H est ouvert.

Montrons que H est fermé. On a clairement $G = \bigcup_{g\in G} gV$ et $H = \bigcup_{h\in H} hV$. Donc ${}^cH = \bigcup_{g\notin H} gV$ (si $g \notin H$, $gV \subset {}^cH$ car s'il existe $gv \in H$, alors on aurait $g \in H$). Comme gV est ouvert pour tout g, ${}^cH$ est ouvert, et donc H est fermé. □

Théorème. *Soit $A \in \mathrm{GL}_n(\mathbb{C})$. On note*

$$\mathbb{C}[A] = \{P(A),\, P \in \mathbb{C}[X]\} \quad \textit{et} \quad U = \mathbb{C}[A] \cap \mathrm{GL}_n(\mathbb{C}).$$

Alors $\exp(\mathbb{C}[A]) = U$.

En particulier, il existe $P \in \mathbb{C}[X]$ tel que $A = e^{P(A)}$, et $\exp : \mathcal{M}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ *est surjective.*

Démonstration. Soit $A \in \mathrm{GL}_n(\mathbb{C})$. D'abord, $\mathbb{C}[A] = \{P(A),\, P \in \mathbb{C}[X]\}$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{C})$ de dimension finie, donc fermé. Pour toute matrice $M \in \mathbb{C}[A]$, comme e^M est limite d'éléments de $\mathbb{C}[A]$, on a $e^M \in \mathbb{C}[A]$. De plus e^M est inversible d'inverse $e^{-M} \in \mathbb{C}[A]$. Comme tous les éléments de $\mathbb{C}[A]$ commutent, on en déduit que exp induit un homomorphisme du groupe additif $(\mathbb{C}[A], +)$ dans le groupe multiplicatif U des inversibles de $\mathbb{C}[A]$. On vérifie bien que U est un groupe : si $M \in U$, $M^{-1} \in \mathbb{C}[M] \subset \mathbb{C}[A]$ par le théorème de Cayley-Hamilton, donc dans U (sinon on peut aussi dire que comme M est inversible, X et Π_M le polynôme minimal de M sont premiers entre eux, et avec le théorème de Bézout, il existe deux polynômes P et Q tels que $PX + Q\Pi_M = 1$, d'où $M^{-1} = P(M) \in \mathbb{C}[M] \subset \mathbb{C}[A]$).

On va maintenant montrer que $\exp(\mathbb{C}[A])$ est un ouvert-fermé de U en utilisant le lemme. On sait d'une part que l'application exp est de classe $\mathcal{C}^1$ sur $\mathcal{M}_n(\mathbb{C})$. On trouve facilement que la différentielle de exp : $\mathbb{C}[A] \to U$ en $0 \in \mathbb{C}[A]$ est $D\exp(0) = \mathrm{Id}$ (en écrivant $e^{0+H} = I_n + H + o(\|H\|)$), qui est inversible. On peut donc appliquer le théorème d'inversion locale : il existe un voisinage ouvert V_0 de 0 dans $\mathbb{C}[A]$ et un voisinage ouvert V de I_n dans U tels que $\exp_{|V_0} : V_0 \to V$ soit un $\mathcal{C}^1$-difféomorphisme. Par le lemme précédent avec $G = U$, $H = \exp(\mathbb{C}[A])$ et $V \subset H$ voisinage ouvert de I_n, on en déduit que $\exp(\mathbb{C}[A])$ est ouvert et fermé dans U.

Montrons à présent que U est connexe. Soient $M, N \in U$ et $f : z \mapsto zM+(1-z)N$ définie pour $z \in \mathbb{C}$, et à valeurs dans $\mathbb{C}[A]$. Le polynôme en z, $\det(zM + (1-z)N)$, s'annule sur un ensemble Z fini. Or $\mathbb{C}\backslash Z$ est connexe par arcs. Donc il existe un chemin continu $\gamma : [0,1] \to \mathbb{C}\backslash Z$ tel que $\gamma(0) = 0$ et $\gamma(1) = 1$. Ainsi $f \circ \gamma : [0,1] \to U$ est un chemin continu qui relie M à N. Donc U est connexe par arcs, donc connexe.

Finalement, $\exp(\mathbb{C}[A])$ est un ouvert-fermé de U qui est connexe, donc

$$\exp(\mathbb{C}[A]) = U.$$

En particulier, comme $A \in U$, il existe $P \in \mathbb{C}[X]$ tel que $A = \exp(P(A))$. □

Corollaire. *Si $A \in \mathrm{GL}_n(\mathbb{C})$ et $k \in \mathbb{N}^*$, alors il existe $B \in \mathrm{GL}_n(\mathbb{C})$ tel que $A = B^k$.*

Démonstration. $B = \exp\left(\frac{P(A)}{k}\right)$ convient. □

Remarque. $\exp : \mathcal{M}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ n'est pas surjective : exp est continue et $\mathrm{GL}_n(\mathbb{R})$ a deux composantes connexes alors que $\mathcal{M}_n(\mathbb{R})$ est connexe.

Application. *Soit $A \in \mathrm{GL}_n(\mathbb{R})$. Alors il existe $M \in \mathcal{M}_n(\mathbb{R})$ telle que $A = e^M$ si et seulement s'il existe $B \in \mathcal{M}_n(\mathbb{R})$ telle que $A = B^2$.*

Démonstration. Le sens direct est évident en prenant $B = e^{\frac{M}{2}}$.

Sens indirect : soit $A \in \mathrm{GL}_n(\mathbb{R})$. On suppose qu'il existe $B \in \mathrm{GL}_n(\mathbb{R})$ telle que $A = B^2$. On a aussi $B \in \mathrm{GL}_n(\mathbb{C})$, donc d'après le théorème, il existe $Q \in \mathbb{C}[X]$ tel que $B = e^{Q(B)}$. Mais comme B est une matrice réelle, on a

$$B = \overline{B} = \overline{e^{Q(B)}} = e^{\overline{Q(B)}}.$$

On obtient alors

$$A = BB = B\overline{B} = e^{Q(B)}e^{\overline{Q(B)}}.$$

Or $Q(B)$ et $\overline{Q(B)} = \overline{Q}(B)$ commutent comme polynômes en B, donc :

$$A = e^{Q(B)+\overline{Q(B)}} = e^M$$

avec $M = Q(B) + \overline{Q(B)} \in \mathcal{M}_n(\mathbb{R})$, d'où le résultat. □

Définition. Un groupe topologique est un triplet $(G, \cdot, T)$ où $(G, \cdot)$ est un groupe, (G, T) est un espace topologique, et ces deux notions sont compatibles : les applications $(g, h) \mapsto gh$ et $g \mapsto g^{-1}$ sont continues.

Définition. Un espace topologique E est connexe si et seulement s'il ne peut pas s'écrire comme réunion de deux ouverts non vides disjoints de E. C'est équivalent à dire que les seuls sous-ensembles de E à la fois ouverts et fermés dans E sont E et $\varnothing$.

Une proposition intéressante liée à l'exponentielle mais sans rapport avec sa surjectivité :

Proposition. $\mathrm{GL}_n(\mathbb{C})$ *n'admet pas de sous-groupes arbitrairement petits.*

Démonstration. Par le théorème d'inversion locale appliqué à exp, il existe U voisinage de 0 dans $\mathcal{M}_n(\mathbb{C})$ et V voisinage de $e^0 = I_n$ dans $\mathrm{GL}_n(\mathbb{C})$ tels que exp réalise un $\mathcal{C}^1$-difféomorphisme de U sur V.

On pose $U' = \frac{1}{2}U$ et $V' = \exp(U')$. Alors V' est ouvert : en effet, exp réalise un $\mathcal{C}^1$-difféomorphisme de U' sur V' et donc en notant $\psi : V' \to U'$ la réciproque de exp qui est continue, on a $V' = \psi^{-1}(U')$ qui est l'image réciproque de l'ouvert U' par ψ continue. De plus, V' est un voisinage de I_n.

Soit maintenant H un sous-groupe de $\mathrm{GL}_n(\mathbb{C})$ inclu dans V' et soit $M \in H$. En particulier, $M \in V'$ et il existe $A \in U'$ tel que $M = e^A$. Si on suppose $A \neq 0$, il existe $k \in \mathbb{N}$ tel que $kA \in U \backslash U'$, et ainsi, $\exp(kA) = M^k \in V \backslash V'$. On a donc $M^k \notin H$, c'est qui est absurde, donc $A = 0$, $M = I_n$ et $H = \{I_n\}$. □

Références :

– Aucune.

2.16 Théorème de Burnside

Théorème. *Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{C})$. Alors G est fini si et seulement si G est d'exposant fini.*

Démonstration. Le sens direct est immédiat : si G est fini de cardinal p, alors par le théorème de Lagrange, $A^p = I_n$ pour tout $A \in G$. Donc G est d'exposant e fini avec e qui divise p.

Supposons maintenant que G est un sous-groupe d'exposant e fini. Alors le polynôme $P = X^e - 1$ annule tous les éléments de G. Or P est scindé à racines simples dans $\mathbb{C}$, donc tous les éléments de G sont diagonalisables. De plus, pour tout $A \in G$, on sait que les racines du polynôme caractéristique χ_A sont aussi racines de tout polynôme annulateur. Donc les racines de χ_A, qui sont les valeurs propres de A, sont des racines de P, *i.e.* des racines e-ième de l'unité.

On va construire une application φ injective de G dans $\varphi(G)$ avec $\varphi(G)$ fini. Ainsi on aura $\mathrm{Card}(G) \leqslant \mathrm{Card}(\varphi(G)) < \infty$.

On note $\mathrm{Vect}(G)$ le sous-espace vectoriel de $\mathcal{M}_n(\mathbb{C})$ engendré par G, qui est donc de dimension finie. Les éléments de G forment une famille génératrice de $\mathrm{Vect}(G)$ dont on peut extraire une famille libre $C_1, \dots, C_r \in G$, base de $\mathrm{Vect}(G)$. Soit

$$\begin{array}{rccc} \varphi : & G & \longrightarrow & \mathbb{C}^r \\ & A & \longmapsto & (\mathrm{tr}(AC_1), \dots, \mathrm{tr}(AC_r)). \end{array}$$

Montrons que $\mathrm{Im}(\varphi)$ est finie. Les valeurs propres de toute matrice $A \in G$ sont des racines e-ième de l'unité qui sont en nombre fini, donc $\{\mathrm{tr}(A),\ A \in G\}$ est fini. Mais comme G est un groupe, $AC_i \in G$ pour tout $i \in [\![1, r]\!]$ et tout $A \in G$. Donc pour tout i et tout A, $\mathrm{tr}(AC_i) \in \{\mathrm{tr}(B),\ B \in G\}$. Les $\mathrm{tr}(AC_i)$ ne peuvent donc prendre qu'un nombre fini de valeurs, d'où $\mathrm{Im}(\varphi)$ est fini.

Montrons maintenant que φ est injectif. Soient $A, B \in G$ tels que $\varphi(A) = \varphi(B)$. Alors pour tout i, $\mathrm{tr}(AC_i) = \mathrm{tr}(BC_i)$. Pour $M \in G \subset \mathrm{Vect}(G)$, il existe $\alpha_1, \dots, \alpha_r$ tels que $M = \alpha_1 C_1 + \cdots + \alpha_r C_r$ et on a :

$$\mathrm{tr}(AM) = \sum_{i=1}^{r} \alpha_i \mathrm{tr}(AC_i) = \sum_{i=1}^{r} \alpha_i \mathrm{tr}(BC_i) = \mathrm{tr}(BM).$$

Soit $N = AB^{-1} - I_n$. Comme $A, B \in G$ et G est un groupe, $AB^{-1} \in G$, donc diagonalisable. Donc N est diagonalisable ($N = PDP^{-1} - I_n = P(D - I_n)P^{-1}$). Si on montre que N est nilpotente, on aura alors $N = 0$, d'où ensuite $A = B$.

Supposons que $\mathrm{tr}(N^p) = 0$ pour tout $p \in [\![1, n]\!]$ et montrons que dans ce cas, N est nilpotente. Soit $\chi_N = X^\alpha (X - \lambda_1)^{\alpha_1} \dots (X - \lambda_r)^{\alpha_r}$ le polynôme caractéristique

de N, où les λ_i sont non nuls et deux à deux distincts, $r \leqslant n$. Pour tout $p \in [\![1, n]\!]$, on a :

$$\mathrm{tr}(N^p) = \sum_{i=1}^{r} \alpha_i \lambda_i^p = 0.$$

Si on suppose que N n'est pas nilpotente, alors 0 n'est pas la seule valeur propre de N, et donc $(\alpha_1, \ldots, \alpha_r)$ est un zéro non trivial du système suivant d'inconnues $x_1, \ldots, x_r$:

$$(S) \begin{cases} \lambda_1 x_1 + \cdots + \lambda_r x_r = 0 \\ \quad \vdots \\ \lambda_1^r x_1 + \cdots + \lambda_r^r x_r = 0 \end{cases}$$

Donc le déterminant V de ce système est nul, soit :

$$V = \det \begin{pmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_r \\ \lambda_1^2 & \lambda_2^2 & \ldots & \lambda_r^2 \\ \vdots & \vdots & \ldots & \vdots \\ \lambda_1^r & \lambda_2^r & \ldots & \lambda_r^r \end{pmatrix} = \lambda_1 \ldots \lambda_r \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_r \\ \vdots & \vdots & \ldots & \vdots \\ \lambda_1^{r-1} & \lambda_2^{r-1} & \ldots & \lambda_r^{r-1} \end{pmatrix} = 0.$$

Or les λ_i étant non nuls et deux à deux distincts, V est non nul car multiple du déterminant de Vandermonde non nul associé aux scalaires λ_i. Donc N est nécessairement nilpotente.

Montrons donc enfin que $\mathrm{tr}(N^p) = 0$ pour tout $p \in [\![1, n]\!]$. Comme AB^{-1} et I_n commutent, on a :

$$N^p = (AB^{-1} - I_n)^p = \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} (AB^{-1})^k.$$

D'autre part, comme $\mathrm{tr}(AM) = \mathrm{tr}(BM)$ pour tout $M \in G$, on a pour $k \in [\![1, n]\!]$,

$$\begin{aligned} \mathrm{tr}((AB^{-1})^k) &= \mathrm{tr}(AB^{-1}(AB^{-1})^{k-1}) \\ &= \mathrm{tr}(BB^{-1}(AB^{-1})^{k-1}) \\ &= \mathrm{tr}((AB^{-1})^{k-1}). \end{aligned}$$

En itérant, on obtient $\mathrm{tr}((AB^{-1})^k) = \mathrm{tr}(I_n) = n$. D'où :

$$\mathrm{tr}(N^p) = \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} n = n(1-1)^p = 0.$$

□

Proposition. *Pour $A \in \mathcal{M}_n(K)$, les racines du polynôme caractéristique χ_A sont aussi racines de tout polynôme annulateur.*

Démonstration. Si λ est racine de χ_A, alors λ est valeur propre de u, et on montre facilement que λ^k est valeur propre de A^k. Par des combinaisons linéaires, on obtient que $P(\lambda)$ est valeur propre de $P(A)$.

Si P est un polynôme annulateur de A, alors $P(A) = 0$. Mais $P(\lambda)$ étant valeur propre de $P(A)$, nécessairement $P(\lambda) = 0$, d'où λ est racine de P. □

Proposition. *Soit $A \in \mathcal{M}_n(K)$. Si Π_A est un polynôme minimal de A, alors les valeurs propres de A sont exactement les racines dans K de Π_A.*

Démonstration. Le polynôme Π_A étant annulateur, les valeurs propres de A doivent être des racines de Π_A d'après la proposition précédente.

Réciproquement, soit α une racine de Π_A : on écrit $\Pi_A = (X - \alpha)Q$. Alors $\Pi_A(A) = (A - \alpha I_n)Q(A) = 0$. Si α n'était pas valeur propre de A, $A - \alpha I_n$ serait injectif. Mais alors $\mathrm{Im}(Q(A)) \subset \mathrm{Ker}(A - \alpha I_n) = \{0\}$, ce qui impliquerait que $Q(A) = 0$ et contredirait la minimalité de Π_A. □

Références :

– Alessandri - *Thèmes de géométrie* - Page 113.

2.17 Théorème de Cartan-Dieudonné

Théorème (Cartan-Dieudonné vectoriel)**.** *Soient E un espace vectoriel euclidien de dimension n et $f \in \mathcal{O}(E)$. On note $F = \mathrm{Ker}(f - \mathrm{Id})$ l'espace des éléments fixés par f et $p(f) = n - \dim(F)$. Alors f s'écrit comme produit de $p(f)$ réflexions et on ne peut pas faire moins.*

Par convention, on dit que Id *est le produit de* 0 *réflexion.*

Démonstration. Raisonnons par récurrence sur $p(f)$. Si $p(f) = 0$, alors $f = \mathrm{Id}$ et donc f est le produit de 0 réflexions.

Supposons que $p(f) \geqslant 1$ et que toute isométrie g telle que $p(g) < p(f)$ s'écrive comme produit de $p(g)$ réflexions et pas moins. Comme $p(f) \geqslant 1$, on a $E = F \oplus F^\perp$ avec $\dim(F^\perp) \geqslant 1$.

Notons $E' = F^\perp$. Alors E' est stable par f car les isométries conservent l'orthogonalité. Donc $f' = f_{|E'}$ est une isométrie de E'. Comme $\dim(E') \geqslant 1$, il existe $x_0 \neq 0$ tel que $x_0 \in E'$, et on a $f'(x_0) \neq x_0$ (sinon $x_0 \in F$). Mais f' étant une isométrie, on a $\|f'(x_0)\| = \|x_0\|$. Soit alors la réflexion s' par rapport à $H' = \mathrm{Vect}(f'(x_0) - x_0)^\perp$ hyperplan de E' (on travaille dans E' et on prend l'orthogonal dans E'). Cette réflexion vérifie

$$s' \circ f'(x_0) = x_0$$

car $s'(f'(x_0) - x_0) = -(f'(x_0) - x_0)$ et $s'(f'(x_0) + x_0) = f'(x_0) + x_0$: l'élément $f'(x_0) + x_0$ est fixé par s' car il est dans H', puisque $(f'(x_0) + x_0, f'(x_0) - x_0) = \|f'(x_0)\|^2 - \|x_0\|^2 = 0$.

On prolonge s' sur E tout entier en posant $s'(x) = x$ si $x \in F$, et s' est maintenant une réflexion de E. On note $g = s' \circ f$ qui est une isométrie de E car une composée d'isométries conserve encore la norme. Par construction, on a $\mathrm{Vect}(x_0) \subset \mathrm{Ker}(g - \mathrm{Id})$ et $F \subset \mathrm{Ker}(g - \mathrm{Id})$, donc $p(g) \leqslant p(f) - 1$. On applique l'hypothèse de récurrence à g : il existe des réflexions s_i de E telles que

$$g = s' \circ f = s_1 \circ \cdots \circ s_{p(g)}.$$

Donc, comme $s' \circ s' = \mathrm{Id}$, on obtient $f = s' \circ s_1 \circ \cdots \circ s_{p(g)}$.

On vient de prouver qu'on peut décomposer f en produit de réflexions. Il reste à faire le lien avec $p(f)$. Soit q le nombre minimal de réflexions qu'on peut utiliser pour décomposer f. On a déjà $q \leqslant p(g) + 1$, et on avait plus haut que $p(g) \leqslant p(f) - 1$, donc $q \leqslant p(f)$. On écrit ensuite $f = s_1 \circ \cdots \circ s_q$ avec s_i réflexion par rapport à un hyperplan H_i. On a d'une part l'inclusion $\bigcap_{i=1}^q H_i \subset F$, et d'autre part $\bigcap_{i=1}^q H_i =$

$\mathrm{Vect}(x_1,\dots,x_q)^\perp$ si on écrit $H_i = \mathrm{Vect}(x_i)^\perp$. On obtient :

$$\underbrace{\dim\left(\mathrm{Vect}(x_1,\dots,x_q)^\perp\right)}_{\geqslant n-q} = \dim\left(\bigcap_{i=1}^{q} H_i\right) \leqslant \dim(F) = n - p(f).$$

On en déduit que $q \geqslant p(f)$, soit finalement $q = p(f)$. □

Théorème (Cartan-Dieudonné affine)**.** *Soient $\mathcal{E}$ un espace affine euclidien et φ une isométrie de $\mathcal{E}$. Alors :*

(i) Si φ a un point fixe, φ peut s'écrire comme produit de $p(\overrightarrow{\varphi})$ réflexions affines.

(ii) Si φ n'a pas de point fixe, φ peut s'écrire comme produit de $p(\overrightarrow{\varphi})+2$ réflexions affines. On a de plus $p(\overrightarrow{\varphi}) < n$.

(on pourrait montrer qu'on ne peut pas écrire φ comme produit de moins de réflexions qu'annoncé ci-dessus).

Démonstration. (i) Il existe $A \in \mathcal{E}$ tel que $\varphi(A) = A$. On a alors, pour tout $M \in \mathcal{E}$,

$$\varphi(M) = A + \overrightarrow{\varphi}(\overrightarrow{AM}).$$

Par le théorème de Cartan-Dieudonné vectoriel, on peut écrire $\overrightarrow{\varphi} = \overrightarrow{s_1} \circ \cdots \circ \overrightarrow{s_p}$ avec $p = p(\overrightarrow{\varphi})$ et les $\overrightarrow{s_i}$ sont des réflexions vectorielles par rapport à des hyperplans vectoriels H_i. On note $s_i : M \mapsto A + \overrightarrow{s_i}(\overrightarrow{AM})$ la réflexion affine d'hyperplan affine associé l'hyperplan passant par A de direction l'hyperplan vectoriel H_i. Comme A est fixé par toutes ces réflexions affines, on obtient :

$$\varphi(M) = s_1 \circ \cdots \circ s_p(M),$$

pour tout $M \in \mathcal{E}$. En effet, on a :

$$\begin{aligned} s_1 \circ s_2(M) &= A + \overrightarrow{s_1}\left(\overrightarrow{As_2(M)}\right) \\ &= A + \overrightarrow{s_1}\left(\overrightarrow{s_2(A)s_2(M)}\right) \\ &= A + \overrightarrow{s_1}\left(\overrightarrow{s_2}(\overrightarrow{AM})\right) \\ &= A + \overrightarrow{s_1} \circ \overrightarrow{s_2}(\overrightarrow{AM}). \end{aligned}$$

(ii) Soit $A \in \mathcal{E}$. Par hypothèse, $\varphi(A) \neq A$ car φ ne fixe aucun point. Soient H l'hyperplan médiateur de $[A, \varphi(A)]$ et σ la réflexion par rapport à H. Alors $\sigma \circ \varphi(A) = A$. D'après le cas précédent, $\sigma \circ \varphi$ s'écrit comme produit d'au plus $p(\overrightarrow{\sigma \circ \varphi}) \leqslant n$ réflexions. En composant par σ, on obtient que φ est produit d'au plus $n+1$ réflexions.

D'autre part, on peut écrire $\varphi = \tau \circ \psi$ où τ est une translation et ψ une isométrie affine possédant un point fixe. De plus, $\overrightarrow{\varphi} = \overrightarrow{\psi}$, donc ψ s'écrit comme produit de $p(\overrightarrow{\varphi})$ réflexions. Il est clair que τ s'écrit comme produit de deux réflexions. Donc φ s'écrit comme produit de $p(\overrightarrow{\varphi}) + 2$ réflexions, et on a, par ce qu'on a dit juste au-dessus, que $p(\overrightarrow{\varphi}) + 2 \leqslant n + 1$, *i.e.* $p(\overrightarrow{\varphi}) < n$.

□

Compléments :

Montrons qu'on ne peut pas écrire φ comme produit de moins de réflexions qu'annoncé dans le théorème de Cartan-Dieudonné affine.

On écrit $\varphi = s_1 \circ \cdots \circ s_k$. Alors $\overrightarrow{\varphi} = \overrightarrow{s_1} \circ \cdots \circ \overrightarrow{s_k}$, donc $k \geqslant p(\overrightarrow{\varphi})$. Ce qui prouve le résultat pour le point (i).

Pour le point (ii), on suppose que $k < p(\overrightarrow{\varphi}) + 2$. Comme le déterminant d'une réflexion vaut -1, on a $\det(\overrightarrow{\varphi}) = (-1)^k = (-1)^{p(\overrightarrow{\varphi})+2}$ puisqu'on sait qu'on peut écrire un produit de $p(\overrightarrow{\varphi}) + 2$ réflexions. Donc nécessairement $k = p(\overrightarrow{\varphi}) < n$.

Soient $\mathcal{H}_i$ l'hyperplan associé à s_i et H_i l'hyperplan associé à $\overrightarrow{s_i}$. On note $\mathcal{F} = \cap\mathcal{H}_i$ et $F = \cap H_i$. Soit $\psi_i \in E^*$ telle que $H_i = \mathrm{Ker}(\psi_i)$. Le point M appartient à $\mathcal{H}_i = A_i + H_i$ SSi $\overrightarrow{A_iM} \in H_i$, SSi $\psi_i(\overrightarrow{A_iM}) = 0$, SSi $\psi_i(\overrightarrow{A_iA}) + \psi_i(\overrightarrow{AM}) = 0$ pour un point A fixé dans $\mathcal{E}$, SSi il existe β_i tel que $\psi_i(\overrightarrow{AM}) = \beta_i$. D'où :

$$\mathcal{F} = \{M \in \mathcal{E} \,/\, \psi_i(\overrightarrow{AM}) = \beta_i, \text{ pour tout } i\}.$$

On a $\dim(F) \geqslant n - k$ (intersection d'hyperplans), et $F \subset \mathrm{Ker}(\overrightarrow{\varphi} - \mathrm{Id})$. Or $\dim(\mathrm{Ker}(\overrightarrow{\varphi} - \mathrm{Id})) = n - k$: en effet, d'après le théorème de Cartan-Dieudonné vectoriel, $\dim(\mathrm{Ker}(\overrightarrow{\varphi} - \mathrm{Id})) = n - p(\overrightarrow{\varphi})$ et on a ici $k = p(\overrightarrow{\varphi})$. Ainsi, $\dim(F) = n - k$, et donc les formes linéaires ψ_i sont indépendantes. On complète les ψ_i en une base $(\psi_1, \ldots, \psi_n)$ de E^*, de base anté-duale $(e_1, \ldots, e_n)$. Soit $M \in \mathcal{E}$ tel que $\overrightarrow{OM} = \beta_1 e_1 + \ldots \beta_n e_n$. Alors $\psi_i(\overrightarrow{OM}) = \beta_i$, pour tout i. On vient donc de prouver que $\mathcal{F}$ n'est pas vide. Mais tout point de $\mathcal{F}$ est point fixe de φ et φ n'a pas de point fixe : c'est absurde.

Définition. On note $\mathcal{O}(E)$ l'ensemble des isométries vectorielles d'un espace vectoriel euclidien E. Il s'agit de l'ensemble des applications linéaires conservant le produit scalaire.

Définition. On appelle *réflexion* une symétrie orthogonale par rapport à un hyperplan.

Proposition. *Soient a et b deux vecteurs distincts d'un espace euclidien E tels que $\|a\| = \|b\|$. Alors il existe une unique réflexion échangeant a et b.*

Démonstration. Unicité : si σ est une réflexion par rapport à un hyperplan H et échangeant a et b, alors $\sigma(a-b) = b-a$ et donc $H = \text{Vect}(b-a)^\perp$.

Existence : soient $H = \text{Vect}(b-a)^\perp$ et σ la réflexion par rapport à H. Alors $\sigma(a-b) = b-a$. D'autre part, on a

$$(a+b, a-b) = \|a\|^2 - \|b\|^2 = 0$$

par l'hypothèse $\|a\| = \|b\|$. Donc $a+b \in H$ et $\sigma(a+b) = a+b$. On en déduit $\sigma(a) = \frac{1}{2}\sigma(a+b) + \frac{1}{2}\sigma(a-b) = b$ et $\sigma(b) = a$. La réflexion σ est solution. □

Théorème. *Pour toute isométrie f de $\mathcal{E}$, il existe un unique couple (t,h) avec t translation de $\mathcal{E}$ et h isométrie de $\mathcal{E}$ tel que h a un point fixe et $f = h \circ t = t \circ h$.*

Proposition. *Soit t une translation de $\mathcal{E}$ de vecteur $\overrightarrow{u}$. Si $(\mathcal{F}, F)$ est un sous-espace de $\mathcal{E}$ tel que $\overrightarrow{u} \in F^\perp$ et si $\mathcal{G} = \mathcal{F} + \frac{1}{2}\overrightarrow{u}$, alors $t = \sigma_\mathcal{G} \circ \sigma_\mathcal{F}$. En particulier, toute translation est composée de deux réflexions.*

Références :

- Cognet - *Algèbre bilinéaire* - page 96.
- Tauvel - *Cours de géométrie* - page 97 (pour Cartan-Dieudonné affine).

2.18 Théorème de Chevalley-Warning

Théorème. *Soient $d, n \in \mathbb{N}^*$ et p un nombre premier. On note $q = p^d$, et pour $P_1, \dots, P_r \in \mathbb{F}_q[X_1, \dots, X_n]$,*

$$Z(P_1, \dots, P_r) = \left\{x \in \mathbb{F}_q^n \ / \ P_i(x) = 0, \forall i \in [\![1, r]\!]\right\}.$$

Alors si $\sum_{i=1}^r \deg(P_i) < n$, on a :

$$\mathrm{Card}(Z(P_1, \dots, P_r)) \equiv 0 \ [p].$$

Démonstration. Soient $P_1, \dots, P_r \in \mathbb{F}_q[X_1, \dots, X_n]$ tels que $\sum_{i=1}^r \deg(P_i) < n$. On introduit le polynôme :

$$S(X_1, \dots, X_n) = \prod_{i=1}^{r}(1 - P_i(X_1, \dots, X_n)^{q-1}) \in \mathbb{F}_q[X_1, \dots, X_n].$$

Si $x \in Z(P_1, \dots, P_r)$, alors $P_i(x) = 0$ pour tout i. Donc $S(x) = 1$.
Si $x \notin Z(P_1, \dots, P_r)$, alors il existe i tel que $P_i(x) \neq 0$. En particulier, $P_i(x) \in \mathbb{F}_q^*$, donc $P_i(x)^{q-1} = 1$ et on en déduit $S(x) = 0$. Au final :

$$S(x) = \begin{cases} 1 \text{ si } x \in Z(P_1, \dots, P_r) \\ 0 \text{ sinon,} \end{cases}$$

d'où :

$$\sum_{x \in \mathbb{F}_q^n} S(x) = \mathrm{Card}(Z(P_1, \dots, P_r))1_{\mathbb{F}_q}.$$

D'autre part, $S \in \mathbb{F}_q[X_1, \dots, X_n]$ s'écrit sous la forme

$$S = \sum_{\alpha=(\alpha_1,\dots,\alpha_n) \in \mathbb{N}^n} c_\alpha X_1^{\alpha_1} \dots X_n^{\alpha_n},$$

où seulement un nombre fini de $c_\alpha \in \mathbb{F}_q$ sont non nuls. On obtient :

$$\begin{aligned} \sum_{x \in \mathbb{F}_q^n} S(x) &= \sum_{x=(x_1,\dots,x_n) \in \mathbb{F}_q^n} \ \sum_{\alpha=(\alpha_1,\dots,\alpha_n) \in \mathbb{N}^n} c_\alpha x_1^{\alpha_1} \dots x_n^{\alpha_n} \\ &= \sum_{\alpha=(\alpha_1,\dots,\alpha_n) \in \mathbb{N}^n} c_\alpha \sum_{x_1 \in \mathbb{F}_q} x_1^{\alpha_1} \dots \sum_{x_n \in \mathbb{F}_q} x_n^{\alpha_n} \\ &= \sum_{\alpha=(\alpha_1,\dots,\alpha_n) \in \mathbb{N}^n} c_\alpha \prod_{i=1}^{n} \sum_{x_0 \in \mathbb{F}_q} x_0^{\alpha_i}. \end{aligned}$$

On majore maintenant le degré de S :

$$\deg(S) = \sum_{i=1}^{r} \deg(1 - P_i^{q-1}) \leqslant \sum_{i=1}^{r} \deg(P_i^{q-1}) = (q-1)\sum_{i=1}^{r} \deg(P_i) < n(q-1),$$

cette dernière inégalité stricte étant donnée par l'hypothèse du théorème. Cette majoration implique que pour tout monôme $X_1^{\alpha_1} \dots X_n^{\alpha_n}$ apparaissant dans l'écriture de S, il existe $i \in [\![1, n]\!]$ tel que $\alpha_i < q-1$, car sinon on aurait $\deg(S) \geqslant \deg(X_1^{\alpha_1} \dots X_n^{\alpha_n}) = \sum_{i=1}^n \alpha_i \geqslant n(q-1)$. Pour un tel i, on a $\alpha_i = 0$ ou $q-1 \nmid \alpha_i$, et nous démontrerons juste après que cela implique :

$$\sum_{x_0 \in \mathbb{F}_q} x_0^{\alpha_i} = 0, \text{ et donc } \prod_{i=1}^n \sum_{x_0 \in \mathbb{F}_q} x_0^{\alpha_i} = 0$$

pour tout $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ tel que $c_\alpha \neq 0$. Finalement, il vient :

$$\sum_{x \in \mathbb{F}_q^n} S(x) = 0 = \mathrm{Card}(Z(P_1, \dots, P_r)) 1_{\mathbb{F}_q},$$

et comme $\mathbb{F}_q$ est de caractéristique p,

$$\mathrm{Card}(Z(P_1, \dots, P_r)) \equiv 0 \ [p].$$

□

Démontrons maintenant le point manquant :

Proposition. *Soient $d \in \mathbb{N}^*$ et p un nombre premier. On note $q = p^d$. Alors :*

$$\sum_{x \in \mathbb{F}_q} x^m = \begin{cases} 0 \text{ si } m = 0 \text{ ou si } q-1 \nmid m \\ -1 \text{ sinon.} \end{cases}$$

Démonstration. Si $m = 0$, alors pour tout $x \in \mathbb{F}_q$, $x^0 = 1$, d'où le résultat : $\sum_{x \in \mathbb{F}_q} x^0 = q = 0$ dans $\mathbb{F}_q$.

Supposons maintenant $q-1 \nmid m$. Le groupe $\mathbb{F}_q^*$ est un groupe cyclique dont on note g un générateur. Comme l'application

$$\begin{array}{rccc} f : & \mathbb{F}_q & \longrightarrow & \mathbb{F}_q \\ & x & \longmapsto & gx \end{array}$$

est bijective, il vient :

$$\sum_{x \in \mathbb{F}_q} x^m = \sum_{x \in \mathbb{F}_q} (gx)^m = g^m \sum_{x \in \mathbb{F}_q} x^m,$$

soit encore

$$(1 - g^m) \sum_{x \in \mathbb{F}_q} x^m = 0.$$

Mais g étant un générateur de $\mathbb{F}_q^*$, g est d'ordre $q-1$, avec $q-1 \nmid m$, donc $g^m \neq 1$. Le corps $\mathbb{F}_q$ étant intègre, on conclut finalement :

$$\sum_{x \in \mathbb{F}_q} x^m = 0.$$

Si on suppose maintenant $q-1 \,|\, m$, il existe $k \in \mathbb{Z}$ tel que $m = (q-1)k$. Or pour tout $x \in \mathbb{F}_q^*$, $x^{q-1} = 1$. Donc $x^m = 1$. Comme $0^m = 0$, on obtient :

$$\sum_{x \in \mathbb{F}_q} x^m = \sum_{x \in \mathbb{F}_q^*} x^m = \sum_{x \in \mathbb{F}_q^*} 1 = q - 1 = -1$$

dans $\mathbb{F}_q$ de caractéristique p où on a $q = 0$. □

Si on a le temps, on peut donner le corollaire et l'application ci-après :

Corollaire. *Soient $d, n \in \mathbb{N}^*$ et p un nombre premier. On note $q = p^d$. Soient $P_1, \ldots, P_r \in \mathbb{F}_q[X_1, \ldots, X_n]$ tels que :*

(i) Chaque P_i est sans terme constant.

(ii) $\sum_{i=1}^{r} \deg(P_i) < n$.

Alors les P_i admettent un zéro commun non trivial.

Démonstration. Comme P_i est sans terme constant, $P_i(0, \ldots, 0) = 0$, et donc $0_{\mathbb{F}_q^n}$ est un zéro commun à tous les P_i. On en déduit que

$$\mathrm{Card}(Z(P_1, \ldots, P_r)) \geqslant 1.$$

D'après le théorème de Chevalley-Warning,

$$\mathrm{Card}(Z(P_1, \ldots, P_r)) \equiv 0 \ [p].$$

Donc $\mathrm{Card}(Z(P_1, \ldots, P_r)) \geqslant p \geqslant 2$ et il existe un autre zéro que $0_{\mathbb{F}_q^n}$ commun à tous les P_i. □

Application. *Toute forme quadratique sur un $\mathbb{F}_q$-espace vectoriel de dimension $\geqslant 3$ est isotrope, i.e. admet un zéro non trivial.*

Démonstration. Soit Q une forme quadratique sur E un $\mathbb{F}_q$-espace vectoriel de dimension $n \geqslant 3$, *i.e.* $Q \in \mathbb{F}_q[X_1, \ldots, X_n]$. Par définition d'une forme quadratique, Q est un polynôme homogène de degré 2. On applique le corollaire précédent avec $r = 1$:

(i) Q est sans terme constant car homogène de degré 2.

(ii) $\deg(Q) = 2 < n$.

Donc Q admet un zéro non trivial. □

Références :

– Jean-Pierre Serre - *Cours d'arithmétique* - Page 13.

2.19 Théorème de Frobenius-Zolotarev

Théorème. *Soient p un nombre premier $\geqslant 3$ et V un espace vectoriel de dimension finie n sur $\mathbb{F}_p$. Soit $u \in \mathrm{GL}(V)$. Si $\varepsilon(u)$ désigne la signature de u en tant qu'élément de S_{p^n}, alors*

$$\varepsilon(u) = \left(\frac{\det(u)}{p}\right),$$

où $\left(\frac{a}{p}\right)$ désigne le symbole de Legendre égal à 1 si a est un carré dans $\mathbb{F}_p$, -1 sinon.

Démonstration. On regarde $\mathrm{GL}(V)$ comme sous-groupe de S_{p^n} : la signature ε induit un morphisme de groupes de $\mathrm{GL}(V)$ sur $\{-1,1\}$ par restriction. On note $D(\mathrm{GL}(V))$ le groupe dérivé de $\mathrm{GL}(V)$ engendré par les commutateurs $[u,v] = uvu^{-1}v^{-1}$. Alors comme $\{-1,1\}$ est un groupe commutatif, $D(\mathrm{GL}(V)) \subset \mathrm{Ker}(\varepsilon)$ et donc le théorème de factorisation nous permet d'écrire $\varepsilon = \overline{\varepsilon} \circ \Pi$ avec

$$\overline{\varepsilon} : \mathrm{GL}(V)/D(\mathrm{GL}(V)) \to \{-1,1\}$$

et Π la projection canonique de $\mathrm{GL}(V)$ sur $\mathrm{GL}(V)/D(\mathrm{GL}(V))$.

Or $D(\mathrm{GL}(V)) = \mathrm{SL}(V)$. En effet, on a clairement $D(\mathrm{GL}(V)) \subset \mathrm{SL}(V)$. Pour l'inclusion inverse, on sait que $\mathrm{SL}(V)$ est engendré par les transvections (on le montre par l'algorithme du pivot de Gauss), donc il suffit de montrer que toute transvection est un commutateur. Or toutes les transvections de $\mathrm{GL}(V)$ sont conjuguées : dans une bonne base, la matrice d'une transvection est :

$$\begin{pmatrix} 1 & \dots & \dots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & 1 & 1 \\ 0 & \dots & \dots & 1 \end{pmatrix},$$

en échangeant correctement les vecteurs de la base et en leur faisant subir des homothéties. Soit $u \in \mathrm{GL}(V)$ une transvection, $u = I_n + \lambda E_{ij}$ avec $i \neq j$. On a :

$$u^2 = (I_n + \lambda E_{ij})^2 = I_n + 2\lambda E_{ij} + \lambda^2 \underbrace{E_{ij}^2}_{=\delta_{ji}E_{ij}=0} = I_n + 2\lambda E_{ij}.$$

Comme $p \geqslant 3$, la caractéristique de $\mathbb{F}_p$ est différente de 2, donc u^2 est encore une transvection. Donc il existe $v \in \mathrm{GL}(V)$ tel que $u^2 = vuv^{-1}$, soit encore $u = vuv^{-1}u^{-1}$, *i.e.* u est un commutateur.

Par ailleurs, le morphisme surjectif $\det : \mathrm{GL}(V) \to \mathbb{F}_p^*$ a pour noyau $\mathrm{SL}(V)$ par définition. Par le théorème de factorisation, il existe $\overline{\det} : \mathrm{GL}(V)/\mathrm{SL}(V) \to \mathbb{F}_p^*$ isomorphisme tel que $\det = \overline{\det} \circ \Pi$.

On obtient finalement un morphisme $\delta = \overline{\varepsilon} \circ \overline{\det}^{-1} : \mathbb{F}_p^* \to \{-1, 1\}$ tel que $\delta \circ \det = \varepsilon$. Le but est de montrer que δ est le symbole de Legendre.

On note d'abord que $\left(\frac{\cdot}{p}\right) : \mathbb{F}_p^* \to \{-1, 1\}$ est un morphisme (il est multiplicatif) et qu'il est non trivial. En effet, il vaut -1 sur les générateurs de $\mathbb{F}_p^*$ (on sait que $\mathbb{F}_p^*$ est cyclique) : si g est un générateur de $\mathbb{F}_p^*$, g est d'ordre $p-1$, et si g était un carré, on aurait $g = h^2$, donc $g^{\frac{p-1}{2}} = h^{p-1} = 1$, ce qui contredirait la définition de l'ordre.

Mais comme $\mathbb{F}_p^*$ est cyclique, tout morphisme de groupes de $\mathbb{F}_p^*$ sur $\{-1, 1\}$ est déterminé par l'image d'un générateur : il n'y en a donc que deux, qui sont le morphisme trivial et le symbole de Legendre.

Il reste à montrer que δ n'est pas le morphisme trivial. Comme V est un $\mathbb{F}_p$-espace vectoriel de dimension n, on a $V \cong \mathbb{F}_p^n \cong \mathbb{F}_{p^n} = \mathbb{F}_q$. On prend g un générateur de $\mathbb{F}_q^*$ et on considère

$$\begin{array}{rccc} \varphi : & \mathbb{F}_q & \longrightarrow & \mathbb{F}_q \\ & x & \longmapsto & gx \end{array}$$

qui est une application $\mathbb{F}_p$-linéaire. Donc $\varphi \in \mathrm{GL}(V)$. De plus, φ est égal en tant que permutation au cycle $(1, g, g^2, \ldots, g^{q-2})$ qui est de longueur paire $q-1$, donc $\varepsilon(\varphi) = -1$. Comme $\varepsilon = \delta \circ \det$, on a $\delta(\det(\varphi)) = -1$ et donc δ n'est pas le morphisme trivial.

Finalement, δ ne peut être que le symbole de Legendre, d'où le résultat. □

Comme V est un $\mathbb{F}_p$-espace vectoriel de dimension n sur $\mathbb{F}_p$, V est un espace de cardinal p^n. Donc tout $u \in \mathrm{GL}(V)$ est une bijection de V sur V qui peut être vue comme une permutation de S_{p^n}. Donc $\mathrm{GL}(V)$ est isomorphe à un sous-groupe de S_{p^n}. Mais on n'a pas $\mathrm{GL}(V) \cong S_{p^n}$ pour des raisons de cardinal : $\mathrm{GL}(V) \cong \mathrm{GL}_n(\mathbb{F}_p) \subset \mathcal{M}_n(\mathbb{F}_p)$, donc $\mathrm{Card}(\mathrm{GL}(V)) \leqslant \mathrm{Card}(\mathcal{M}_n(\mathbb{F}_p)) = p^{n^2}$, et comme $\mathrm{Card}(S_{p^n}) = (p^n)!$, on aurait $(p^n)! \leqslant p^{n^2}$, ce qui n'est pas le cas. En effet :

$$\begin{aligned} (p^n)! &= 1 \ldots p(p+1) \ldots p^2(p^2+1) \ldots p^n \\ &= kp(p+1)p^2(p^2+1)p^{n-1}(p^{n-1}+1)p^n \\ &\geqslant kpppp^2p^2 \ldots p^{n-1}p^{n-1}p^n \\ &= kp^{\frac{n(n+1)}{2}}p^{\frac{n(n-1)}{2}} = kp^{n^2}, \end{aligned}$$

et on voit bien que $k > 1$.

Proposition. *Le groupe dérivé de G est un sous-groupe distingué de G.*

Démonstration. Il s'agit de montrer que si $g \in G$ et $[x,y] = xyx^{-1}y^{-1} \in D(G)$, alors $g[x,y]g^{-1} \in D(G)$. On a le résultat en écrivant :

$$g[x,y]g^{-1} = (gxg^{-1})(gyg^{-1})(gx^{-1}g^{-1})(gy^{-1}g^{-1}) = [gxg^{-1}, gyg^{-1}].$$

Cela revient aussi à dire que $D(G)$ est stable par les automorphismes intérieurs, *i.e.* les $i_g : x \mapsto gxg^{-1}$. □

Proposition. *Le symbole de Legendre est un morphisme.*

Démonstration. Dans le document *Théorème des deux carrés*, on montre que

$$\mathbb{F}_q^{*2} = \left\{ x \in \mathbb{F}_q \,/\, x^{\frac{q-1}{2}} = 1 \right\}.$$

On prend ici $q = p$. Comme $\left(x^{\frac{p-1}{2}}\right)^2 = x^{p-1} = 1$ et que le polynôme $X^2 - 1$ admet au plus deux racines (qui sont en fait 1 et -1), alors pour tout $x \in \mathbb{F}_p^*$, on a soit $x^{\frac{p-1}{2}} = 1$ (lorsque x est un carré), soit $x^{\frac{p-1}{2}} = -1$ (lorsque x n'est pas un carré). Donc

$$\left(\frac{x}{p}\right) = x^{\frac{p-1}{2}}.$$

Ainsi il s'agit bien d'un morphisme :

$$\left(\frac{ab}{p}\right) = (ab)^{\frac{p-1}{2}} = a^{\frac{p-1}{2}} b^{\frac{p-1}{2}} = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

□

Références :

– Beck, Malick, Peyré - *Objectif Agrégation* - Page 251.

2.20 Théorème de Krein-Milman

Proposition. *Soit E un espace euclidien et C un convexe fermé non vide de E. On note p la projection sur C (qui est continue, voir document "Projection sur un convexe fermé"). Soit c un point de la frontière de C. Alors il existe un hyperplan d'appui de C en c.*

Démonstration. Le point c est adhérent à $E\backslash C$, donc il existe une suite (x_k) de $E\backslash C$ qui converge vers c. Par continuité de p, la suite $p(x_k)$ converge vers $p(c) = c$. On sait de plus qu'en chacun des points $p(x_k)$, on dispose d'un hyperplan d'appui dont un vecteur normal unitaire est

$$u_k = \frac{x_k - p(x_k)}{\|x_k - p(x_k)\|}.$$

La suite (u_k) est dans la sphère unité de E qui est compacte, on peut donc en extraire une sous-suite convergente, qu'on appellera encore (u_k). On note u sa limite et on considère l'hyperplan affine

$$H = \{x \in E \ / \ (u, x) = (u, c)\}.$$

Cet hyperplan contient c et on va montrer qu'il s'agit d'un hyperplan d'appui en c. Soit $z \in C$. Pour tout $k \in \mathbb{N}$, on a

$$(u_k, z - p(x_k)) \leqslant 0.$$

En passant à la limite quand $k \to \infty$, on obtient $(u, z - p(c)) \leqslant 0$, *i.e.* $(u, z) \leqslant (u, c)$. Ceci prouve que C est inclu dans un des demi-espaces fermés délimité par H et H est donc un hyperplan d'appui. □

Théorème (Krein-Milman)**.** *Soit E un espace euclidien et K un convexe compact non vide de E. Alors K est égal à l'enveloppe convexe de l'ensemble de ses points extrémaux.*

Démonstration. On commence par montrer un résultat préliminaire : soit $a \in K$ tel que a appartienne à un hyperplan d'appui H de K. Alors a est un point extrémal de K si et seulement si a est un point extrémal du convexe compact $K \cap H$.

D'abord, $K \cap H$ est non vide (il contient a), convexe (intersection de deux convexes), et compact (intersection du compact K et du fermé H).

On suppose que a est un point extrémal de K. Alors $K\backslash\{a\}$ est convexe, donc $K\backslash\{a\} \cap H = (K \cap H)\backslash\{a\}$ aussi comme intersection de deux convexes. Par conséquent, a est un point extrémal de $K \cap H$.

Réciproquement, on suppose que a est un point extrémal de $K \cap H$. Soient $u, v \in K$ tels que $a = \frac{u+v}{2}$. On veut montrer que $u = v = a$. Soit φ une forme

linéaire sur E telle que $H = \{x \in E \,/\, \varphi(x) = \lambda\}$. Comme K est contenu dans un demi-espace fermé délimité par H, on a par exemple $\varphi(u) \leqslant \lambda$ et $\varphi(v) \leqslant \lambda$. Mais $\varphi(a) = \frac{\varphi(u)+\varphi(v)}{2} = \lambda$, donc nécessairement $\varphi(u) = \varphi(v) = \lambda$, ce qui signifie que u et v sont dans H, donc dans $K \cap H$. Par conséquent, comme a est un point extrémal de $K \cap H$, on a $u = v = a$.

On va montrer que K est l'enveloppe convexe de l'ensemble de ses points extrémaux par récurrence sur la dimension p du sous-espace affine engendré par K. Si $p = 0$, alors K est un singleton et le résultat est immédiat. On suppose le résultat vrai jusqu'au rang $p-1$ et on se donne un convexe compact K qui engendre un sous-espace affine de dimension p. Quitte à translater K et à remplacer E par un de ses sous-espaces vectoriels, on peut supposer que $\dim(E) = p$ (on regarde K comme convexe compact de l'espace vectoriel engendré par ses éléments). Soit $c \in K$. On cherche à écrire c comme combinaison convexe de points extrémaux de K.

Distinguons deux cas :

(i) Si c appartient à la frontière de K, par la proposition précédente, il existe un hyperplan d'appui H de K en c. On pose $K' = K \cap H$. Il s'agit d'un convexe compact qui engendre un sous-espace affine de dimension $\leqslant p-1$ (car $\dim(H) = p-1$ puisqu'on a supposé que $\dim(E) = p$) auquel on peut appliquer l'hypothèse de récurrence. Comme les points extrémaux de K' sont des points extrémaux de K d'après le premier résultat montré au-dessus, c est bien combinaison convexe de points extrémaux de K.

(ii) Si maintenant c est intérieur à K, on considère une droite quelconque D qui passe par c. Alors $D \cap K$ est une partie convexe de D qui est de plus compact (car D est fermée). Il s'agit donc d'un segment $[a, b]$. Les points a et b sont clairement sur la frontière de K et on peut appliquer le cas précédent. Par associativité, c est barycentre à coefficients positifs de points extrémaux de K.

□

Soient E un espace euclidien et K une partie non vide, convexe et fermée de E. D'après le théorème de projection sur un convexe fermé, pour tout $x \in E$, il existe un unique point $p(x) \in K$ tel que $d(x, K) = \|x - p(x)\|$. De plus, on a $(x - p(x), y - p(x)) \leqslant 0$ pour tout $y \in K$, ce qui signifie que si $x \notin K$, l'hyperplan affine H passant par $p(x)$ et de vecteur normal $x - p(x)$ sépare K et x. En effet, on a

$$H = \{y \in E \;/\; (x - p(x), y - p(x)) = 0\}$$

et $K \subset \{y \in E \ / \ (x - p(x), y - p(x)) \leqslant 0\}$ alors que $(x - p(x), x - p(x)) = \|x - p(x)\|^2 > 0$.

Définition. On dit que H est un hyperplan d'appui de K en $p(x)$.

Remarque. Il n'y a pas nécessairement unicité de l'hyperplan d'appui en un point : par exemple dans le plan, lorsque c est l'un des sommets d'un carré.

Définition. Un point x d'un convexe X est dit *extrémal* si $X \backslash \{x\}$ est encore convexe. Cela équivaut à dire que x ne peut pas s'écrire comme milieu de deux points distincts de X.

Exemple. *Les points extrémaux d'un carré plein de $\mathbb{R}^2$ sont ses quatre sommets.*

Références :

– Francinou, Gianella, Nicolas - *Oraux X-ENS, Analyse 3* - Page 101.

2.21 Théorème de Kronecker

Théorème. *Soit $P \in \mathbb{Z}[X]$ unitaire tel que $P(0) \neq 0$. On suppose que les racines complexes $\alpha_1, \ldots, \alpha_n$ de P sont de module $\leqslant 1$. Alors les α_i sont des racines de l'unité.*

Démonstration. On note $\sigma_1, \ldots, \sigma_n$ les fonctions symétriques élémentaires en les α_i. Par les relations coefficients-racines, sachant que P est unitaire, on a :

$$P = X^n - \sigma_1 X^{n-1} + \cdots + (-1)^n \sigma_n.$$

Les σ_i sont dans $\mathbb{Z}$ car $P \in \mathbb{Z}[X]$. Comme $|\alpha_i| \leqslant 1$ pour tout i, on a :

$$|\sigma_k| = \left| \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \alpha_{i_1} \ldots \alpha_{i_k} \right| \leqslant \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} 1 = \binom{n}{k}.$$

On note Ω_n l'ensemble des polynômes unitaires de $\mathbb{Z}[X]$ de degré n dont les racines complexes sont de module $\leqslant 1$. La majoration précédente montre qu'il n'y a qu'un nombre fini de choix pour les coefficients de P (les σ_i sont entiers et bornés). Donc Ω_n est fini.

On définit maintenant, pour $k \in \mathbb{N}^*$, les polynômes P_k et Q_k par

$$P_k(X) = \prod_{i=1}^{n} (X - \alpha_i^k) \quad \text{et} \quad Q_k(X, Y) = X^k - Y \in \mathbb{Z}[X, Y].$$

On pose $R_k = \mathrm{Res}_X(P, Q_k)$ le résultant des polynômes P et Q_k vus comme polynômes à coefficients dans $\mathbb{Z}[Y]$, de l'indéterminée X. On a donc $R_k \in \mathbb{Z}[Y]$. De plus, la relation entre le résultant et les racines des polynômes nous donne :

$$R_k = \prod_{i=1}^{n} Q_k(\alpha_i) = \prod_{i=1}^{n} (\alpha_i^k - Y) = (-1)^n P_k(Y).$$

Finalement, $P_k(X) = (-1)^n R_k(X) \in \mathbb{Z}[X]$. Comme $|\alpha_i| \leqslant 1$, on a encore $|\alpha_i|^k \leqslant 1$, d'où enfin : $P_k \in \Omega_n$, pour tout $k \in \mathbb{N}^*$.

Comme Ω_n est fini et que chaque polynôme de Ω_n admet au plus n racines complexes distinctes, l'ensemble des racines des polynômes de Ω_n est fini. Donc $\{\alpha_i^k, k \in \mathbb{N}^*\}$ est fini pour tout i. Par conséquent, pour tout i, il existe $p \neq q$, par exemple $p > q$, tels que $\alpha_i^p = \alpha_i^q$. Comme $P(0) \neq 0$, on a $\alpha_i \neq 0$, et donc $\alpha_i^{p-q} = 1$. Les α_i sont donc des racines de l'unité. □

Application. *En gardant les mêmes hypothèses que le théorème précédent, on a de plus : si P est irréductible, alors P est un polynôme cyclotomique.*

Démonstration. D'après le théorème précédent, les racines α_i de P sont des racines de l'unité, disons racines l_i-ième de l'unité. Comme P est irréductible sur $\mathbb{Z}$ et unitaire, il est irréductible sur $\mathbb{Q}$. Et comme $\mathbb{Q}$ est de caractéristique nulle, c'est un corps parfait, donc P étant irréductible, il est scindé à racines simples sur $\mathbb{C}$.

On peut aussi dire que P étant irréductible sur $\mathbb{Q}$, P et P' sont premiers entre eux dans $\mathbb{Q}[X]$. Donc il existe $U, V \in \mathbb{Q}[X]$ tels que $PU + P'V = 1$. Cette relation étant vraie dans $\mathbb{C}[X]$, P et P' sont premiers entre eux dans $\mathbb{C}[X]$, donc n'ont pas de racine commune, et donc P est scindé à racines simples sur $\mathbb{C}$.

Notons $m = \text{PPCM}(l_1, \dots, l_n)$. Alors pour tout i, α_i est racine de $X^m - 1 = \prod_{d|m} \Phi_d(X)$. Comme les racines α_i de P sont simples, P divise $X^m - 1$ dans $\mathbb{C}[X]$: on écrit $X^m - 1 = PQ$ avec $Q \in \mathbb{C}[X]$. En développant, on voit directement que les coefficients de Q sont dans $\mathbb{Q}$. Donc P divise $X^m - 1$ dans $\mathbb{Q}[X]$. Enfin, comme P est unitaire et irréductible sur $\mathbb{Q}$ et que les Φ_d aussi, il existe d diviseur de m tel que $P = \Phi_d$. □

Remarque. Dans la démonstration du théorème, on peut ne pas utiliser le résultant pour montrer que $P_k \in \mathbb{Z}[X]$. On peut voir P_k comme un polynôme symétrique en les α_i, à coefficients dans $\mathbb{Z}[X]$. D'après le théorème de structure des polynômes symétriques, il est égal à un polynôme en les fonctions symétriques élémentaires des α_i, à coefficients dans $\mathbb{Z}[X]$. Mais les fonctions symétriques élémentaires des α_i sont dans $\mathbb{Z}$ d'après les relations coefficients-racines et le fait que P soit unitaire. Donc $P_k \in \mathbb{Z}[X]$.

Théorème (Structure des polynômes symétriques)**.** *Soient A un anneau et $P \in A[X_1, \dots, X_n]$ symétrique. Alors il existe un unique polynôme $Q \in A[X_1, \dots, X_n]$ tel que*

$$P(X_1, \dots, X_n) = Q(\Sigma_1, \dots, \Sigma_n),$$

où les Σ_i sont les polynômes symétriques élémentaires de $A[X_1, \dots, X_n]$.

Théorème. *Soit $\overline{K}$ une clotûre algébrique d'un corps K. On note $x_1, \dots, x_m$ les racines de $P = \sum_{i=0}^{m} a_i X^i \in K[X]$ et $y_1, \dots, y_n$ celles de $Q = \sum_{j=0}^{n} b_j X^j \in K[X]$. On a alors :*

$$\text{Res}_X(P, Q) = a_m^n b_n^m \prod_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} (x_i - y_j) = a_m^n \prod_{i=1}^{m} Q(x_i) = (-1)^{mn} b_n^m \prod_{j=1}^{n} P(y_j).$$

Démonstration. On a $P = a_m(X - x_1)\dots(X - x_m)$ et $Q = b_n(X - y_1)\dots(X - y_n)$ qu'on regarde comme polynômes de $A = \mathbb{Z}[a_m, b_n, x_1, \dots, x_m, y_1, \dots, y_n, X]$. Il s'agit de montrer que $x_i - y_j$ divise $\mathrm{Res}_X(P, Q) \in \mathbb{Z}[a_m, b_n, x_1, \dots, x_m, y_1, \dots, y_n]$. Pour se fixer les idées, on prend $i = 1$ et $j = 1$, et on effectue la division euclidienne par $x_1 - y_1$, par rapport à l'indéterminée x_1 :

$$\mathrm{Res}_X(P, Q) = (x_1 - y_1)S + T$$

avec $T \in \mathbb{Z}[a_m, b_n, x_2, \dots, x_m, y_1, \dots, y_n]$. On spécialise les indéterminées en les choisissant dans $\mathbb{C}^{2+m+n}$ telles que $x_1 = y_1$. Alors P et Q ont une racine commune, ils ne sont donc pas premiers entre eux, donc leur résultant est nul. On obtient alors $T(a_m, b_n, x_2, \dots, x_m, y_1, \dots, y_n) = 0$ pour tout $a_m, b_n, x_2, \dots, y_n \in \mathbb{C}$. Donc T est le polynôme nul et on a le résultat.

Les $x_i - y_j$ sont deux à deux premiers entre eux dans A et ils divisent tous $R = \mathrm{Res}_X(P, Q)$. Remarque : on les regarde en tant qu'indéterminées, il n'y a donc pas de problème selon qu'il existe des racines multiples ou des racines nulles. Par factorialité de A (car $\mathbb{Z}$ est factoriel), le produit $\prod_{i,j}(x_i - y_j)$ divise R :

$$R = F \prod_{i,j}(x_i - y_j).$$

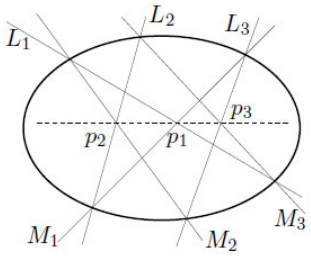
Le coefficient constant de Q (vu dans $K[X]$) vaut $(-1)^n b_n y_1 \dots y_n$ et ses autres coefficients sont de degré strictement plus petit en les y_i. On se rend donc compte en développant le déterminant de la matrice de Sylvester que la partie homogène de plus haut degré en les y_i de R est donnée par la diagonale, que l'on trouve égale à $a_m^n b_0^m = a_m^n((-1)^n b_n y_1 \dots y_n)^m$. Comme la partie homogène de plus haut degré en les y_i de $\prod_{i,j}(x_i - y_j)$ vaut $(-1)^{nm}(y_1 \dots y_n)^m$, on en déduit $F = a_m^n b_n^m$ (une considération sur les degrés montre que F ne dépend ni des x_i, ni des y_j), d'où le résultat. □

Références :

- Szpirglas - *Mathématiques L3 Algèbre* - Page 573.
- Francinou, Gianella, Nicolas - *Oraux X-ENS, Algèbre 1* - Page 213.

2.22 Théorème de Pascal

Théorème. *Soit C une conique propre. On fixe six droites distinctes L_1, L_2, L_3, M_1, M_2, M_3 telles que pour tout $i, j \in \{1,2,3\}$, $i \neq j$, les droites L_i et M_j se coupent en un point de C. Alors les trois points $p_i = L_i \cap M_i$, $i \in \{1,2,3\}$ sont alignés.*



Démonstration. Fixons un repère du plan : on notera (x, y) les coordonnées d'un point dans ce repère. Pour tout $i \in \{1,2,3\}$, on note $l_i = 0$ (resp. $m_i = 0$) l'équation de la droite L_i (resp. M_i) : les l_i et m_i sont des polynômes de degré 1 en x, y. On introduit le polynôme f_λ suivant, dépendant d'un paramètre $\lambda \in \mathbb{R}$:

$$f_\lambda = l_1 l_2 l_3 + \lambda m_1 m_2 m_3.$$

Montrons que f_λ est de degré exactement 3 en x, y. Supposons que son degré soit < 3. On écrit $l_i = a_i x + b_i y + c_i$ et $m_i = \alpha_i x + \beta_i y + \gamma_i$. Alors la partie homogène de degré 3 de f_λ est

$$(a_1 x + b_1 y)(a_2 x + b_2 y)(a_3 x + b_3 y) + \lambda(\alpha_1 x + \beta_1 y)(\alpha_2 x + \beta_2 y)(\alpha_3 x + \beta_3 y).$$

Comme elle est nulle, que $\mathbb{R}[x, y]$ est factoriel, et que les $a_i x + b_i y$ et $\alpha_i x + \beta_i y$ sont irréductibles dans $\mathbb{R}[x, y]$, on obtient que pour tout i, il existe j tel que l_i et m_j soient égaux à une constante multiplicative près. Par conséquent, L_i serait parallèle à M_j, ce qui est contraire aux hypothèses (les deux droites se coupent en un point appartenant à C, en particulier pas à l'infini).

Finalement, l'équation $f_\lambda(x, y) = 0$ définit une cubique F_λ. De façon évidente, les six points d'intersection $L_i \cap M_j$, $i, j \in \{1,2,3\}$, $i \neq j$, appartiennent à F_λ.

Soit $p \in C$ distinct des six points $L_i \cap M_j$, $i \neq j$: p n'appartient donc à aucune des droites M_i, car une droite coupe une conique en au plus deux points (une conique a une équation de degré 2 en x, y, donc si on y injecte $y = ax + b$, on obtient que x est racine d'un polynôme de degré 2). Par conséquent, le polynôme $m_1 m_2 m_3$ ne s'annule pas en p. On peut alors poser

$$\lambda = -\frac{(l_1 l_2 l_3)(p)}{(m_1 m_2 m_3)(p)}.$$

Avec cette valeur de λ, on a $p \in C \cap F_\lambda$: l'intersection $C \cap F_\lambda$ contient donc au moins 7 points distincts. D'après la forme faible du théorème de Bézout que l'on démontrera ensuite, cela implique que la conique C est incluse dans F_λ. On note $c = 0$ l'équation de la conique C. Alors c divise f_λ, et pour des raisons de degré, il existe un polynôme l de degré 1 en x, y tel que

$$f_\lambda = cl.$$

Or, pour $i \in \{1, 2, 3\}$, le point p_i appartient à F_λ mais pas à C (car l'intersection $L_i \cap C$ contient au plus deux points, qui sont $L_i \cap M_j$ pour $j \neq i$). Par conséquent, pour tout i, on a $l(p_i) = 0$, ce qui exprime que les p_i sont alignés sur la droite d'équation $l = 0$. □

Théorème (Forme faible du théorème de Bézout)**.** *Soit C une conique propre et F une cubique propre (ensemble des solutions dans $\mathbb{R}^2$ d'un polynôme de degré 3). Alors soit C est contenue dans F, soit $C \cap F$ contient au plus 6 points distincts.*

Démonstration. Supposons que C n'est pas contenue dans F. Comme la conique est propre, on peut choisir un repère dans lequel son équation est

$$y^2 = f(x)$$

avec f polynôme de degré au plus 2. Dans ce repère, F a une équation de la forme $e(x, y) = 0$ avec e polynôme de degré 3 en x, y. Ainsi

$$(x, y) \in C \cap F \iff \begin{cases} y^2 = f(x) \\ e(x, y) = 0 \end{cases} \iff (*) \begin{cases} y^2 = f(x) \\ g(x)y + h(x) = 0, \end{cases}$$

en ayant remplacé les y^2 par $f(x)$ dans l'équation $e(x, y) = 0$, g et h sont alors des polynômes tels que $\deg(g) \leqslant 2$ et $\deg(h) \leqslant 3$. Notre but est de montrer qu'il y a au plus six solutions au système $(*)$ d'inconnues x, y.

Il va falloir distinguer les solutions (x, y) telles que $g(x) = 0$ (qui impliquent $h(x) = 0$) et celles telles que $g(x) \neq 0$. Soient $x_1, \ldots, x_k$ les racines communes à g et h, répétées avec multiplicité. Comme g est de degré au plus 2, on a $k \leqslant 2$. On factorise g et h :

$$g(x) = \prod_{i=1}^{k} (x - x_i)\tilde{g}(x), \quad h(x) = \prod_{i=1}^{k} (x - x_i)\tilde{h}(x),$$

avec $\tilde{g}$ et $\tilde{h}$ deux polynômes sans racines communes, $\deg(\tilde{g}) \leqslant 2-k$ et $\deg(\tilde{h}) \leqslant 3-k$. Comme l'équation $y^2 = f(x)$ est de degré 2 en y, pour chaque x_i, il existe au plus deux valeurs de y telles que $y^2 = f(x_i)$. Donc le système $(*)$ possède au plus $2k$ solutions dont l'abscisse x vérifie $g(x) = 0$.

Les solutions de $(*)$ dont l'abscisse x est telle que $g(x) \neq 0$ sont aussi solutions de

$$(**) \begin{cases} y^2 = f(x) \\ \tilde{g}(x)y + \tilde{h}(x) = 0 \\ \tilde{g}(x) \neq 0. \end{cases}$$

Si (x, y) est solution de $(**)$, alors x est solution de

$$\tilde{h}(x)^2 = \tilde{g}(x)^2 f(x),$$

équation polynomiale de degré $d \leqslant \max(2\deg(\tilde{h}), 2\deg(\tilde{g}) + \deg(f)) \leqslant \max(2(3-k), 2(2-k)+2) = 6-2k$. Si le polynôme $\tilde{h}^2 - \tilde{g}^2 f$ n'est pas nul, $(**)$ admet au plus $6-2k$ solutions. Au final, le système $(*)$ admet au plus $6-2k+2k = 6$ solutions.

Montrons donc que le polynôme $\tilde{h}^2 - \tilde{g}^2 f$ n'est pas nul. Supposons $\tilde{h}^2 = \tilde{g}^2 f$. Quitte à simplifier l'équation par les facteurs irréductibles de degré 2 communs à $\tilde{g}$ et $\tilde{h}$, on peut supposer $\tilde{g}$ et $\tilde{h}$ premiers entre eux. Alors $\tilde{h}^2 = \tilde{g}^2 f$ implique que $\tilde{g}$ est constant. Si $\tilde{g}$ était la constante nulle, $\tilde{h}$ aussi, et g et h seraient nuls également, ce qui est absurde car $g(x)y + h(x) = 0$ définit la cubique F de degré 3. Donc

$$\tilde{f} = \frac{1}{\tilde{g}^2}\tilde{h}^2.$$

Comme $\deg(f) \leqslant 2$, on obtient $\deg(\tilde{h}) \leqslant 1$. Par conséquent, la conique C a pour équation

$$y^2 = f(x) = \frac{1}{\tilde{g}^2}\tilde{h}^2(x),$$

soit encore $y = \pm\frac{1}{\tilde{g}}\tilde{h}(x)$. Donc C est une conique dégénérée (la réunion de deux droites), ce qui est absurde car on l'a supposé propre. □

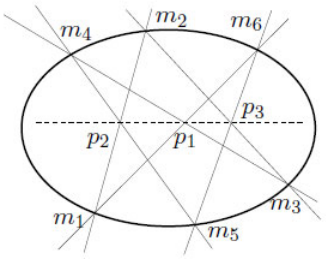
Remarque. 6 vient de $2 * 3$, le produit des degrés des polynômes qui définissent les courbes.

Le théorème de Pascal entraîne sa réciproque :

Proposition. *Soient six points du plan $m_1, m_2, m_3, m_4, m_5, m_6$. On note*

$$p_2 = m_1m_2 \cap m_4m_5, \;\; p_3 = m_2m_3 \cap m_5m_6, \;\; p_1 = m_3m_4 \cap m_1m_6.$$

Alors si p_1, p_2 et p_3 sont alignés, il existe une conique qui contient tous les m_i, $i \in \{1, \ldots, 6\}$.



Démonstration. On sait qu'il existe une unique conique passant par 5 points donnés. Notons C la conique passant par m_1, m_2, m_3, m_4, m_5. Soit m_6' le point d'intersection de C avec m_5p_3 autre que m_5, et soit $p_1' = m_3m_4 \cap m_6'm_1$.

Par le théorème de Pascal appliqué à $m_1, m_2, m_3, m_4, m_5, m_6'$ (p_3 est l'intersection de m_2m_3 avec m_5m_6', etc...), on a $p_1' \in p_2p_3$. On en déduit que $p_1' = p_2p_3 \cap m_3m_4$, et comme p_1 est aussi à cette unique intersection, $p_1' = p_1$. Mais

$$m_6 = m_1p_1 \cap m_5p_3 \quad \text{et} \quad m_6' = m_1p_1' \cap m_5p_3,$$

ce qui prouve que $m_6' = m_6$, et donc que $m_6 \in C$. □

Théorème (Pappus)**.** *Si C est la réunion de deux droites, le théorème de Pascal est encore vrai.*

Références :

- Un document PDF de Jérôme Germoni.
- Shafarevich - *Basic algebraic geometry* - Page 20.

2.23 Théorème de Rothstein-Trager

Théorème. *Soient $P, Q \in \mathbb{Q}[X]$ premiers entre eux tels que $\deg(P) < \deg(Q)$. On suppose de plus Q unitaire et sans facteur carré. Soit K une extension de $\mathbb{Q}$ dans laquelle on ait*

$$\int \frac{P}{Q} = \sum_{i=1}^{n} c_i \ln(P_i)$$

avec les $c_i \in K^*$ et les P_i unitaires non constants (car $\deg(P) < \deg(Q)$) sans facteur carré, et premiers entre eux deux à deux. Quitte à regrouper, on peut supposer les c_i deux à deux distincts. Alors :

(i) Pour tout i, on a $P_i = \mathrm{PGCD}(P - c_i Q', Q)$.

(ii) Les c_i sont exactement les racines du polynôme

$$R(Y) = \mathrm{Res}_X(P - YQ', Q) \in K[Y].$$

Démonstration. L'existence de l'extension K est donnée par le théorème de décomposition en éléments simples.

Pour $i \in [\![1, n]\!]$, on pose $U_i = \prod_{j \neq i} P_j$. On dérive ensuite formellement la relation de l'énoncé :

$$\frac{P}{Q} = \sum_{i=1}^{n} c_i \frac{P_i'}{P_i}.$$

On en déduit en multipliant par $Q \prod_{i=1}^{n} P_i$,

$$P \prod_{i=1}^{n} P_i = Q \sum_{i=1}^{n} c_i P_i' U_i. \quad (*)$$

Par conséquent, dans $K[X]$, on a que Q divise $\prod_{i=1}^{n} P_i$ et P_j divise $Q \sum_{i=1}^{n} c_i P_i' U_i$ pour tout j. Soit $i \in [\![1, n]\!]$. Comme pour tout $j \neq i$, P_i divise U_j, on en déduit que P_i divise $c_i Q P_i' U_i$. Or P_i est sans facteur carré, donc P_i est premier avec P_i', et de plus, P_i est premier avec U_i car il est premier avec tous les P_j pour $j \neq i$. Finalement, on en conclut que P_i divise Q. Ceci étant vrai pour tout i et que les P_i sont premiers entre eux deux à deux, il vient

$$\prod_{i=1}^{n} P_i \text{ divise } Q.$$

Comme les deux polynômes sont unitaires, on a donc montré que

$$Q = \prod_{i=1}^{n} P_i,$$

puis par $(*)$, que

$$P = \sum_{i=1}^{n} c_i P_i' U_i.$$

Soit $i \in [\![1, n]\!]$. Montrons que P_i divise $P - c_i Q'$. En dérivant Q, il vient $Q' = \sum_{j=1}^{n} P_j' U_j$. Donc

$$P - c_i Q' = \sum_{j=1}^{n} c_j P_j' U_j - c_i \sum_{j=1}^{n} P_j' U_j = \sum_{j=1}^{n} (c_j - c_i) P_j' U_j = \sum_{j \neq i} (c_j - c_i) P_j' U_j.$$

Mais pour $j \neq i$, P_i divise U_j. On a donc bien que P_i divise $P - c_i Q'$. Ensuite,

$$\mathrm{PGCD}(P - c_i Q', Q) = \mathrm{PGCD}\left(P - c_i Q', \prod_{j=1}^{n} P_j\right) = \prod_{j=1}^{n} \mathrm{PGCD}\left(P - c_i Q', P_j\right)$$

car les P_j sont premiers entre eux deux à deux. Or si $j \neq i$,

$$\mathrm{PGCD}\left(P - c_i Q', P_j\right) = \mathrm{PGCD}\left(\sum_{k \neq i} (c_k - c_i) P_k' U_k, P_j\right) = \mathrm{PGCD}\left((c_j - c_i) P_j' U_j, P_j\right)$$

car P_j divise tous les autres termes de la somme. Mais comme P_j est premier avec P_j' et premier avec U_j, on obtient

$$\mathrm{PGCD}\left(P - c_i Q', P_j\right) = 1.$$

Finalement,

$$\mathrm{PGCD}(P - c_i Q', Q) = \prod_{j=1}^{n} \mathrm{PGCD}\left(P - c_i Q', P_j\right) = \mathrm{PGCD}\left(P - c_i Q', P_i\right) = P_i$$

car P_i divise $P - c_i Q'$.

Le fait que c_i soit racine du résultant R découle directement du fait que $P - c_i Q'$ et Q ont un PGCD non constant.

Réciproquement, soit c une racine de R dans une extension L de K. Alors $S = \mathrm{PGCD}(P - cQ', Q) \in L[X]$ est non constant. Soit T un facteur irréductible de S dans $L[X]$, on a donc

$$T \text{ divise } P - cQ' \text{ et } T \text{ divise } Q.$$

Comme $Q = \prod_{i=1}^{n} P_i$ et comme les P_i sont deux à deux premiers entre eux (dans $K[X]$, donc dans $L[X]$ par l'identité de Bézout), T doit nécessairement diviser un et un seul des P_i, qu'on note P_{i_0}. Mais

$$P - cQ' = \sum_{j=1}^{n} (c_j - c) P_j' U_j.$$

Comme T divise P_{i_0}, il divise tous les U_i pour $i \neq i_0$. Et puisque T divise $P - cQ'$, on a alors

$$T \text{ divise } (c_{i_0} - c)P'_{i_0}U_{i_0}.$$

Or T ne divise pas U_{i_0}, car sinon il diviserait un P_i pour $i \neq i_0$. Donc si on suppose que c est différent de tous les c_i, alors T divise P'_{i_0}, et donc P_{i_0} admet un facteur carré, ce qui est absurde. Donc il existe i tel que $c = c_i$. □

Pour calculer l'intégrale d'une fration rationnelle $\frac{P}{Q}$ avec Q sans facteur carré, on a donc deux méthodes : décomposer en éléments simples ou utiliser le théorème de Rothstein-Trager. Pour cette deuxième méthode, l'avantage est qu'on n'a pas besoin de connaître les racines de Q, même si en revanche on devra calculer celles du résultant $R(Y)$.

Réduction de Hermite :

Par le théorème de décomposition en éléments simples, la recherche de primitives de fractions rationnelles est ramenée à la recherche de primitives de $\frac{P}{Q^i}$ où Q est sans facteur carré et $\deg(P) < \deg(Q)$. Comme Q est sans facteur carré, Q et Q' sont premiers entre eux, et par l'identité de Bézout, il existe U, V tels que $UQ+VQ' = 1$. En multipliant par P, il vient

$$P = PUQ + PVQ'.$$

On peut réduire davantage en faisant les deux divisions euclidiennes suivantes : $PU = AQ' + R$ avec $\deg(R) < \deg(Q')$ et $PV = BQ + S$ avec $\deg(S) < \deg(Q)$. Alors

$$P = (A + B)QQ' + (RQ + SQ'),$$

avec $\deg(RQ + SQ') \leqslant \max(\deg(RQ), \deg(SQ')) < \deg(QQ')$. Or la partie entière de $\frac{P}{Q}$ est nulle, car $\deg(P) < \deg(Q)$. Cela impose $A+B = 0$ (on a $\frac{P}{Q} = (A+B)Q' + R + \frac{SQ'}{Q}$, mais R n'est pas nécessairement nul puisque $\frac{SQ'}{Q}$ admet une partie entière qui peut compenser R, les degrés le permettent). Finalement

$$P = RQ + SQ'$$

avec $\deg(R) < \deg(Q')$ et $\deg(S) < \deg(Q)$.

Pour $i > 1$, on a donc

$$\int \frac{P}{Q^i} = \int \frac{R}{Q^{i-1}} + \int \frac{SQ'}{Q^i}$$

En intégrant par partie, il vient

$$\begin{aligned}\int \frac{P}{Q^i} &= \int \frac{R}{Q^{i-1}} + \frac{S}{(1-i)Q^{i-1}} - \int \frac{S'}{(1-i)Q^{i-1}} \\ &= \frac{S}{(1-i)Q^{i-1}} + \int \frac{(i-1)R + S'}{(i-1)Q^{i-1}},\end{aligned}$$

où l'on a $\deg((i-1)R + S') < \deg(Q') < \deg(Q)$. En itérant ce procédé, on est ramené à chercher des primitives de $\frac{P}{Q}$ avec Q sans facteur carré et $\deg(P) < \deg(Q)$. Ce procédé est nommé *réduction de Hermite*. Le théorème de Rothstein-Trager intervient à ce niveau-là : il nous permet de calculer ces primitives sans factoriser Q sur $\mathbb{C}$, en utilisant des PGCD et des résultants.

Références :

– Saux-Picart - *Cours de calcul formel, Algorithmes fondamentaux* - Page 153.

2.24 Théorème de Wantzel

Commençons par donner la structure des extensions de degré 2, dites aussi extensions quadratiques : elles sont obtenues par adjonction d'une racine carrée.

Proposition. *Soit E un sous-corps de $\mathbb{R}$ (plus généralement un corps de caractéristique différente de 2). Soit $j : E \to F$ une extension de degré 2. Alors il existe un élément $a \in F\backslash E$ tel que $a^2 \in E$ et $F = E[a]$.*

Démonstration. Soit $x \in F\backslash E$. La famille $(1, x)$ est libre sur E, donc une base de F comme E-espace vectoriel. La famille $(1, x, x^2)$ est alors liée : il existe $a, b, c \in E$ non tous nuls tels que $ax^2 + bx + c = 0$. Comme $(1, x)$ est libre, $a \neq 0$, et on peut écrire :

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}.$$

Posons $\delta = 2ax + b$. Alors $\delta^2 = b^2 - 4ac \in E$ est le discriminant du polynôme $aX^2 + bX + c$. Comme $x = \frac{\delta - b}{2a}$, la famille $(1, \delta)$ est une base de F sur E. □

Théorème (Wantzel)**.** *Soit E un sous-corps de $\mathbb{R}$. Un réel x est constructible à la règle et au compas à partir de E si et seulement s'il existe une suite $E_0, \ldots, E_n$ de sous-corps de $\mathbb{R}$ telle que*

(i) $E = E_0 \subset E_1 \subset \cdots \subset E_n$.

(ii) $[E_i : E_{i-1}] = 2$, *pour tout* $i \in [\![1, n]\!]$.

(iii) $x \in E_n$.

Démonstration. Soit $\alpha \in \mathbb{R}$ constructible à partir de E. Raisonnons par récurrence sur le nombre r d'étapes pour construire α. Montrons le résultat pour $r = 1$. Dans ce cas, il existe un point P dont une des coordonnées est égale à α, qui est constructible en une étape à partir de l'ensemble F des points du plan $\mathbb{R}^2$ dont les coordonnées sont des éléments de E. Cela signifie que α est à l'intersection de deux droites, d'une droite et d'un cercle, ou de deux cercles constructibles (*i.e.* les droites doivent passer par deux points de F et les cercles doivent avoir pour centre un point de F et passer en un autre point de F).

Tout d'abord, une droite D passant par deux points $A = (a, a')$ et $B = (b, b')$ dont les coordonnées sont dans E possède une équation à coefficients dans E :

$$\begin{aligned} M = (x, y) \in D &\iff \det(\overrightarrow{AM}, \overrightarrow{AB}) = 0 \\ &\iff \det \begin{pmatrix} x - a & b - a \\ y - b & b' - a' \end{pmatrix} = 0. \end{aligned}$$

Un cercle C de centre A et de rayon AB possède aussi une équation à coefficients dans E :

$$M = (x, y) \in C \iff (x - a)^2 + (y - a')^2 = (b - a)^2 + (b' - a')^2.$$

Les coordonnées du point à l'intersection de deux droites concourantes sont donc des expressions rationnelles en les coefficients des équations des droites. Elles sont donc dans E.

Soit $M = (x, y)$ un point à une des intersections d'une droite et d'un cercle. Alors

$$\begin{cases} x^2 + y^2 + Ax + By + C = 0 \\ Dx + Ey + F = 0, \end{cases}$$

avec $A, B, C, D, E, F \in E$. Supposons par exemple $E \neq 0$. Alors en éliminant y dans la première équation, on obtient une équation du second degré à coefficients dans E dont x est solution. Son discriminant Δ appartient à E. Donc x appartient à l'extension $E(\sqrt{\Delta})$ qui est de degré au plus 2 sur K. Enfin, on a également $y \in E(\sqrt{\Delta})$.

Si M est à l'intersection de deux cercles, on soustrait les deux équations de cercles et on est ramené au cas précédent (intersection d'un cercle et d'une droite).

Finalement, soit $\alpha \in E$, soit $\alpha \in E(\sqrt{\Delta})$ et on a le résultat. La récurrence est immédiate.

Réciproquement, soit $E_0, \ldots, E_n$ une suite de sous-corps de $\mathbb{R}$ vérifiant les hypothèses, avec $x \in E_n$. Pour montrer que x est constructible, il suffit de montrer que si $E \subset F$ est une extension de degré 2, tout élément de F est constructible à partir de E.

D'après la proposition démontrée juste avant, il existe $a \in F$ tel que $F = E[a]$ et $a^2 \in E$. Or la racine carrée d'un nombre constructible est constructible. Donc $a = \pm\sqrt{a^2}$ est constructible. Ainsi, tout élément de F est constructible, car il s'écrit $x + ay$ avec $x, y \in E$. □

Remarque. Comme l'ensemble des nombres constructibles est un corps, il revient au même de dire que x est constructible à partir d'une partie E contenant 0 et 1 que de dire qu'il est constructible à partir du corps engendré par E dans $\mathbb{R}$. En particulier, être constructible à partir de $\{0, 1\}$ et l'être à partir de $\mathbb{Q}$ sont deux notions équivalentes.

Proposition. *L'ensemble K des nombres réels constructibles est un sous-corps de $\mathbb{R}$.*

Démonstration. Soient $a, b \in K$. On a facilement que $a+b \in K$ et $a-b \in K$. Les réels $x = ab$ et $y = \frac{1}{a}$ sont obtenues avec le théorème de Thalès :

$$\frac{a}{x} = \frac{1}{b} \quad \text{et} \quad \frac{y}{1} = \frac{1}{a},$$

où la première fraction est le quotient des longueurs des deux segments parallèles (faire le dessin). □

Proposition. *Si x est constructible, alors $\sqrt{x}$ est constructible.*

Démonstration. On place x et -1 sur l'axe des abscisses et on trace le cercle passant par ces deux points : son centre est à l'abscisse $\frac{x-1}{2}$ et son rayon vaut $\frac{x+1}{2}$. On note h le nombre à l'intersection de ce cercle et du demi-axe positif $[Oy)$. Alors

$$h = \sqrt{\left(\frac{x+1}{2}\right)^2 - \left(\frac{x-1}{2}\right)^2} = \sqrt{x}.$$

□

Corollaire. *Soit E un sous-corps de $\mathbb{R}$ et soit x un nombre réel constructible à partir de E. Alors x est algébrique sur E et son degré est une puissance de 2.*

Démonstration. Soit $E = E_0 \subset E_1 \subset \cdots \subset E_n \subset \mathbb{R}$ une chaîne d'extensions quadratiques avec $x \in E_n$. La multiplicativité des degrés entraîne que

$$[E_n : E] = [E_n : E_1][E_1 : E_0] = 2[E_n : E_1] = \cdots = 2^n.$$

En considérant les extensions $E \subset E[x] \subset E_n$, on obtient que le degré d de $E[x]$ sur E doit diviser 2^n, c'est donc une puissance de 2.

La famille $(1, x, \ldots, x^d)$ est donc liée sur $E[x]$ vu comme E-espace vectoriel. Donc x est algébrique sur E. □

Références :

– Audin - *Géométrie* - Page 130.

2.25 Théorème de Wedderburn

Théorème. *Tout corps fini est commutatif.*

Démonstration. Soit K un corps fini dont on note Z le centre :

$$Z = \{x \in K \,/\, xy = yx \text{ pour tout } y \in K\}.$$

Cet ensemble Z est un sous-corps commutatif de K et on note q son cardinal. On sait que $q \geqslant 2$ car $0, 1 \in Z$. Comme on peut voir K comme un Z-espace vectoriel de dimension finie n, on a $\mathrm{Card}(K) = q^n$ (on prend $e_1 \in K \backslash Z$ et on obtient $Z + Ze_1 \subset K$, puis $e_2 \in K \backslash (Z + Ze_1)$, et ainsi de suite jusqu'à avoir $K = Z + Ze_1 + \cdots + Ze_{n-1}$).

On suppose que K n'est pas commutatif (donc que $n > 1$). Le groupe multiplicatif K^* opère sur lui-même par conjugaison. Pour $x \in K^*$, on note O_x l'orbite de x et

$$k_x = \{y \in K \,/\, yx = xy\}.$$

Alors k_x est un sous-corps de K (pas nécessairement commutatif) et on a $k_x^* = \mathrm{Stab}_{K^*}(x)$. Comme $Z \subset k_x$, on peut voir k_x comme un Z-espace vectoriel de dimension finie d et on a $\mathrm{Card}(k_x) = q^d$.

Comme $k_x^* \subset K^*$, on a $q^d - 1 \,|\, q^n - 1$, et donc $d \,|\, n$ (voir la proposition plus bas). Enfin, on a

$$\mathrm{Card}(O_x) = \frac{\mathrm{Card}(K^*)}{\mathrm{Card}(k_x^*)} = \frac{q^n - 1}{q^d - 1}.$$

Soit Φ_m le m-ième polynôme cyclotomique. Grâce à la relation

$$X^n - 1 = \prod_{m \,|\, n} \Phi_m,$$

on obtient

$$\frac{q^n - 1}{q^d - 1} = \frac{\prod_{m \,|\, n} \Phi_m(q)}{\prod_{m \,|\, d} \Phi_m(q)} = \prod_{m \,|\, n,\, m \nmid d} \Phi_m(q).$$

Donc si $d \neq n$, $\Phi_n(q) \,|\, \frac{q^n - 1}{q^d - 1}$.

On écrit maintenant l'équation aux classes :

$$\mathrm{Card}(K^*) = \mathrm{Card}(Z^*) + \sum_{x \in E} \mathrm{Card}(O_x)$$

où E désigne un ensemble de représentants des orbites tel que pour tout $x \in E$, $x \notin Z$. Mais si $x \notin Z$, alors $k_x \neq K$, et donc $d \neq n$. On obtient ainsi :

$$q^n - 1 = q - 1 + \sum \frac{q^n - 1}{q^d - 1}$$

où la somme porte sur un certain nombre de diviseurs stricts de n et d'après ce qui précède, tous les termes de cette somme sont divisibles par $\Phi_n(q)$. Cette égalité prouve que $\Phi_n(q) \,|\, q-1$, et en particulier, $|\Phi_n(q)| \leqslant q-1$.

Par définition du n-ième polynôme cyclotomique, on a $\Phi_n(q) = (q-\omega_1)\dots(q-\omega_{\varphi(n)})$ où les ω_i sont les racines primitives n-ièmes de l'unité dans $\mathbb{C}$. Comme $n \neq 1$, $\omega_i \neq 1$ pour tout i. Mais alors, pour tout i, $|q-\omega_i| > q-1$ (cela se voit géométriquement sur un dessin), et donc

$$|\Phi_n(q)| > (q-1)^{\varphi(n)} \geqslant q-1,$$

ce qui est une contradiction. □

Proposition. *Soient q un entier $\geqslant 2$, et $d, n \in \mathbb{N}$. Si $q^d-1 \,|\, q^n-1$, alors $d \,|\, n$.*

Démonstration. On effectue la division euclidienne de n par d : il existe $k, r \in \mathbb{N}$ tels que $n = kd+r$ et $0 \leqslant r < d$. On se rappelle ensuite de la formule :

$$a^n - b^n = (a-b)\sum_{i=0}^{n-1} a^i b^{n-1-i}.$$

On a donc :

$$\begin{aligned} q^n-1 &= q^{kd+r}-1 \\ &= q^{kd}q^r - q^r + q^r - 1 \\ &= q^r(q^{kd}-1) + (q^r-1) \\ &= q^r(q^d-1)\left(\sum_{i=0}^{k-1} q^{id}(-1)^{k-1-i}\right) + (q^r-1). \end{aligned}$$

Mais par hypothèse, $q^d-1 \,|\, q^n-1$, donc nécessairement, $q^d-1 \,|\, q^r-1$. Comme $r<d$, cela implique $r=0$, et donc $n=kd$. □

Références :

– Perrin - *Cours d'algèbre* - Page 82.

2.26 Théorème des deux carrés

Le problème est de déterminer quels entiers $n \in \mathbb{N}$ sont somme de deux carrés : $n = a^2 + b^2$ avec $a, b \in \mathbb{N}$. On pose :

$$\Sigma = \{n \in \mathbb{N} \,/\, \exists a, b \in \mathbb{N} \text{ avec } n = a^2 + b^2\}.$$

On introduit $\mathbb{Z}[i] = \{a + ib \in \mathbb{C} \,/\, a, b \in \mathbb{Z}\}$ anneau euclidien, muni de la "norme" $N(a + ib) = |z|^2 = a^2 + b^2$, qui est clairement multiplicative.

Théorème (Théorème des deux carrés, version faible)**.** *Soit $p \in \mathbb{N}$ un nombre premier. On a l'équivalence :*

$$p \in \Sigma \iff p = 2 \text{ ou } p \equiv 1\,[4].$$

Démonstration. Nous allons pouvoir raisonner uniquement par équivalence. D'abord :

$$p \in \Sigma \iff p \text{ n'est pas irréductible dans } \mathbb{Z}[i].$$

En effet, si $p \in \Sigma$, $p = a^2 + b^2 = (a + ib)(a - ib)$. Comme p est premier, on a nécessairement $a \neq 0$ et $b \neq 0$. Donc $a + ib$ et $a - ib$ ne sont pas dans $\mathbb{Z}[i]^* = \{1, -1, i, -i\}$. Donc p n'est pas irréductible dans $\mathbb{Z}[i]$.

Réciproquement, si $p = zz'$ avec $z, z' \notin \mathbb{Z}[i]^* = \{1, -1, i, -i\}$, on a $N(p) = N(z)N(z') = p^2$ et comme $N(z), N(z') \neq 1$, on a $p = N(z) = N(z')$. Donc en particulier, $p = N(z) = N(a + ib) = a^2 + b^2 \in \Sigma$.

Ensuite, $\mathbb{Z}[i]$ est euclidien, donc principal, donc factoriel, et donc :

$$\begin{aligned} p \text{ est réductible dans } \mathbb{Z}[i] &\iff (p) \text{ n'est pas premier} \\ &\iff \mathbb{Z}[i]/(p) \text{ n'est pas intègre (par définition).} \end{aligned}$$

Par ailleurs, on a $\mathbb{Z}[i] \cong \mathbb{Z}[X]/(X^2 + 1)$. Pour le voir, on définit $\varphi : \mathbb{Z}[X] \to \mathbb{Z}[i]$ avec $\varphi(X) = i$. Alors φ est un morphisme d'anneaux surjectif. Si $P \in \mathbb{Z}[X]$ est tel que $\varphi(P) = 0$, on effectue la division euclidienne de P par X^2+1 : $P = (X^2+1)Q+R$. Donc $\varphi(P) = 0$ implique $R(i) = 0$, mais R est de degré $\leqslant 1$ et la famille $(1, i)$ est libre, donc $R = 0$. Finalement $\mathrm{Ker}(\varphi) = (X^2 + 1)$, et $\mathbb{Z}[X]/\mathrm{Ker}(\varphi) \cong \mathbb{Z}[i]$.

On obtient alors les isomorphismes suivants par le théorème d'isomorphisme :

$$\begin{aligned} \mathbb{Z}[i]/(p) &\cong \big(\mathbb{Z}[X]/(X^2 + 1)\big)/(p) \\ &\cong \mathbb{Z}[X]/(X^2 + 1, p) \\ &\cong \big(\mathbb{Z}[X]/(p)\big)/(X^2 + 1) \\ &\cong (\mathbb{Z}/p\mathbb{Z})[X]/(X^2 + 1) = \mathbb{F}_p[X]/(X^2 + 1). \end{aligned}$$

On poursuit alors nos équivalences :

$$\begin{aligned}\mathbb{Z}[i]/(p) \text{ n'est pas intègre} &\iff \mathbb{F}_p[X]/(X^2+1) \text{ n'est pas intègre}\\ &\iff (X^2+1) \text{ n'est pas premier}\\ &\iff X^2+1 \text{ est réductible sur } \mathbb{F}_p,\end{aligned}$$

La dernière équivalence est justifiée par le fait que $\mathbb{F}_p$ est factoriel (car c'est un corps), donc aussi $\mathbb{F}_p[X]$. On poursuit encore :

$$\begin{aligned}X^2+1 \text{ est réductible sur } \mathbb{F}_p &\iff X^2+1 \text{ admet une racine dans } \mathbb{F}_p\\ &\iff -1 \text{ est un carré dans } \mathbb{F}_p^*\\ &\iff p=2 \text{ ou } p\equiv 1\,[4].\end{aligned}$$

Pour justifier la dernière équivalence lorsque $p\neq 2$, on considère, pour $q=p$ ou plus généralement $q=p^r$ avec $r\geqslant 1$, le morphisme de groupes (groupes multiplicatifs)

$$\begin{array}{rccc}\varphi: & \mathbb{F}_q^* & \longrightarrow & \mathbb{F}_q^{*2}\\ & x & \longmapsto & x^2\end{array}$$

qui est surjectif ($\mathbb{F}_q^{*2}=\{x\in\mathbb{F}_q^*\,/\,\exists y\in\mathbb{F}_q^*,\, x=y^2\}$). Son noyau vaut $\{-1,1\}$ car le polynôme X^2-1 admet au plus deux racines, ce sont donc -1 et 1. Donc $\mathrm{Im}(\varphi)=\mathbb{F}_q^{*2}\cong\mathbb{F}_q^*/\mathrm{Ker}(\varphi)$. En prenant les cardinaux, on obtient $\mathrm{Card}(\mathbb{F}_q^{*2})=\frac{q-1}{2}$. On montre ensuite que

$$\mathbb{F}_q^{*2}=\left\{x\in\mathbb{F}_q\,/\,x^{\frac{q-1}{2}}=1\right\}.$$

En effet, en notant A le deuxième ensemble, son cardinal est $\leqslant\frac{q-1}{2}$ (nombre de racines maximum du polynôme $X^{\frac{q-1}{2}}-1$). D'autre part si $x\in\mathbb{F}_q^{*2}$, alors $x=y^2$, et donc $x^{\frac{q-1}{2}}=y^{q-1}=1$. Donc $\mathbb{F}_q^{*2}\subset A$. Pour une raison de cardinal, on conclut l'égalité.

A présent, on a :

$$-1\in\mathbb{F}_p^{*2}\ \Leftrightarrow\ (-1)^{\frac{p-1}{2}}=1\ \Leftrightarrow\ \frac{p-1}{2}\text{ est pair}\ \Leftrightarrow\ p\equiv 1\,[4].$$

□

Théorème (Théorème des deux carrés)**.** *Soit $n\in\mathbb{N}^*$, $n\neq 1$. On décompose n en facteurs premiers : $n=\prod_{p\in\mathcal{P}}p^{v_p(n)}$. Alors on a l'équivalence :*

$$n\in\Sigma\iff v_p(n)\ \textit{est pair pour tout}\ p\equiv 3\,[4].$$

Démonstration. On traduit la propriété $n\in\Sigma$ en termes d'entiers de Gauss :

$$n\in\Sigma\iff\exists z\in\mathbb{Z}[i]\text{ tel que }n=N(z).$$

Alors comme N est multiplicative, si $n, n' \in \Sigma$, alors $n = N(z)$ et $n' = N(z')$, et donc $nn' = N(zz') \in \Sigma$. Le sens indirect de l'équivalence est donc immédiat grâce à la version faible du théorème des deux carrés : n est un produit de nombres premiers p_i, avec $p_i \in \Sigma$ si $p_i = 2$ ou $p_i \equiv 1\,[4]$, et si $p_i \equiv 3\,[4]$, on a $p_i^2 \in \Sigma$ car un carré est toujours dans Σ (en effet, $p_i^2 = p_i^2 + 0$).

Montrons le sens direct. Soit $p \equiv 3\,[4]$. On va montrer par récurrence sur $v_p(n)$ que $v_p(n)$ est pair. Si $v_p(n) = 0$, c'est clair. Sinon, p divise $n = a^2 + b^2 = (a+ib)(a-ib)$. D'après les résultats précédents, $p \notin \Sigma$ (car $p \equiv 3\,[4]$), et ceci est équivalent à p est irréductible dans $\mathbb{Z}[i]$. Donc p divise $(a+ib)$ ou $(a-ib)$. Or p est entier, donc p divise a et b. On écrit $a = pa'$ et $b = pb'$. Alors $\frac{n}{p^2} = a'^2 + b'^2 \in \Sigma$. Mais $v_p\left(\frac{n}{p^2}\right) = v_p(n) - 2$ est pair d'après l'hypothèse de récurrence, donc aussi $v_p(n)$. □

Théorème (Théorème d'isomorphisme)**.** *Soit $f : A \to B$ un homomorphisme d'anneaux. On note $I = \mathrm{Ker}(f)$. Soient $J \subset I$ un idéal de A et $p : A \to A/J$ la projection canonique. Alors :*

(i) Il existe un unique homomorphisme $\overline{f} : A/J \to B$ tel que $f = \overline{f} \circ p$.

(ii) $\overline{f}$ est injectif si et seulement si $J = I$.

(iii) $\overline{f}$ est surjectif si et seulement si f l'est.

En particulier, on a $\mathrm{Im}(f) \cong A/\mathrm{Ker}(f)$.

Remarque. Justification de $\mathbb{F}_p$ est factoriel : $\mathbb{F}_p$ est un corps, donc $\mathbb{F}_p$ est engendré par 1 en tant qu'idéal. Tout idéal de $\mathbb{F}_p$ non réduit à $\{0\}$ contient 1 (s'il contient $x \neq 0$, il contient $x^{-1}x = 1$), donc vaut $\mathbb{F}_p$. Les idéaux de $\mathbb{F}_p$ sont $\{0\} = (0)$ et $\mathbb{F}_p = (1)$ qui sont bien principaux. Donc $\mathbb{F}_p$ est principal, donc factoriel.

En fait, on a même que tout corps K est un anneau euclidien, donc principal, donc factoriel. En effet, si $x \in K$ et $y \in K^*$, alors $x = (xy^{-1})y + 0$. Le reste de la division euclidienne est toujours nul et on peut définir n'importe quel stathme.

Proposition. *On a $\mathbb{Z}[i]^* = \{-1, 1, -i, i\}$.*

Démonstration. Si $z \in \mathbb{Z}[i]^*$, il existe $z' \in \mathbb{Z}[i]$ tel que $zz' = 1$. Donc $N(zz') = N(z)N(z') = N(1) = 1$. Comme $N(z), N(z') \in \mathbb{N}$, cela implique $N(z) = N(z') = 1$. En écrivant $z = a + ib$, on a donc $a^2 + b^2 = 1$, d'où le résultat. □

Proposition. *L'anneau $\mathbb{Z}[i]$ est euclidien (relativement à N), donc principal.*

Démonstration. Soient $z, t \in \mathbb{Z}[i]\backslash\{0\}$. Pour faire la division euclidienne de z par t, on commence par considérer $\frac{z}{t} \in \mathbb{C}$. On approxime $\frac{z}{t}$ par un entier de Gauss q : si $\frac{z}{t} = x + iy$, on prend $q = a + ib$ où a et b sont les entiers les plus proches de x et y. On a ainsi :

$$\left|\frac{z}{t} - q\right| = \sqrt{(x-a)^2 + (y-b)^2} \leqslant \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \frac{\sqrt{2}}{2} < 1.$$

Alors $r = z - qt \in \mathbb{Z}[i]$ est tel que $|r| = |t|\left|\frac{z}{t} - q\right| < |t|$. En élevant au carré, on obtient $N(r) < N(t)$. On a donc bien écrit $z = qt + r$ avec $N(r) < N(t)$. □

Remarque. On peut expliciter un isomorphisme d'anneaux pour montrer que

$$\mathbb{Z}[i]/(p) \cong (\mathbb{Z}/p\mathbb{Z})[X]/(X^2+1).$$

En effet, soit

$$\begin{array}{rccc} \varphi : & \mathbb{Z}[i] & \longrightarrow & (\mathbb{Z}/p\mathbb{Z})[X]/(X^2+1) \\ & a+ib & \longmapsto & \overline{a} + \overline{b}\Pi(X) \end{array}$$

où $\Pi : (\mathbb{Z}/p\mathbb{Z})[X] \to (\mathbb{Z}/p\mathbb{Z})[X]/(X^2+1)$ est la surjection canonique. On vérifie alors que φ est un morphisme d'anneau surjectif (facile), de noyau $(p) = p\mathbb{Z}[i]$ (facile aussi).

Proposition. *Soit A un anneau intègre. On a :*

(i) L'idéal (p) est premier implique p est irréductible.

(ii) Si A est factoriel, alors la réciproque est vraie : p est irréductible implique l'idéal (p) est premier.

Démonstration. (i) Soit (p) un idéal premier. Alors si $p = ab$, on a $ab \in (p)$, et comme (p) est premier, cela implique $a \in (p)$ ou $b \in (p)$. Si par exemple $a \in (p)$, alors p divise a : $a = pc$, $c \in A$. On a donc $p = ab = pbc$. Comme A est intègre, cela implique $bc = 1$, *i.e.* $b \in A^*$. Donc p est irréductible.

(ii) Soit p irréductible. Soit $ab \in (p)$. Alors il existe $k \in A$ tel que $ab = kp$. Comme A est factoriel, on peut décomposer a, b et k de façon unique en produit de facteurs irréductibles à une unité de l'anneau près. En comparant les décompositions de ab et kp, on obtient que p divise a ou b. Donc $a \in (p)$ ou $b \in (p)$, ce qui prouve que (p) est premier.

□

Références :

– Perrin - *Cours d'algèbre* - Page 56.

2.27 Topologie des orbites de l'action de Steinitz

Propriété. *Soit $K = \mathbb{R}$ ou $\mathbb{C}$. On note O_r l'orbite des matrices de $\mathcal{M}_{n,p}(K)$ de rang r (on a $0 \leqslant r \leqslant \min(n,p)$) pour l'action de Steinitz. Alors*

$$\overline{O_r} = \bigsqcup_{k=0}^{r} O_k.$$

Démonstration. On commence par montrer que $\bigsqcup_{k=0}^{r} O_k$ est un fermé de $\mathcal{M}_{n,p}(K)$. Soient $I \subset \{1,\dots,n\}$ et $J \subset \{1,\dots,p\}$ tels que $|I| = |J|$. On note

$$\begin{array}{rccc} \Delta_{I,J} : & \mathcal{M}_{n,p}(K) & \longrightarrow & K \\ & (a_{ij})_{i\leqslant n, j\leqslant n} & \longmapsto & \det((a_{ij})_{i\in I, j\in J}), \end{array}$$

application mineur d'indice (I,J). On considère ensuite l'application

$$\begin{array}{rccc} \delta : & \mathcal{M}_{n,p}(K) & \longrightarrow & K^{\binom{p}{r+1}\binom{n}{r+1}} \\ & A & \longmapsto & (\Delta_{I,J}(A))_{I\subset\{1,\dots,n\}, J\subset\{1,\dots,p\}, |I|=|J|=r+1}. \end{array}$$

Comme le rang d'une matrice A est l'ordre de son plus grand mineur non nul, on a :

$$\begin{aligned} \bigsqcup_{k=0}^{r} O_k &= \{A \in \mathcal{M}_{n,p}(K) \,/\, \mathrm{rg}(A) \leqslant r\} \\ &= \{A \in \mathcal{M}_{n,p}(K) \,/\, \Delta_{I,J}(A) = 0, \forall I \subset \{1,\dots,n\}, \\ &\quad \forall J \subset \{1,\dots,p\}, |I| = |J| \geqslant r+1\} \\ &= \{A \in \mathcal{M}_{n,p}(K) \,/\, \Delta_{I,J}(A) = 0, \forall I \subset \{1,\dots,n\}, \\ &\quad \forall J \subset \{1,\dots,p\}, |I| = |J| = r+1\} \\ &= \delta^{-1}(\{0\}). \end{aligned}$$

Les applications $\Delta_{I,J}$ étant continues car polynomiales en les coefficients de la matrice, δ est continue et $\bigsqcup_{k=0}^{r} O_k$ est un fermé de $\mathcal{M}_{n,p}(K)$.

Comme $O_r \subset \bigsqcup_{k=0}^{r} O_k$ et comme $\overline{O_r}$ est le plus petit fermé contenant O_r, on obtient

$$\overline{O_r} \subset \bigsqcup_{k=0}^{r} O_k.$$

Soit maintenant $A \in \bigsqcup_{k=0}^{r} O_k$. Il existe $k \in [\![0,r]\!]$ tel que $A \in O_k$, *i.e.* A s'écrit $A = PJ_kQ^{-1}$ avec $(P,Q) \in \mathrm{GL}_n(K) \times \mathrm{GL}_p(K)$. Pour $n \in \mathbb{N}^*$, on définit

$$A_n = P \begin{pmatrix} I_k & 0 & 0 \\ 0 & \frac{1}{n} I_{r-k} & 0 \\ 0 & 0 & 0 \end{pmatrix} Q^{-1}.$$

Pour tout $n \in \mathbb{N}^*$, $A_n \in O_r$ et $\lim_{n\to\infty} A_n = A$, donc $A \in \overline{O_r}$. On a donc démontré l'autre inclusion. □

Application. $\mathrm{GL}_n(K)$ *est dense dans* $\mathcal{M}_n(K)$.

Démonstration. On a $\mathrm{GL}_n(K) = O_n$ l'orbite des matrices de rang n dans $\mathcal{M}_n(K)$, donc d'après ce qui précède,

$$\overline{\mathrm{GL}_n(K)} = \overline{O_n} = \bigsqcup_{k=0}^{n} O_k = \mathcal{M}_n(K).$$

□

Proposition. *On suppose* $K = \mathbb{C}$. *Alors pour tout* $r \in [\![0, \min(n,p)]\!]$, O_r *est connexe.*

Démonstration. Soit l'application

$$\begin{array}{rccc} \varphi_r : & \mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_p(\mathbb{C}) & \longrightarrow & \mathcal{M}_{n,p}(\mathbb{C}) \\ & (P,Q) & \longmapsto & PJ_rQ^{-1} \end{array}$$

qui est continue car les coefficients de PJ_rQ^{-1} sont polynomiaux en les coefficients de P et Q (on rappelle que $Q^{-1} = \frac{1}{\det(Q)}{}^t\mathrm{com}(Q)$).

Ensuite, comme $\mathrm{GL}_n(\mathbb{C})$ et $\mathrm{GL}_p(\mathbb{C})$ sont connexes, leur produit $\mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_p(\mathbb{C})$ l'est aussi.

Enfin,

$$O_r = \varphi_r\big(\mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_p(\mathbb{C})\big)$$

est connexe comme image d'un connexe par une application continue. □

K désigne un corps commutatif.

Définition. Le groupe $G = \mathrm{GL}_n(K) \times \mathrm{GL}_p(K)$ agit à gauche sur $\mathcal{M}_{n,p}(K)$ par :

$$\begin{array}{rcl} G \times \mathcal{M}_{n,p}(K) & \longrightarrow & \mathcal{M}_{n,p}(K) \\ ((P,Q),M) & \longmapsto & PMQ^{-1}. \end{array}$$

On dit que deux matrices $A, B \in \mathcal{M}_{n,p}(K)$ sont équivalentes si elles sont dans la même orbite pour cette action, appelée *action de Steinitz*.

Proposition. *Soit* $r \in [\![1, \min(n,p)]\!]$ *et* $J_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. *L'orbite de* J_r *est l'ensemble des matrices de rang* r *de* $\mathcal{M}_{n,p}(K)$. *De plus, toute matrice de* $\mathcal{M}_{n,p}(K)$ *est dans l'orbite d'une unique matrice* J_r.

Théorème (Théorème du rang)**.** *Deux matrices* $A, B \in \mathcal{M}_{n,p}(K)$ *sont dans la même orbite si et seulement si elles ont même rang.*

Références :

- Caldero, Germoni - *Histoires hédonistes de groupes et de géométries.*
- Francinou, Gianella, Nicolas - *Oraux X-ENS, Algèbre 2* - Page 217 (autre version).

Chapitre 3

Développements d'analyse

3.1 Calcul de l'intégrale $\int_0^\infty t^{\alpha-1}e^{it}\,\mathrm{d}t$

Soit $\alpha \in]0,1[$. On veut calculer $J = \int_0^\infty t^{\alpha-1}e^{it}\,\mathrm{d}t$. Son existence est justifiée par la règle d'Abel, ou par une intégration par parties :

$$\int_a^b t^{\alpha-1}e^{it}\,\mathrm{d}t = \left[\frac{1-e^{it}}{-i}t^{\alpha-1}\right]_a^b - \int_a^b \frac{1-e^{it}}{-i}\frac{\alpha-1}{t^{2-\alpha}}\,\mathrm{d}t.$$

Comme $|1-e^{it}| \sim t$ quand $t \to 0$, on en déduit que la limite quand $a \to 0$ existe. La limite quand $b \to \infty$ existe aussi.

Pour $x \geqslant 0$, on pose

$$I(x) = \int_0^\infty t^{\alpha-1}e^{ixt}e^{-t}\,\mathrm{d}t.$$

On va commencer par calculer $I(x)$. Vérifions les hypothèses du théorème de dérivation des intégrales à paramètre :

(i) Pour tout $x \geqslant 0$, $t \mapsto t^{\alpha-1}e^{ixt}e^{-t}$ est intégrable sur $]0,\infty[$ car continue et majorée en module par $t^{\alpha-1}e^{-t}$ qui est intégrable.

(ii) Pour tout $t \in]0,\infty[$, $x \mapsto t^{\alpha-1}e^{ixt}e^{-t}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}_+$, de dérivée $x \mapsto it^\alpha e^{ixt}e^{-t}$, majorée en module par $e^{-t}t^\alpha$ indépendante de x et intégrable sur $]0,\infty[$.

On en conclut que I est dérivable sur $\mathbb{R}_+$, et pour tout $x \geqslant 0$,

$$I'(x) = i\int_0^\infty t^\alpha e^{ixt}e^{-t}\,\mathrm{d}t.$$

Par ailleurs, on intègre $I(x)$ par parties :

$$I(x) = \left[e^{(-1+ix)t}\frac{t^\alpha}{\alpha}\right]_0^\infty - \frac{-1+ix}{\alpha}\int_0^\infty t^\alpha e^{ixt}e^{-t}\,\mathrm{d}t = -\frac{i+x}{\alpha}I'(x),$$

pour tout $x \geqslant 0$. On sait intégrer cette équation différentielle :

$$\begin{aligned} I(x) &= I(0)\exp\left(\int_0^x \frac{-\alpha}{i+t}\,\mathrm{d}t\right) \\ &= I(0)\exp\left(-\alpha\int_0^x \frac{t-i}{1+t^2}\,\mathrm{d}t\right) \\ &= I(0)\exp\left(-\alpha\left(\frac{1}{2}\ln(1+x^2) - i\arctan(x)\right)\right) \\ &= I(0)\left(1+x^2\right)^{-\frac{\alpha}{2}}e^{i\alpha\arctan(x)}. \end{aligned}$$

On remarque enfin que $I(0) = \Gamma(\alpha)$.

Pour retrouver l'intégrale J, on fait le changement de variable $u = xt$ pour $x > 0$ dans l'intégrale définissant $I(x)$. On obtient :

$$I(x) = x^{-\alpha}\int_0^\infty e^{-\frac{u}{x}}e^{iu}u^{\alpha-1}\,\mathrm{d}u.$$

On a donc envie de faire tendre x vers l'infini à l'intérieur de l'intégrale pour obtenir J, soit encore de faire tendre y vers 0 dans l'intégrale :

$$K(y) = \int_0^\infty e^{-uy}e^{iu}u^{\alpha-1}\,\mathrm{d}u.$$

Mais le théorème de continuité sous le signe intégral ne s'applique pas car on ne peut dominer $e^{-uy}e^{iu}u^{\alpha-1}$ par une fonction intégrable indépendante de y que sur $[\varepsilon,\infty[$ pour $\varepsilon > 0$. On prouve ainsi que K est continue sur $[\varepsilon,\infty[$ pour tout $\varepsilon > 0$, donc sur $\mathbb{R}_+^*$, mais on ne peut pas conclure que K est continue en 0.

Pour montrer que K est continue en 0, on va l'exprimer comme limite uniforme de fonctions continues en 0. On considère la suite de fonctions $(K_n)_{n\in\mathbb{N}^*}$ où les K_n sont définies par :

$$K_n(y) = \int_0^n e^{-uy}e^{iu}u^{\alpha-1}\,\mathrm{d}u$$

pour tout $y \geqslant 0$. Comme on intègre sur un compact, le théorème de continuité sous le signe intégral s'applique immédiatement : les K_n sont continues sur $\mathbb{R}_+$. Il reste à montrer que la convergence vers K est uniforme. On calcule la différence :

$$\begin{aligned} K_n(y) - K(y) &= \int_n^\infty e^{-uy}e^{iu}u^{\alpha-1}\,\mathrm{d}u \\ &= \left[\frac{e^{(i-y)u}}{i-y}u^{\alpha-1}\right]_n^\infty - \int_n^\infty \frac{e^{(i-y)u}}{i-y}\frac{\alpha-1}{u^{2-\alpha}}\,\mathrm{d}u \\ &= \frac{1}{i-y}\left(\frac{e^{(i-y)n}}{n^{1-\alpha}} + (\alpha-1)\int_n^\infty \frac{e^{(i-y)u}}{u^{2-\alpha}}\,\mathrm{d}u\right). \end{aligned}$$

Comme $|e^{(i-y)n}| \leqslant 1$ et $\left|\frac{1}{i-y}\right| = \sqrt{\frac{1}{1+y^2}} \leqslant 1$, on obtient :

$$|K_n(y) - K(y)| \leqslant \frac{1}{n^{1-\alpha}} + (1-\alpha)\int_n^\infty \frac{1}{u^{2-\alpha}}\,\mathrm{d}u$$

et cette majoration tend vers 0 quand n tend vers l'infini indépendamment de y.

Au final, on a :

$$I(x) = \frac{1}{x^\alpha}K\left(\frac{1}{x}\right) = \Gamma(\alpha)\frac{1}{(1+x^2)^{\frac{\alpha}{2}}}e^{i\alpha\arctan(x)},$$

pour tout $x > 0$. En faisant tendre x vers l'infini, on obtient :

$$J = K(0) = \Gamma(\alpha)e^{i\alpha\frac{\pi}{2}}.$$

Application. *On effectue le changement de variable* $t = x^2$ *dans l'intégrale définissant* J*, qui est un* $\mathcal{C}^1$*-difféomorphisme de* $]0,\infty[\to]0,\infty[$ *:*

$$J = 2\int_0^\infty x^{2\alpha-1}e^{ix^2}\,\mathrm{d}x.$$

En prenant $\alpha = \frac{1}{2}$*, on trouve :*

$$J = 2\int_0^\infty e^{ix^2}\,\mathrm{d}x = \Gamma\left(\frac{1}{2}\right)e^{i\frac{\pi}{4}} = \sqrt{\pi}e^{i\frac{\pi}{4}}.$$

D'où la valeur de l'intégrale de Fresnel :

$$\int_{-\infty}^\infty e^{ix^2}\,\mathrm{d}x = 2\int_0^\infty e^{ix^2}\,\mathrm{d}x = \sqrt{\pi}e^{i\frac{\pi}{4}}.$$

Références :

- Gourdon - *Analyse* - Page 164 (exercice 4, question 2).

3.2 Complétude de l'espace $L^p(\mu)$

Théorème. *Pour $1 \leqslant p \leqslant \infty$ et pour toute mesure positive μ, $L^p(\mu)$ est un espace métrique complet.*

Démonstration. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy de $L^p(\mu)$.

Supposons d'abord $p < \infty$. On peut extraire une sous-suite $(f_{n_i})_{i\in\mathbb{N}}$ telle que

$$\|f_{n_{i+1}} - f_{n_i}\|_p \leqslant \frac{1}{2^i}.$$

On pose

$$g_k = \sum_{i=1}^{k} |f_{n_{i+1}} - f_{n_i}| \quad \text{et} \quad g = \sum_{i=1}^{\infty} |f_{n_{i+1}} - f_{n_i}|.$$

D'après l'inégalité de Minkowski, $\|g_k\|_p \leqslant \sum_{i=1}^{\infty} \frac{1}{2^i} \leqslant 1$ pour tout k. On applique le lemme de Fatou à la suite $(g_k^p)_{k\in\mathbb{N}^*}$:

$$\int_X \underbrace{\liminf\,(g_k^p)}_{=g^p} \mathrm{d}\mu \leqslant \liminf \int_X g_k^p \,\mathrm{d}\mu = \liminf \|g_k\|_p{}^p \leqslant 1,$$

qui nous donne $\|g\|_p \leqslant 1$. En particulier, on a $g(x) < \infty$ presque partout, et donc que la série

$$f_{n_1} + \sum_{i\geqslant 1} \big(f_{n_{i+1}}(x) - f_{n_i}(x)\big)$$

est absolument convergente pour presque tout x. On note alors $f(x)$ la somme de cette série lorsqu'elle converge et on pose $f(x) = 0$ sur l'ensemble restant qui est de mesure nulle. Comme

$$f_{n_1} + \sum_{i=1}^{k-1} \big(f_{n_{i+1}}(x) - f_{n_i}(x)\big) = f_{n_k},$$

on a directement $f(x) = \lim_{i\to\infty} f_{n_i}(x)$ pour presque tout x. Nous avons donc trouvé une fonction f qui est la limite simple presque partout de (f_{n_i}). Il faut maintenant démontrer que cette fonction est la limite de (f_n) au sens de $L^p(\mu)$.

Soit $\varepsilon > 0$. Il existe N tel que pour tout $n, m \geqslant N$, $\|f_n - f_m\|_p \leqslant \varepsilon$. On applique à nouveau le lemme de Fatou :

$$\int_X \underbrace{\liminf_{i\to\infty} |f_{n_i} - f_m|^p}_{=|f-f_m|^p} \mathrm{d}\mu \leqslant \liminf_{i\to\infty} \int_X |f_{n_i} - f_m|^p \,\mathrm{d}\mu \leqslant \varepsilon^p.$$

Cette inégalité montre que $f - f_m \in L^p(\mu)$, que $f \in L^p(\mu)$ (en effet, on écrit $f = (f - f_m) + f_m$ et on applique l'inégalité de Minkowski), et que $\|f - f_m\|_p \to 0$ quand $m \to \infty$. La démonstration est donc terminée dans le cas où $p < \infty$.

Supposons maintenant $p = \infty$. On note A_k l'ensemble sur lequel $|f_k(x)| > \|f_k\|_\infty$ et $B_{n,m}$ l'ensemble sur lequel $|f_n(x) - f_m(x)| > \|f_n - f_m\|_\infty$. Alors les A_k et les $B_{n,m}$ sont de mesure nulle par définition de la norme infini sur $L^\infty(\mu)$. On note :

$$E = \left(\bigcup_{k=1}^{\infty} A_k\right) \cup \left(\bigcup_{n=1}^{\infty}\bigcup_{m=1}^{\infty} B_{n,m}\right).$$

L'ensemble E est de mesure nulle car c'est une union dénombrable d'ensembles de mesure nulle. Pour $x \in {}^cE$,

$$|f_n(x) - f_m(x)| \leqslant \|f_n - f_m\|_\infty,$$

donc $(f_n(x))_{n\in\mathbb{N}}$ est une suite de Cauchy dans $\mathbb{C}$, qui converge donc vers $f(x) \in \mathbb{C}$.

D'autre part, on sait qu'il existe N tel que pour tout $n, m \geqslant N$,

$$|f_n(x) - f_m(x)| \leqslant \|f_n - f_m\|_\infty \leqslant \varepsilon.$$

Alors pour tout $n, m \geqslant N$, on a aussi

$$|f_m(x)| \leqslant |f_n(x)| + |f_n(x) - f_m(x)| \leqslant \|f_n\|_\infty + \|f_n - f_m\|_\infty \leqslant \|f_n\|_\infty + \varepsilon.$$

En faisant tendre m vers l'infini, on en déduit que pour tout $x \in {}^cE$,

$$|f_n(x) - f(x)| \leqslant \varepsilon \quad \text{et} \quad |f(x)| \leqslant \|f_n\|_\infty + \varepsilon,$$

en particulier f est bornée sur ${}^cE$.

Pour $x \in E$, on pose $f(x) = 0$. Alors $f \in L^\infty(\mu)$ et $\|f_n - f\|_\infty \to 0$ quand $n \to \infty$, ce qui achève la démonstration. □

Dans ce qui suit, on désigne par X un espace mesuré quelconque muni d'une mesure positive μ.

Définition $(L^p(\mu), 1 < p < \infty)$**.** Soit $1 < p < \infty$. Pour toute fonction f mesurable à valeurs complexes définie sur X, on pose

$$\|f\|_p = \left(\int_X |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}}$$

et on appelle $L^p(\mu)$ l'ensemble de toutes les fonctions f pour lesquelles $\|f\|_p < \infty$.

Définition (Borne essentielle)**.** Soient $g : X \to [0, \infty]$ mesurable et S l'ensemble de tous les nombres réels α tels que

$$\mu\big(g^{-1}(]\alpha, \infty])\big) = 0.$$

Pour $S = \varnothing$, on pose $\beta = \infty$, et pour $S \neq \varnothing$, $\beta = \inf(S)$. Comme

$$g^{-1}(]\beta, \infty]) = \bigcup_{n=1}^{\infty} g^{-1}\left(\left]\beta + \frac{1}{n}, \infty\right]\right)$$

et que la réunion dénombrable d'ensembles de mesure nulle est de mesure nulle, alors $\beta \in S$. On appelle β la *borne essentielle* de g.

Définition ($L^\infty(\mu)$)**.** Pour toute fonction f mesurable à valeurs complexes définie sur X, on pose

$$\|f\|_\infty = \text{ la borne essentielle de } |f|$$

et on appelle $L^\infty(\mu)$ l'ensemble de toutes les fonctions f mesurables pour lesquelles $\|f\|_\infty < \infty$.

On peut noter que le début de la démonstration de la complétude de $L^p(\mu)$ contient le résultat intéressant suivant :

Théorème. *Soit $1 \leqslant p \leqslant \infty$. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans $L^p(\mu)$, qui possède donc une limite $f \in L^p(\mu)$. Alors il existe une sous-suite de (f_n) qui converge ponctuellement presque partout vers f.*

Démonstration. Pour $p < \infty$, c'est la suite $(f_{n_i})_{i\in\mathbb{N}}$ de la démonstration. Pour $p = \infty$, la suite toute entière, $(f_n)_{n\in\mathbb{N}}$, convient. □

Références :

– Rudin - *Analyse réelle et complexe* - Page 82.

3.3 Comportement des solutions d'une équation différentielle linéaire

Théorème. *Soient $A \in \mathcal{M}_n(\mathbb{R})$, $B \in \mathcal{C}(\mathbb{R}_+, \mathcal{M}_n(\mathbb{R}))$, $f \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}^n)$. On suppose que l'application $t \mapsto e^{tA}$ est bornée sur $\mathbb{R}_+$ et que*

$$\int_0^\infty \|B(s)\| \, \mathrm{d}s < \infty \quad et \quad \int_0^\infty \|f(s)\| \, \mathrm{d}s < \infty.$$

Alors :

(i) Toutes les solutions (maximales) du système $x' = (A + B(t))x + f(t)$ sont bornées.

(ii) Toutes les solutions maximales de $x' = B(t)x$ ont une limite finie en $+\infty$.

(iii) En supposant de plus $t \mapsto e^{tA}$ bornée sur $\mathbb{R}$ tout entier, on a que pour toute solution maximale de $x' = (A + B(t))x$, l'application $y : t \mapsto e^{-tA}x(t)$ a une limite en $+\infty$.

Démonstration. (i) Soit M un majorant de $\|e^{tA}\|$. Par la formule de Duhamel appliquée avec le terme source $B(t)x + f(t)$ dépendant de x, on a

$$x(t) = e^{tA}x(0) + \int_0^t e^{(t-s)A}(B(s)x(s) + f(s)) \, \mathrm{d}s.$$

On note $u(t) = \|x(t)\|$, $F(t) = \int_0^t \|f(s)\| \, \mathrm{d}s$ et $\beta(t) = \int_0^t \|B(s)\| \, \mathrm{d}s$. On obtient alors la majoration suivante :

$$\begin{aligned} u(t) &\leqslant Mu(0) + \int_0^t M(\beta'(s)u(s) + f(s)) \, \mathrm{d}s \\ &= M(u(0) + F(t)) + M\int_0^t \beta'(s)u(s) \, \mathrm{d}s. \end{aligned}$$

D'après le lemme de Gronwall, en notant $c(t) = u(0) + F(t)$,

$$u(t) \leqslant Mc(t) + M\int_0^t c(\tau)\beta'(\tau)e^{\int_\tau^t \beta'(s) \, \mathrm{d}s} \, \mathrm{d}\tau.$$

Il ne reste plus qu'à majorer tous les termes :

$$\begin{aligned} u(t) &\leqslant M(u(0) + F(\infty)) + M(u(0) + F(\infty)) \int_0^t \beta'(\tau)e^{\beta(t)-\beta(\tau)} \, \mathrm{d}\tau \\ &\leqslant M(u(0) + F(\infty)) + M(u(0) + F(\infty))e^{\beta(\infty)} \int_0^t \beta'(\tau) \, \mathrm{d}\tau \\ &\leqslant M(u(0) + F(\infty)) + M(u(0) + F(\infty))e^{\beta(\infty)}\beta(\infty). \end{aligned}$$

(*ii*) D'après (*i*) appliqué à $A = 0$ et $f = 0$, les solutions maximales de $x' = B(t)x$ sont bornées. De plus, on sait que les solutions maximales sont globales, c'est-à-dire ici définies sur $\mathbb{R}_+$. Soit x une telle solution. Alors pour tout $t, s \in \mathbb{R}_+$ avec $s \leqslant t$,

$$x(t) - x(s) = \int_s^t B(\tau)x(\tau)\,\mathrm{d}\tau.$$

En notant C un majorant de $\|x(\tau)\|$ pour $\tau \in \mathbb{R}_+$, on obtient

$$\|x(t) - x(s)\| \leqslant C \int_s^t \|B(\tau)\|\,\mathrm{d}\tau.$$

Ce majorant tend vers 0 quand s et t tendent vers l'infini car $\int_0^\infty \|B(\tau)\|\,\mathrm{d}\tau < \infty$. Donc x satisfait le critère de Cauchy en $+\infty$, ce qui implique que x a une limite finie en $+\infty$.

(*iii*) Soit x une solution maximale de $x' = (A + B(t))x$. L'application $y : t \mapsto e^{-tA}x(t)$ est solution de $y' = e^{-tA}B(t)e^{tA}y$. On note M' un majorant de $\|e^{tA}\|$ pour $t \in \mathbb{R}$. On a :

$$\int_0^\infty \|e^{-sA}B(s)e^{sA}\|\,\mathrm{d}s \leqslant M'^2 \int_0^\infty \|B(s)\|\,\mathrm{d}s < \infty.$$

La question précédente s'applique donc à l'équation différentielle $y' = e^{-tA}B(t)e^{tA}y$, ce qui prouve que y a une limite finie en $+\infty$. □

Application. *Soient $q : \mathbb{R}_+ \to \mathbb{R}$ continue telle que $\int_0^\infty |q(t)|\,\mathrm{d}t < \infty$ et x une solution de*

$$x'' + (1 + q(t))x = 0.$$

Alors il existe $\alpha, \beta \in \mathbb{R}$ tels que $\lim_{t\to\infty} x(t) - \alpha\cos(t) - \beta\sin(t) = 0$.

Démonstration. On pose $X(t) = \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix}$. Alors X est solution du système $X' = (A + B(t))X$ avec

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{et} \quad B(t) = \begin{pmatrix} 0 & 0 \\ -q(t) & 0 \end{pmatrix}.$$

Or

$$e^{tA} = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$

est bornée pour $t \in \mathbb{R}$. Donc la question précédente s'applique : $Y(t) = e^{-tA}X(t)$ a une limite finie $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ à l'infini, *i.e.*

$$e^{-tA}X(t) - \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \underset{t\to\infty}{\longrightarrow} 0.$$

Comme e^{tA} est bornée, en multipliant, on obtient

$$X(t) - e^{tA}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \underset{t\to\infty}{\longrightarrow} 0.$$

En particulier, la première composante, $x(t) - \alpha\cos(t) - \beta\sin(t)$ tend vers 0 quand $t \to +\infty$. □

Formule de Duhamel :

En appliquant la méthode de variation de la constante, on trouve que les solutions de l'équation $u' = Au + b(t)$ sont donnés par la formule de Duhamel :

$$u(t) = e^{(t-t_0)A}u(t_0) + \int_{t_0}^{t} e^{(t-s)A}b(s)\,\mathrm{d}s.$$

La formule de Duhamel peut s'appliquer à des équations dont le terme source (ici $b(t)$) dépend aussi de $u(t)$, ce qui fournit une équation implicite en u.

Théorème (Lemme de Gronwall)**.** *Soit $u \in \mathcal{C}(I, \mathbb{R}_+)$. On suppose qu'il existe $a \in \mathcal{C}(I, \mathbb{R}_+)$ et $c \in \mathcal{C}(I, \mathbb{R})$ tels que pour tout $t \in I$,*

$$u(t) \leqslant c(t) + \int_0^t a(\tau)u(\tau)\,\mathrm{d}\tau.$$

Alors pour tout $t \in I$,

$$u(t) \leqslant c(t) + \int_0^t c(\tau)a(\tau)e^{\int_\tau^t a(s)\,\mathrm{d}s}\,\mathrm{d}\tau.$$

Démonstration. Soit $v(t) = \int_0^t a(\tau)u(\tau)\,\mathrm{d}\tau$. En dérivant, on obtient

$$v'(t) = a(t)u(t) \leqslant a(t)(c(t) + v(t)).$$

On intègre cette inéquation différentielle et on obtient le résultat voulu. □

Définition. Soit λ une valeur propre d'une matrice A. On dit que λ est *semi-simple* si les dimensions de l'espace propre et de l'espace caractéristique associés sont égales, *i.e.* si

$$\dim(\mathrm{Ker}(A - \lambda I_n)) = \dim(\mathrm{Ker}((A - \lambda I_n)^m))$$

où m est la multiplicité de λ. Comme $\mathrm{Ker}(A-\lambda I_n) \subset \mathrm{Ker}((A-\lambda I_n)^m)$, cela implique l'égalité de ces sous-espaces.

On dit que λ est *simple* si

$$\dim(\mathrm{Ker}(A - \lambda I_n)) = \dim(\mathrm{Ker}((A - \lambda I_n)^m)) = 1.$$

Proposition. *L'application $t \mapsto e^{tA}$ est bornée sur $\mathbb{R}_+$ si et seulement si les valeurs propres de A sont toutes de partie réelle négative ou nulle et si les valeurs propres imaginaires pures sont semi-simples. Elle tend vers 0 à l'infini si et seulement si les valeurs propres de A sont toutes de partie réelle strictement négative.*

Démonstration. Voir le document "Théorème de Liapounov". Le cas où les valeurs propres imaginaires pures sont semi-simples s'en déduit facilement. En effet, si λ est une valeur propre imaginaire pure semi-simple, on a $\mathrm{Ker}(A-\lambda I_n) = \mathrm{Ker}((A-\lambda I_n)^m)$, on peut donc prendre $m = 1$ dans la décomposition par le lemme des noyaux. □

Références :

- Benzoni-Gavage - *Calcul différentiel et équations différentielles* - Page 210 (exercice 6.3).

3.4 Densité des fonctions continues et nulle part dérivables

Proposition. *Soit E l'ensemble des fonctions continues sur $[0,1]$ et à valeurs dans F complet. On munit E de la norme infinie. Alors le sous-ensemble A de E des fonctions continues sur $[0,1]$ qui ne sont dérivables en aucun point de $[0,1]$ est dense dans E.*

Démonstration. L'espace E est un espace métrique complet (car F est complet et on travaille avec la norme infinie), c'est donc un espace de Baire. Il nous suffit donc de trouver une suite dénombrable d'ouverts $(O_n)_{n\in\mathbb{N}}$ denses dans E tels que $\cap_{n=1}^{\infty} O_n \subset A$, car par le théorème de Baire, $\cap_{n=1}^{\infty} O_n$ sera dense dans E, donc A aussi. Par passage au complémentaire, on pose $B = {}^cA$ et on va trouver une suite dénombrable $(F_n)_{n\in\mathbb{N}}$ de fermés d'intérieur vide tels que $B \subset \cup_{n=1}^{\infty} F_n$.

Remarquons que B est l'ensemble des fonctions continues dérivables en au moins un point de $[0,1]$. Si f est dérivable en un point x, alors la quantité $\frac{f(y)-f(x)}{y-x}$ est bornée quand $y \to x$. Posons alors pour $n \in \mathbb{N}^*$:

$$F_n = \{f \in E \ / \ \exists x \in [0,1] \text{ tel que } \forall y \in [0,1], |f(x) - f(y)| \leqslant n|x-y|\}.$$

Alors clairement, $B \subset \cup_{n=1}^{\infty} F_n$. Il reste à montrer que F_n est fermé et $\mathring{F}_n = \varnothing$.

Montrons que F_n est fermé. Soit (f_k) une suite de F_n convergeant vers $f \in E$. Pour tout $k \in \mathbb{N}$, il existe x_k tel que pour tout y, $|f_k(y) - f_k(x_k)| \leqslant n|y - x_k|$. Le segment $[0,1]$ étant compact dans $\mathbb{R}$, il existe une suite extraite $(x_{\varphi(k)})$ qui converge vers $x \in [0,1]$. Pour alléger les notations, on note encore (x_k) cette suite extraite.

Soit $y \in [0,1]$. On veut alors montrer que $f_k(y) - f_k(x_k) \to f(y) - f(x)$ quand $k \to \infty$, car ainsi, en passant à la limite dans l'inégalité plus haut, on aura $|f(y) - f(x)| \leqslant n|y-x|$, pour tout y, *i.e.* $f \in F_n$. On a :

$$\begin{aligned} \big|\big(f_k(y) - f_k(x_k)\big) - \big(f(y) - f(x)\big)\big| &\leqslant |f_k(y) - f(y)| + |f_k(x_k) - f(x_k)| \\ &\quad + |f(x_k) - f(x)| \\ &\leqslant 2\|f_k - f\|_\infty + |f(x_k) - f(x)|, \end{aligned}$$

et ce majorant tend vers 0 car $|f(x_k) - f(x)| \to 0$ par continuité de f.

Montrons maintenant que $\mathring{F}_n = \varnothing$. Soit $f \in F_n$. Il s'agit de montrer que toute boule $B(f,\varepsilon)$ rencontre ${}^cF_n$, *i.e.* il existe $g \in E$ tel que $\|f - g\|_\infty \leqslant \varepsilon$ et pour tout $x \in [0,1]$, il existe $y \in [0,1]$ avec $|g(y) - g(x)| > n|y-x|$.

Les polynômes étant denses dans E pour la norme infinie, il existe un polynôme P tel que $\|P - f\|_\infty < \frac{\varepsilon}{2}$. L'idée est d'ajouter à P une fonction g_0 de E assez petite

de sorte à avoir $P + g_0 \in {}^cF_n$. Soit alors $N \in \mathbb{N}^*$. On écrit :

$$[0,1] = \bigcup_{k=0}^{N-1} \left[\frac{k}{N}, \frac{k+1}{N}\right],$$

et on considère la fonction en dents de scie g_0, périodique de période $\frac{1}{N}$, qui vaut :

$$g_0(x) = \begin{cases} \varepsilon N x \text{ pour } 0 \leqslant x \leqslant \frac{1}{2N} \\ \varepsilon - \varepsilon N x \text{ pour } \frac{1}{2N} \leqslant x \leqslant \frac{1}{N}. \end{cases}$$

La fonction g_0 est continue sur $[0,1]$, dérivable sauf en un nombre fini de points, et aux points où elle est dérivable, $|g_0'(x)| = \varepsilon N$. De plus, $\|g_0\|_\infty = \frac{\varepsilon}{2}$. Si on pose $g = P + g_0$, alors $\|f - g\|_\infty \leqslant \|f - P\|_\infty + \|g_0\|_\infty < \varepsilon$. Il reste à voir si $g \in {}^cF_n$. Pour $x, y \in [0,1]$, on effectue la minoration :

$$|g(y) - g(x)| \geqslant |g_0(y) - g_0(x)| - |P(y) - P(x)|.$$

Par l'inégalité des accroissements finis, on a $|P(y) - P(x)| \leqslant M|x - y|$, où $M = \|P'\|_\infty$. Prenons y assez proche de x de sorte que g_0 soit dérivable sur $]x, y[$. Alors par l'égalité des accroissements finis, il existe $c \in]x, y[$ tel que

$$|g_0(y) - g_0(x)| = |g_0'(c)||x - y| = \varepsilon N |x - y|.$$

Ainsi :

$$|g(y) - g(x)| \geqslant (\varepsilon N - M)|x - y|.$$

En reprenant plus haut, on fixe l'entier N tel que $N \geqslant \frac{n+1+M}{\varepsilon}$, ce qui implique :

$$|g(y) - g(x)| \geqslant (n+1)|x - y|.$$

Finalement, pour tout $x \in [0,1]$, on peut trouver $y \in [0,1]$ tel que $|g(y) - g(x)| > n|y - x|$, *i.e.* $g \in {}^cF_n$. □

Exemple (Fonction de Weierstrass)**.** *Soient $a \in]0,1[$ et $b \in \mathbb{N}^*$ entier pair tel que $ab > 1 + \pi$. Alors la fonction f définie pour $x \in \mathbb{R}$ par*

$$f(x) = \sum_{n=1}^{\infty} a^n e^{ib^n \pi x}$$

est continue mais nulle part dérivable.

Démonstration. La série définissant f est normalement convergente, donc f est continue. Soient $x \in \mathbb{R}$ et $h \neq 0$. On regarde le taux décroissement

$$\frac{f(x+h)-f(x)}{h} = \sum_{n=1}^{\infty} a^n e^{ib^n \pi x} \frac{e^{i\pi b^n h}-1}{h}.$$

Prenons $h = \frac{1}{b^m}$ avec $m \in \mathbb{N}$. Comme b est pair, pour $n > m$, on a $e^{i\pi b^n h} - 1 = e^{i\pi b^{n-m}} - 1 = 0$. Alors

$$\frac{f(x+h)-f(x)}{h} = S_{m-1} - 2a^m b^m e^{i\pi b^m x},$$

avec

$$\begin{aligned} |S_{m-1}| &= \left| \sum_{n=1}^{m-1} a^n e^{ib^n \pi x} \frac{e^{i\pi b^n h}-1}{h} \right| \\ &\leqslant \sum_{n=1}^{m-1} a^n \frac{2}{h} \left| \sin\left(\frac{\pi b^n h}{2}\right) \right| \\ &\leqslant \pi \sum_{n=0}^{m-1} a^n b^n = \pi \frac{a^m b^m - 1}{ab-1} \end{aligned}$$

où on a utilisé que $|\sin(x)| \leqslant x$ pour $x \in \mathbb{R}_+$. Mais $ab > 1 + \pi$, donc

$$\pi \frac{a^m b^m - 1}{ab-1} < \pi \frac{a^m b^m - 1}{\pi} = a^m b^m - 1 < a^m b^m.$$

On a donc $|S_{m-1}| \leqslant a^m b^m$. On en déduit :

$$\left|S_{m-1} - 2a^m b^m e^{i\pi b^m x}\right| \geqslant \left|2a^m b^m e^{i\pi b^m x}\right| - |S_{m-1}| \geqslant 2a^m b^m - a^m b^m = a^m b^m.$$

Ainsi,

$$\left| \frac{f\left(x + \frac{1}{b^m}\right) - f(x)}{\frac{1}{b^m}} \right| \geqslant (ab)^m \underset{m\to\infty}{\longrightarrow} \infty,$$

ce qui prouve que f n'est pas dérivable en x. □

Exemple (Fonction de Bolzano)**.** *Sur $[0,1]$, on part de la fonction $x \mapsto x$. On découpe $[0,1]$ en trois, on double les pentes des deux morceaux de la fonction aux extrémités et on les rejoint par une ligne droite. On recommence à l'infini.*

Exemple (Fonction de Van der Waerden)**.** *Pour $x \in \mathbb{R}$, on note $\{x\}$ la distance de x à l'entier le plus proche. Alors la fonction f définie par*

$$f(x) = \sum_{n=0}^{\infty} \frac{\{10^n x\}}{10^n}$$

est continue sur $\mathbb{R}$ mais dérivable en aucun point de $\mathbb{R}$. Noter que $x \mapsto \{x\}$ est 1-périodique et continue sur $[0,1]$ avec $\{0\} = \{1\} = 0$, donc continue sur $\mathbb{R}$. Par conséquent les fonctions $x \mapsto \frac{\{10^n x\}}{10^n}$ sont continues sur $\mathbb{R}$.

Théorème (Théorème de Baire)**.** *Soit E est un espace métrique complet.*

(i) Si $(O_n)_{n\in\mathbb{N}}$ est une suite d'ouverts denses de E, alors $\cap_{n=1}^{\infty} O_n$ est encore dense dans E.

(ii) Si $(F_n)_{n\in\mathbb{N}}$ est une suite de fermés d'intérieur vide de E, alors $\cup_{n=1}^{\infty} F_n$ est encore d'intérieur vide dans E.

Démonstration. On ne va que démontrer que (i) implique (ii) : il s'agit juste de passer au complémentaire en remarquant qu'une partie A est dense si et seulement si son complémentaire est d'intérieur vide. En effet, A est dense si et seulement si tout ouvert rencontre A, ce qui est équivalent à ce qu'aucun ouvert n'est contenu dans le complémentaire de A, et donc à ce que ${}^cA$ soit d'intérieur vide. On remarque aussi que

$$ {}^c\left(\bigcup_{n=1}^{\infty} F_n\right) = \bigcap_{n=1}^{\infty} {}^cF_n. $$

En effet, $x \in {}^c(\cup_{n=1}^{\infty} F_n)$ équivaut à $x \notin \cup_{n=1}^{\infty} F_n$, qui est équivalent à $x \notin F_n$ pour tout n, équivalent à $x \in {}^cF_n$ pour tout n, équivalent à $x \in \cap_{n=1}^{\infty} {}^cF_n$. □

Références :

- Zuily et Queffélec - *Eléments d'analyse (2ème édition)* - Page 263.
- Valiron - Page 160 (pour l'exemple de la fonction de Weierstrass).

3.5 Densité des polynômes orthogonaux

Théorème. *Soient I un intervalle de $\mathbb{R}$ et ρ une fonction poids. On suppose qu'il existe $\alpha > 0$ tel que*

$$\int_I e^{\alpha|x|}\rho(x)\,\mathrm{d}x < \infty.$$

Alors les polynômes orthogonaux associés à ρ forment une base hilbertienne de $L^2(I, \rho\mathrm{d}\lambda)$, espace de Hilbert muni du produit scalaire

$$(f, g) = \int_I f(x)g(x)\rho(x)\,\mathrm{d}x.$$

Démonstration. On commence par noter que tout polynôme appartient à $L^2(I, \rho)$, donc en particulier les polynômes orthogonaux : par définition de la fonction poids, $x \mapsto x^n \in L^1(I, \rho)$ pour tout n, donc les carrés de ces fonctions aussi. Le résultat s'ensuit par linéarité.

Par définition, les polynômes orthogonaux P_n associés à ρ forment une famille orthonormale pour le produit scalaire $(\cdot, \cdot)$. Pour prouver qu'ils forment une base hilbertienne, nous allons montrer que

$$\{x \mapsto P_n(x),\, n \in \mathbb{N}\}^\perp = \{0\},$$

ce qui démontrera le résultat (c'est l'une des caractérisations des bases hilbertiennes lorsque la famille est orthonormée). Mais les X^k s'expriment sur les P_n et réciproquement, c'est donc équivalent à montrer que $\{x \mapsto x^n,\, n \in \mathbb{N}\}^\perp = \{0\}$.

Soit alors $f \in L^2(I, \rho)$ telle que $\int_I x^n f(x)\rho(x)\,\mathrm{d}x = 0$ pour tout $n \in \mathbb{N}$. On veut prouver que $f = 0$ presque partout.

On définit la fonction φ sur $\mathbb{R}$ par :

$$\varphi(x) = \begin{cases} f(x)\rho(x) \text{ si } x \in I \\ 0 \text{ sinon.} \end{cases}$$

On a la majoration évidente suivante :

$$|f(x)|\rho(x) \leqslant \left(1 + |f(x)|^2\right)\rho(x),$$

pour tout $x \in I$. Comme ρ et $f^2\rho$ sont intégrables sur I, on obtient $\varphi \in L^1(\mathbb{R})$. On peut donc considérer la transformée de Fourier de φ :

$$\hat{\varphi}(\omega) = \int_I f(x)e^{-i\omega x}\rho(x)\,\mathrm{d}x$$

pour $\omega \in \mathbb{R}$.

On pose maintenant $B_\alpha = \left\{z \in \mathbb{C} \,/\, |\mathrm{Im}(z)| < \frac{\alpha}{2}\right\}$ et on va montrer que $\hat{\varphi}$ se prolonge en une fonction holomorphe sur B_α. Soient $g(z,x) = e^{-izx} f(x)\rho(x)$ pour $x \in I$ et $z \in B_\alpha$, et

$$F(z) = \int_I g(z,x)\,\mathrm{d}x.$$

Montrons que F est bien définie et holomorphe sur B_α. On vérifie pour cela les hypothèses du théorème d'holomorphie sous le signe intégral :

(i) Pour tout $x \in I$, $z \mapsto g(z,x)$ est holomorphe.

(ii) Pour tout $z \in B_\alpha$, $x \mapsto g(z,x)$ est mesurable (comme produit de fonctions mesurables).

(iii) Pour tout $z \in B_\alpha$, $|g(z,x)| \leqslant e^{\frac{\alpha}{2}|x|}|f(x)|\rho(x)$ qui est une fonction intégrable sur I et indépendante de z. En effet, par l'inégalité de Cauchy-Schwarz :

$$\int_I e^{\frac{\alpha}{2}|x|}|f(x)|\rho(x)\,\mathrm{d}x \leqslant \left(\int_I e^{\alpha|x|}\rho(x)\,\mathrm{d}x\right)^{\frac{1}{2}} \left(\int_I |f(x)|^2\rho(x)\,\mathrm{d}x\right)^{\frac{1}{2}} < \infty.$$

La fonction F est donc holomorphe sur B_α et pour tout $n \in \mathbb{N}$,

$$F^{(n)}(z) = \int_I \frac{\partial^n g(z,x)}{\partial z^n}\,\mathrm{d}x = (-i)^n \int_I x^n e^{-izx} f(x)\rho(x)\,\mathrm{d}x.$$

En évaluant en zéro, on obtient :

$$F^{(n)}(0) = (-i)^n \int_I x^n f(x)\rho(x)\,\mathrm{d}x = 0$$

par l'hypothèse faite sur f.

Par unicité du développement en série entière, on en déduit que F est nulle sur un voisinage de 0. Par le théorème du prolongement analytique, il s'ensuit que F est nulle sur tout le connexe B_α. En particulier, F est nulle sur l'axe réel, *i.e.* $\hat{\varphi} = 0$. Par injectivité de la transformée de Fourier, cela implique $\varphi = 0$ dans $L^1(\mathbb{R})$. Comme $\rho(x) > 0$ pour tout x, on en déduit que $f(x) = 0$ pour presque tout $x \in I$. □

On peut donner l'exemple suivant pour justifier que l'hypothèse d'existence d'un $\alpha > 0$ tel que

$$\int_I e^{\alpha|x|}\rho(x)\,\mathrm{d}x < \infty$$

est nécessaire.

Exemple. *Soient $I =]0,\infty[$ et la fonction de poids $\rho(x) = x^{-\ln(x)}$. Alors les polynômes orthogonaux associés au poids ρ ne forment pas une base hilbertienne de $L^2(I,\rho)$.*

Démonstration. Soit $f : x \mapsto \sin(2\pi \ln(x))$ définie sur I. En notant $g_n : x \mapsto x^n$, on calcule :

$$(f, g_n) = \int_0^\infty f(x)x^n\rho(x)\,\mathrm{d}x = e^{\frac{(n+1)^2}{4}} \int_{-\infty}^\infty e^{-\left(y-\frac{n+1}{2}\right)^2} \sin(2\pi y)\,\mathrm{d}y$$

par le changement de variables $y = \ln(x)$ qui est bien un $\mathcal{C}^1$-difféomorphisme de $\mathbb{R}_+^*$ sur $\mathbb{R}$. Par le deuxième changement de variables affine $t = y - \frac{n+1}{2}$, on obtient

$$(f, g_n) = (-1)^n e^{\frac{(n+1)^2}{4}} \int_{-\infty}^\infty \sin(2\pi t)e^{-t^2}\,\mathrm{d}t = 0$$

car la fonction intégrée est impaire.

On en conclut que la famille des g_n n'est pas dense dans $L^2(I, \rho)$. En effet, si P est un polynôme,

$$\|f - P\|_2^2 = (f - P, f - P) = \|f\|_2^2 + \|P\|_2^2 \geqslant \|f\|_2^2 > 0,$$

on ne pourra donc pas approcher f par un polynôme à ε près. La famille des polynômes orthogonaux associés au poids ρ n'est alors pas dense dans $L^2(I, \rho)$, ce n'est donc pas une base hilbertienne. □

Définition (Fonction poids)**.** Soit I un intervalle de $\mathbb{R}$. On appelle fonction poids une fonction $\rho : I \to \mathbb{R}$ mesurable, strictement positive et telle que

$$\int_I |x|^n \rho(x)\,\mathrm{d}x < \infty$$

pour tout $n \in \mathbb{N}$.

Remarque. Par définition de la fonction poids, la mesure $\rho\mathrm{d}\lambda$ sur I est finie. En particulier, on a l'emboîtement décroissant des espaces $L^p(I, \rho\mathrm{d}\lambda)$, *i.e.* $L^q(I, \rho\mathrm{d}\lambda) \subset L^p(I, \rho\mathrm{d}\lambda)$ si $q \geqslant p$.

Définition (Polynôme orthogonaux)**.** On définit le k-ième polynôme orthogonal associé au poids ρ par le polynôme unitaire de norme 1 qui dirige la droite vectorielle orthogonale à $\mathbb{R}_{k-1}[X]$ dans $\mathbb{R}_k[X]$. Cette famille s'obtient également en appliquant le procédé d'orthonormalisation de Gram-Schmidt sur la base canonique de $\mathbb{R}[X]$. Les polynômes orthogonaux forment une famille échelonnée qui existe et est unique.

Théorème (Caractérisation des bases hilbertiennes)**.** *Soient $\big(H, (\cdot, \cdot)\big)$ un espace de Hilbert séparable et $(e_n)_{n\in\mathbb{N}}$ une famille orthonormée de H. Les propriétés suivantes sont équivalentes :*

(i) La famille orthonormée (e_n) est une base hilbertienne, i.e.

$$H = \overline{\mathrm{Vect}(e_n,\, n \in \mathbb{N})}.$$

(ii) Pour tout $x \in H$, $x = \sum_{n=0}^{\infty}(x, e_n)e_n$, ce qui signifie

$$\lim_{N\to\infty} \left\| x - \sum_{n=0}^{N}(x, e_n)e_n \right\| = 0.$$

(iii) Pour tout $x \in H$, $\|x\|^2 = \sum_{n=0}^{\infty} |(x, e_n)|^2$.

(iv) On a : $\{e_n,\, n \in \mathbb{N}\}^{\perp} = \{0\}$.

Théorème. *La transformation de Fourier*

$$\begin{array}{rccc} \mathcal{F}: & L^1(\mathbb{R}) & \longrightarrow & L^\infty(\mathbb{R}) \\ & f & \longmapsto & \hat{f} : \omega \mapsto \int_{\mathbb{R}} f(x)e^{-i\omega x}\,\mathrm{d}x \end{array}$$

est un morphisme d'algèbres de la $\mathbb{C}$-algèbre de Banach $(L^1, +, \cdot, *, \|\cdot\|_1)$ sur la $\mathbb{C}$-algèbre normée $(L^\infty, +, \cdot, \times, \|\cdot\|_\infty)$. C'est un opérateur injectif dont l'image est dense dans L^∞. Il est de plus continu et sa norme d'opérateur vaut 1.

Références :

– Beck, Malick et Peyré - *Objectif agrégation* - Page 140.

3.6 Ellipsoïde de John-Loewner

Théorème. *Tout compact K de $\mathbb{R}^n$ contenant 0 dans son intérieur est contenu dans un unique ellipsoïde de volume minimal.*

Démonstration. Tout ellipsoïde de $\mathbb{R}^n$ centré en 0 est une boule unité fermée pour un produit scalaire défini par $S \in \mathcal{S}_n^{++}(\mathbb{R})$, qu'on note B_S. Si B est la boule unité fermée canonique de $\mathbb{R}^n$, alors en notant $A = \sqrt{S^{-1}}$,

$$x \in B \Leftrightarrow {}^t xx \leqslant 1 \Leftrightarrow {}^t(Ax)S(Ax) \leqslant 1 \Leftrightarrow Ax \in B_S.$$

Par le changement de variables $\varphi : B \to B_S$, $x \mapsto Ax$, $\mathcal{C}^1$-difféomorphisme, on calcule le volume $V(B_S)$ de B_S :

$$V(B_S) = \int_{B_S} \mathrm{d}\lambda(x) = \int_B |J_\varphi(x)|\,\mathrm{d}\lambda(x) = |\det(A)|V(B) = \frac{V(B)}{\sqrt{\det(S)}}.$$

Soit l'application

$$\begin{array}{rccc} \mu : & \mathcal{S}_n^{++}(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ & S & \longmapsto & \frac{1}{\sqrt{\det(S)}}. \end{array}$$

Le but est de minimiser μ sur l'ensemble des $S \in \mathcal{S}_n^{++}(\mathbb{R})$ telles que $K \subset B_S$. On va montrer que μ est strictement convexe sur le convexe $\mathcal{S}_n^{++}(\mathbb{R})$. Soient $R, S \in \mathcal{S}_n^{++}(\mathbb{R})$. Par le théorème de pseudo-réduction simultanée pour des matrices symétriques définies positives, il existe $P \in \mathrm{GL}_n(\mathbb{R})$ telle que

$$R = {}^t PP \quad \text{et} \quad S = {}^t P\,\mathrm{Diag}(s_1, \ldots, s_n)\,P$$

où $s_i > 0$ pour tout i. Soit $t \in [0,1]$. Pour y voir clair dans nos raisonnements de convexité, on note $r_i = 1$ pour tout i, de sorte que $R = {}^t P\,\mathrm{Diag}(r_1, \ldots, r_n)\,P$. On a

$$\begin{aligned} \mu(tR + (1-t)S) &= \mu\big({}^t P\,\mathrm{Diag}(tr_1 + (1-t)s_1, \ldots, tr_n + (1-t)s_n)\,P\big) \\ &= \left(\det(P)^2 \prod_{i=1}^n (tr_i + (1-t)s_i)\right)^{-\frac{1}{2}} \\ &= \frac{1}{|\det(P)|} \prod_{i=1}^n \exp\left(-\frac{1}{2}(\ln(tr_i + (1-t)s_i)\right) \\ &\leqslant \frac{1}{|\det(P)|} \prod_{i=1}^n \exp\left(-\frac{1}{2}(t\ln(r_i) + (1-t)\ln(s_i))\right), \end{aligned}$$

par concavité du logarithme. Ensuite,

$$|\det(P)|\mu(tR + (1-t)S) \leqslant e^{-\frac{1}{2}\left(t\ln\left(\prod_{i=1}^n r_i\right) + (1-t)\ln\left(\prod_{i=1}^n s_i\right)\right)}.$$

Or $x \mapsto e^{-\frac{1}{2}x}$ est convexe, donc

$$\begin{aligned} |\det(P)|\mu(tR+(1-t)S) &\leqslant te^{-\frac{1}{2}\ln\left(\prod_{i=1}^n r_i\right)} + (1-t)e^{-\frac{1}{2}\ln\left(\prod_{i=1}^n s_i\right)} \\ &= t|\det(P)|\mu(R) + (1-t)|\det(P)|\mu(S), \end{aligned}$$

car $\prod_{i=1}^n r_i = \frac{\det(R)}{\det(P)^2}$. On a donc la convexité de μ. De plus, la stricte concavité du logarithme entraîne que s'il existe i tel que $r_i \neq s_i$, alors l'inégalité est stricte. On a donc égalité si et seulement si $r_i = s_i$ pour tout i, *i.e.* $R = S$. Donc μ est strictement convexe.

Comme K est compact, il est borné : il existe $r > 0$ tel que $K \subset rB$. Mais $rB = B_{S_0}$ où $S_0 = \frac{1}{r^2} I_n$. On considère alors l'ensemble

$$E = \{S \in \mathcal{S}_n^{++}(\mathbb{R}) \,/\, K \subset B_S \text{ et } \mu(S) \leqslant \mu(S_0)\}.$$

L'ensemble E est :

(i) **Non vide :** $S_0 \in E$.

(ii) **Convexe :** car μ est convexe. On vérifie que si $S_1, S_2 \in E$, $tS_1+(1-t)S_2 \in E$: $\mu(tS_1 + (1-t)S_2) \leqslant \mu(S_0)$ par convexité de μ. De plus, pour tout $x \in K$, ${}^t x(tS_1 + (1-t)S_2)x \leqslant 1$, donc $K \subset B_{tS_1+(1-t)S_2}$.

(iii) **Fermé :** le seul point non trivial est que la limite d'une suite d'éléments de E reste *définie* positive, ce qui est assuré par la condition $\mu(S) \leqslant \mu(S_0)$ qui implique $\det(S) \geqslant \det(S_0) > 0$.

(iv) **Borné :** comme 0 est un point intérieur de K, il existe $\lambda > 0$ tel que K contienne λB. Soit $S \in E$. On a $\lambda B \subset K \subset B_S$. Si $x \in B$, on a $\lambda x \in B_S$, *i.e.* ${}^t(\lambda x)S(\lambda x) \leqslant 1$, *i.e.* $\|\sqrt{S}x\|^2 \leqslant \frac{1}{\lambda^2}$. Donc $\|\sqrt{S}\| = \sup_{\|x\|\leqslant 1} \|\sqrt{S}x\| \leqslant \frac{1}{\lambda}$. Il s'ensuit alors que $\|S\| \leqslant \|\sqrt{S}\|\|\sqrt{S}\| \leqslant \frac{1}{\lambda^2}$.

(v) **Compact :** car fermé borné dans $\mathcal{M}_n(\mathbb{R})$ de dimension finie.

La fonction μ est continue (le déterminant et la racine carrée sont continues) sur le compact E, donc y atteint son minimum. L'unicité du minimum est acquise par la stricte convexité de μ sur le convexe E. □

Application. *Tout sous-groupe compact de* $\mathrm{GL}_n(\mathbb{R})$ *est conjugué à un sous-groupe de* $\mathcal{O}_n(\mathbb{R})$.

Démonstration. Soit G un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$. On note B la boule unité fermée de $\mathbb{R}^n$ et on pose

$$K = \bigcup_{A \in G} AB.$$

Alors $K = \mathrm{Im}(\varphi)$ avec

$$\begin{array}{rccc} \varphi : & G \times B & \longrightarrow & \mathbb{R}^n \\ & (A, X) & \longmapsto & AX. \end{array}$$

L'application φ est continue sur le compact $G \times B$, donc K est compact. De plus, K contient 0 dans son intérieur car $B \subset K$. D'après le théorème, K est donc contenu dans un unique ellipsoïde B_S de volume minimal.

On va montrer que $G \subset \mathcal{O}(S) = \{M \in \mathcal{M}_n(\mathbb{R}) \,/\, {}^tMSM = S\}$ le groupe orthogonal associé à la matrice symétrique définie positive S.

Soit $M \in G$. D'une part, on a immédiatement $MK = K$ d'après la définition de K. D'autre part, on considère la suite $(M^p)_{p \in \mathbb{N}}$ d'éléments de G. Comme G est compact, il existe une suite extraite $(M^{p_k})_{k \in \mathbb{N}}$ qui converge vers $N \in G$. Donc $|\det(M)| \geqslant 1$ car sinon on aurait $\det(N) = \lim_{k \to \infty} \det(M)^{p_k} = 0$ alors que N est inversible. De plus, l'application det est continue sur le compact G, donc y est bornée, et donc nécessairement $|\det(M)| \leqslant 1$. Finalement, $|\det(M)| = 1$.

Soit $R = {}^tMSM$. Alors :

(i) $R \in \mathcal{S}_n^{++}(\mathbb{R})$. En effet, ${}^txRx = {}^t(Mx)S(Mx) \geqslant 0$ car $S \in \mathcal{S}_n^{++}(\mathbb{R})$, et on égalité si et seulement si $Mx = 0$, *i.e.* $x = 0$ puisque M est inversible.

(ii) $K \subset B_R$. En effet, si $x \in K$, ${}^txRx = {}^t(Mx)S(Mx) \leqslant 1$ car $Mx \in MK = K \subset B_S$.

(iii) $\det(R) = \det(B)^2 \det(S) = \det(S)$ car $|\det(B)| = 1$. Donc $\mu(R) = \mu(S)$ où $\mu : Q \mapsto \frac{1}{\sqrt{\det(Q)}}$ est la fonction définie dans la démonstration du théorème.

Par unicité de l'ellipsoïde de volume minimal contenant K, on a donc $R = S$. Ainsi ${}^tMSM = S$, ce qui signifie que $M \in \mathcal{O}(S)$. □

Remarque. Le résultat reste vrai dans un espace vectoriel réel E de dimension finie quelconque avec la même démonstration : il suffit de munir E d'une structure euclidienne quelconque.

Théorème (Pseudo-réduction simultanée)**.** *Soient S_1 et S_2 deux matrices symétriques réelles avec S_1 définie positive. Alors il existe $P \in \mathrm{GL}_n(\mathbb{R})$ et D diagonale réelle telles que*

$$S_1 = {}^tPP \quad et \quad S_2 = {}^tPDP.$$

Démonstration. Comme S_1 est symétrique réelle, il existe $A \in \mathcal{O}_n(\mathbb{R})$ et D_1 diagonale telles que $S_1 = {}^tAD_1A$, et comme elle est définie positive, on peut écrire $D_1 = D_1'^2$. Alors en posant $B = D_1'A$, on a

$$S_1 = {}^tAD_1'^2A = {}^t(D_1'A)D_1'A = {}^tBB.$$

Ensuite, la matrice ${}^tB^{-1}S_2B^{-1}$ est symétrique réelle, il existe donc $C \in \mathcal{O}_n(\mathbb{R})$ et D_2 diagonale telles que ${}^tB^{-1}S_2B^{-1} = {}^tCD_2C$. En posant $P = CB$, on obtient

$$S_2 = {}^tB{}^tCD_2CB = {}^tPD_2P.$$

Enfin, on vérifie que

$${}^tPP = {}^tB{}^tCCB = {}^tBB = S_1.$$

□

Références :

- Alessandri - *Thèmes de géométrie.*
- Francinou, Gianella, Nicolas - *Oraux X-ENS, Algèbre 3* - Page 229 (si on veut quelques variations dans la méthode).

3.7 Equation de la chaleur

Considérons une barre métallique. Connaissant, à l'instant initial, la température en chaque point de la barre et, à tout instant, la température aux deux extrémités, peut-on déterminer, à tout moment et en tout point, la température de la barre ? Ce problème a été modélisé et la température u qui est une fonction du temps t et du point x est solution d'une équation de la chaleur.

On peut supposer que la barre est le segment $[0, L]$. Notons $Q =]0, L[\times]0, \infty[$, $\overline{Q} = [0, L] \times [0, \infty[$.

On peut résoudre le problème simplifié suivant à l'aide des séries de Fourier : on cherche s'il existe une fonction u continue sur $\overline{Q}$, de classe $\mathcal{C}^2$ sur Q telle que :

$$\begin{cases} \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0 \text{ sur } Q, \\ u(0,t) = u(L,t) = 0, \text{ pour tout } t \in [0, \infty[, \\ u(x,0) = h(x), \text{ pour tout } x \in [0, L], \end{cases}$$

où h est une fonction $\mathcal{C}^1$ sur $[0, L]$ telle que $h(0) = h(L) = 0$.

Le problème est simplifié car ici la température aux extrémités est imposée constante (l'équation étant linéaire, on a fixé cette constante à 0).

Proposition. *Le problème précédent admet une unique solution u. De plus, on a $u \in \mathcal{C}^\infty(Q)$.*

Démonstration. Nous ne montrerons que l'existence. Cherchons une solution de la forme $u(x,t) = f(x)g(t)$. Alors l'équation différentielle vérifiée par u est équivalente à

$$f(x)g'(t) = f''(x)g(t).$$

On peut essayer de chercher u telle que $u(x,t) \neq 0$ pour tout $x \in]0, L[$ et tout $t \in]0, \infty[$. On a alors :

$$\frac{f''(x)}{f(x)} = \frac{g'(t)}{g(t)}.$$

Comme $\frac{g'(t)}{g(t)}$ ne dépend pas de x, on en déduit qu'il existe $\lambda \in \mathbb{R}$ tel que $f''(x) = \lambda f(x)$ pour tout $x \in]0, L[$. On a aussi $g'(t) = \lambda g(t)$ pour tout $t \in]0, \infty[$.

Si $\lambda > 0$, la solution générale pour f est

$$f(x) = Ae^{\sqrt{\lambda}x} + Be^{-\sqrt{\lambda}x}.$$

La condition $u(0,t) = u(L,t) = 0$ conduit à $A + B = 0$ et $Ae^{\sqrt{\lambda}L} + Be^{-\sqrt{\lambda}L} = 0$, puis $A = B = 0$. On aurait donc $u = 0$ qui ne satisfait pas $u(x,0) = h(x)$.

Si $\lambda = 0$, on a $f(x) = Ax + B$ et la condition $u(0,t) = u(L,t) = 0$ donne encore $A = B = 0$.

Considérons maintenant le cas $\lambda = -\alpha^2 < 0$. Alors

$$f(x) = A\cos(\alpha x) + B\sin(\alpha x) \quad \text{et} \quad g(t) = Ce^{-\alpha^2 t}.$$

La condition $u(0,t) = 0$ conduit à $A = 0$ et la condition $u(L,t) = 0$ donne $B\sin(\alpha L) = 0$, donc $\alpha = \frac{n\pi}{L}$ avec $n \in \mathbb{Z}$. On a donc une famille de solutions :

$$u_n(x,t) = b_n \sin\left(\frac{n\pi}{L}x\right) e^{-\frac{n^2\pi^2}{L^2}t}.$$

Afin de pouvoir avoir $u(x,0) = h(x)$, on va envisager une somme infinie de ces solutions :

$$u(x,t) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}x\right) e^{-\frac{n^2\pi^2}{L^2}t}$$

et essayer d'écrire $h(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}x\right)$. Les b_n doivent alors être les coefficients de Fourier d'une fonction périodique $\tilde{h}$ égale à h sur $[0,L]$. Il faudra aussi que la série de Fourier de $\tilde{h}$ converge vers $\tilde{h}$ sur $[0,L]$.

Soit h_1 la fonction définie sur $[-L,L]$ par

$$h_1(x) = \begin{cases} h(x) \text{ si } x \in [0,L] \\ -h(-x) \text{ si } x \in [-L,0]. \end{cases}$$

Soit $\tilde{h}$ la fonction $2L$-périodique égale à h_1 sur $[-L,L]$. Comme $h(L) = 0$, $\tilde{h}$ est continue sur $\mathbb{R}$ et $\mathcal{C}^1$ par morceaux. La série de Fourier de $\tilde{h}$ converge donc normalement et

$$\tilde{h}(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}x\right) \text{ avec } b_n = \frac{2}{L}\int_0^L h(x)\sin\left(\frac{n\pi}{L}x\right)\,\mathrm{d}x,$$

pour tout $x \in \mathbb{R}$ (les coefficients de Fourier a_n sont nuls car $\tilde{h}$ est impaire). On note que la convergence normale signifie ici

$$\sum_{n=1}^{\infty} |b_n| < \infty.$$

Alors la fonction u définie par

$$u(x,t) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}x\right) e^{-\frac{n^2\pi^2}{L^2}t}$$

est définie et continue sur $\overline{Q}$. En effet, la série la définissant est normalement convergente (son terme général est majoré par $|b_n|$). La condition $u(0,t) = u(L,t) = 0$ est trivialement vérifiée. De plus, si $t \in [\varepsilon, M]$ avec $\varepsilon > 0$, $k, k_1, k_2 \in \mathbb{N}$ avec $k_1 + k_2 = k$, on a :

$$\left|\frac{\partial^k u_n}{\partial x^{k_1}\partial t^{k_2}}\right| \leqslant C_k |b_n| n^{2k} e^{-\frac{n^2\pi^2}{L^2}\varepsilon}$$

qui est le terme général d'une série convergente. La série $\sum_{n\geqslant 1}\frac{\partial^k u_n}{\partial x^{k_1}\partial t^{k_2}}$ converge donc normalement, donc uniformément (sur tout compact de Q), et donc $u\in\mathcal{C}^k(Q)$. On en déduit que $u\in\mathcal{C}^\infty(Q)$. □

Pour montrer l'unicité, nous avons besoin du lemme suivant :

Lemme (Principe du maximum pour l'équation de la chaleur)**.** *Soit u une fonction continue sur $\overline{Q}$ et $\mathcal{C}^2$ sur Q telle que $Pu(x,t)\geqslant 0$ sur Q, où $P=\frac{\partial^2}{\partial x^2}-\frac{\partial}{\partial t}$. Soient $T>0$ et $K=[0,L]\times[0,T]$. Alors*

$$\sup_K u=\sup_{K\cap\partial Q} u.$$

Démonstration. Soient $\varepsilon>0$ et $u_\varepsilon(x,t)=u(x,t)+\varepsilon x^2$. On a $Pu_\varepsilon=Pu+2\varepsilon\geqslant 2\varepsilon$ sur Q. Soit $m_\varepsilon=(x_\varepsilon,t_\varepsilon)$ un point de K où u_ε atteint son maximum sur K et supposons que $m_\varepsilon\notin K\cap\partial Q$. Alors $x_\varepsilon\in]0,L[$. En notant $g:x\mapsto u_\varepsilon(x,t_\varepsilon)$, on a

$$g(x)-g(x_\varepsilon)=g'(x_\varepsilon)(x-x_\varepsilon)+g''(x_\varepsilon)\frac{(x-x_\varepsilon)^2}{2}+o\left((x-x_\varepsilon)^2\right).$$

Comme g atteint son maximum en x_ε, $g(x)-g(x_\varepsilon)\leqslant 0$, et en regardant l'égalité pour x assez proche de x_ε (des deux côtés de x_ε, grâce au fait que $x_\varepsilon\in]0,L[$), ceci implique $g'(x_\varepsilon)=0$ et $g''(x_\varepsilon)\leqslant 0$. On obtient donc

$$\frac{\partial u_\varepsilon}{\partial x}(m_\varepsilon)=0\quad\text{et}\quad\frac{\partial^2 u_\varepsilon}{\partial x^2}(m_\varepsilon)\leqslant 0.$$

D'autre part, comme $t_\varepsilon\in]0,T]$, on peut toujours envisager son taux d'accroissement en t_ε par valeur inférieure :

$$\frac{\partial u_\varepsilon}{\partial t}(m_\varepsilon)=\lim_{\substack{h\to 0\\ h<0}}\frac{u_\varepsilon(x_\varepsilon,t_\varepsilon+h)-u_\varepsilon(x_\varepsilon,t_\varepsilon)}{h}\geqslant 0.$$

Il découle de tout cela que $Pu_\varepsilon(m_\varepsilon)\leqslant 0$, ce qui contredit le fait que $Pu_\varepsilon\geqslant 2\varepsilon$ sur Q.

Donc $m_\varepsilon\in K\cap\partial Q$ et :

$$\sup_{K\cap\partial Q} u\leqslant\sup_K u\leqslant\sup_K u_\varepsilon=\sup_{K\cap\partial Q} u_\varepsilon\leqslant\sup_{K\cap\partial Q} u+\varepsilon L^2.$$

En faisant tendre ε vers 0, on obtient

$$\sup_K u=\sup_{K\cap\partial Q} u.$$

□

Démontrons maintenant l'unicité de notre problème. Soient u et v deux solutions et $w = v - u$. Alors w est continue sur $\overline{Q}$, de classe $\mathcal{C}^2$ sur Q et telle que $Pw = 0$. De plus, $w = 0$ sur ∂Q du fait des conditions initiales vérifiées par u et v. Fixons $T > 0$. D'après le lemme, on a donc $w(x,t) \leqslant 0$ sur $[0,L] \times [0,T]$. Mais $Pw = 0$ implique qu'on a aussi $P(-w) = 0$, et donc toujours d'après le lemme, $w(x,t) = 0$ sur $[0,L] \times [0,T]$. Comme T est arbitraire, on en déduit $w = 0$ sur Q, *i.e.* $u = v$.

Remarque.
1. On peut montrer que le problème reste vrai si h est seulement continue sur $[0,1]$.
2. La méthode utilisée permet de traiter d'autres équations comme l'équation aux ondes :
$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0$$
qui intervient dans tous les phénomènes de propagation d'ondes (son, lumière,...) et l'équation de Laplace :
$$\frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial x^2} = 0$$
qui intervient par exemple en électromagnétisme.

Théorème. *Si f est 2π-périodique, continue et $\mathcal{C}^1$ par morceaux, alors la série de Fourier de f converge normalement et f est égale à sa série de Fourier.*

Démonstration. Comme f est $\mathcal{C}^1$ par morceaux, on a $c_n(f') = inc_n(f)$ et $f' \in L^2(\mathbb{T})$. D'après le théorème de Parseval, la série $\sum |c_n(f')|^2$ converge et
$$\sum_{n\in\mathbb{Z}} |c_n(f')|^2 = \sum_{n\in\mathbb{Z}} n^2 |c_n(f)|^2 = \|f'\|_2^2.$$
L'inégalité de Cauchy-Schwarz donne :
$$\begin{aligned}
\sum_{1\leqslant|k|\leqslant n} |c_k(f)| &= \sum_{1\leqslant|k|\leqslant n} \frac{1}{k} k|c_k(f)| \\
&\leqslant \left(\sum_{1\leqslant|k|\leqslant n} \frac{1}{k^2}\right)^{\frac{1}{2}} \left(\sum_{1\leqslant|k|\leqslant n} k^2|c_k(f)|^2\right)^{\frac{1}{2}} \\
&\leqslant \left(2\sum_{k=1}^{\infty} \frac{1}{k^2}\right)^{\frac{1}{2}} \|f'\|_2 \\
&= \frac{\pi}{\sqrt{3}} \|f'\|_2.
\end{aligned}$$

En faisant tendre n vers l'infini, on obtient :

$$\sum_{n\in\mathbb{Z}} |c_n(f)| \leqslant |c_0(f)| + \frac{\pi}{\sqrt{3}}\|f'\|_2.$$

La série de Fourier de f est donc en particulier normalement convergente. Dans ce cas, la série de Fourier de f converge uniformément (donc est continue), et f étant continue, f est égal à sa série de Fourier (on montre que la série de Fourier de f et f ont les mêmes coefficients de Fourier, donc sont égales car continues). □

Références :

– Zuily et Queffélec - *Eléments d'analyse (2ème édition)* - Page 103.

3.8 Etude du folium de Descartes

Nous allons faire l'étude complète de la courbe C donnée par l'ensemble des $(x, y) \in \mathbb{R}^2$ tels que

$$x^3 + y^3 - 3xy = 0,$$

appelée *folium de Descartes*.

Représentation paramétrique rationnelle :

Si $(x, y) \in C$ et $x = 0$, alors $y = 0$. On suppose maintenant que x et y sont non nuls. On résout

$$\begin{cases} y = tx \\ x^3 + y^3 - 3xy = 0. \end{cases}$$

On obtient l'équation $x^2((1 + t^3)x - 3t) = 0$ et $y = tx$. Comme $x \neq 0$, cela implique

$$(x, y) = \left(\frac{3t}{1 + t^3}, \frac{3t^2}{1 + t^3} \right).$$

En $t = 0$ on retrouve le point $(0, 0)$. On a donc obtenu une représentation paramétrique rationnelle de C.

Symétries :

Les fonctions x et y sont définies sur $\mathbb{R}\backslash\{-1\}$. On remarque que $x\left(\frac{1}{t}\right) = y(t)$ et $y\left(\frac{1}{t}\right) = x(t)$. La courbe est donc symétrique par rapport à la première bissectrice. On peut donc se contenter de faire l'étude sur $I =]-1, 1]$.

Variations :

On calcule les dérivées :

$$x'(t) = 3\frac{1 - 2t^3}{(t^3 + 1)^2} \quad \text{et} \quad y'(t) = 3t\frac{2 - t^3}{(t^3 + 1)^2}.$$

On obtient le tableau de variations suivant :

t	-1		0		$2^{-\frac{1}{3}}$		1
$x(t)$	$-\infty$	$\nearrow$	0	$\nearrow$	$2^{\frac{2}{3}}$	$\searrow$	$\frac{3}{2}$
$y(t)$	$+\infty$	$\searrow$	0	$\nearrow$	$2^{\frac{1}{3}}$	$\nearrow$	$\frac{3}{2}$

On note qu'il n'y a pas de point stationnaire car x' et y' ne s'annulent pas simultanément (pas de point d'inflexion à étudier).

Branches infinies :

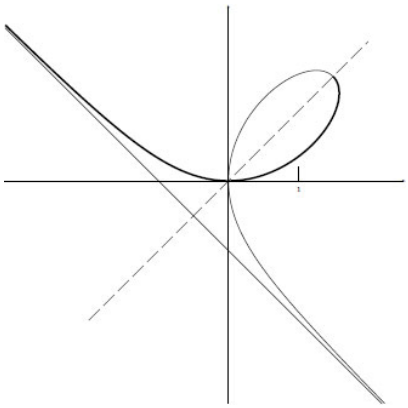
Notons que l'on passe par l'origine en $t = 0$ où on a une tangente horizontale puisque $\frac{y(t)}{x(t)} = t \to 0$.

Lorsque t tend vers -1, on a $\frac{y(t)}{x(t)} = t \to -1$. On étudie donc

$$x(t) + y(t) = 3\frac{t^2+t}{1+t^3} = \frac{3t}{1-t+t^2} \to -1.$$

Donc la droite $y = -x - 1$ est asymptote à C, et la courbe est au-dessus de son aymptote car

$$x(t) + y(t) + 1 = \frac{3t}{t^2-t+1} + 1 = \frac{(t+1)^2}{t^2-t+1} \geqslant 0.$$



Courbe non lisse en 0 :

Comme $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$, le théorème des fonctions implicites ne s'applique pas.

Pour tout voisinage U de 0, le complémentaire de 0 dans $C \cap U$ a au moins 3 composantes connexes : il y en a en effet au moins une dans chacun des trois quadrants ouverts du plan, $x > 0$ et $y > 0$, $x > 0$ et $y < 0$, $x < 0$ et $y > 0$ (que l'on distingue par les signes de $x(t)$ et $y(t)$). Il n'y a donc pas homémorphisme local du folium de Descartes avec une droite. Le folium de Descartes n'est donc pas une sous-variété.

Application du lemme de Morse :

La fonction $f : (x, y) \mapsto x^3 + y^3 - 3xy$ est de classe $\mathcal{C}^3$ sur $\mathbb{R}^2$, et 0 est un point critique de f non dégénéré : en effet, on a $Df(0) = 0$ et $D^2f(0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$ dont les valeurs propres sont -3 et 3. La forme quadratique hessienne $D^2f(0)$ étant de signature $(1, 1)$, le lemme de Morse nous dit qu'il existe un changement de coordonnées local $\varphi : (x, y) \mapsto (u, v)$ de classe $\mathcal{C}^1$ entre deux voisinages de l'origine dans $\mathbb{R}^n$ tel que $\varphi(0) = 0$ et

$$f(x, y) - f(0, 0) = u^2(x, y) - v^2(x, y).$$

Alors $f(x,y)=0$ équivaut à $u^2(x,y)-v^2(x,y)=0$, soit encore

$$u(x,y)+v(x,y)=0 \quad \text{ou} \quad u(x,y)-v(x,y)=0.$$

La courbe de niveau de f se décompose donc en deux courbes C^+ et C^- définies implicitement par $u+v=0$ et $u-v=0$: il y a un point double à l'origine.

Aire délimitée par la boucle :

On choisit la formule suivante pour l'aire délimitée par la boucle, qui va nous donner une intégrale calculable :

$$A=\frac{1}{2}\int_C (x\,\mathrm{d}y - y\,\mathrm{d}x).$$

On note que la boucle est parcourue pour t variant entre 0 et $+\infty$ et que ce sens de parcours donne bien le sens trigonométrique. Ensuite, on tient compte du fait que $y(t)=tx(t)$: on a alors $x\,\mathrm{d}y = x(t)(x(t)+tx'(t))\,\mathrm{d}t$, et $y\,\mathrm{d}x = tx(t)x'(t)\,\mathrm{d}t$. On en déduit :

$$A=\frac{1}{2}\int_0^\infty x(t)^2\,\mathrm{d}t=\frac{1}{2}\int_0^\infty \frac{9t^2}{(1+t^3)^2}\,\mathrm{d}t=-\frac{3}{2}\left[\frac{1}{1+t^3}\right]_0^\infty=\frac{3}{2}.$$

Théorème (Théorème des sous-variétés)**.** *Soient V un sous-ensemble de $\mathbb{R}^n$, $a\in V$ et $d\in\mathbb{N}$. Les quatre propriétés suivantes sont équivalentes :*

(i) *V est lisse en a, de dimension d.*

(ii) (Définition implicite) *Il existe un voisinage ouvert U de a dans $\mathbb{R}^n$ et $n-d$ fonctions $f_i : U\to\mathbb{R}$ de classe $\mathcal{C}^1$ telles que*

$$x\in V\cap U \iff x\in U \text{ et } f_1(x)=0,\dots,f_{n-d}(x)=0,$$

et les différentielles $Df_1(a),\dots,Df_{n-d}(a)$ sont linéairement indépendantes.

(iii) (Graphe) *Il existe un voisinage ouvert U de a dans $\mathbb{R}^n$, un voisinage ouvert U' de $(a_1,\dots,a_d)$ dans $\mathbb{R}^d$ et $n-d$ fonctions $g_i : U'\to\mathbb{R}$ de classe $\mathcal{C}^1$ telles que, après permutation éventuelle des coordonnées x_i,*

$$x\in V\cap U \iff \begin{cases}(x_1,\dots,x_d)\in U'\\ x_{d+1}=g_1(x_1,\dots,x_d),\dots,x_n=g_{n-d}(x_1,\dots,x_d).\end{cases}$$

(iv) (Définition paramétrique) *Il existe un voisinage ouvert U de a dans $\mathbb{R}^n$, un voisinage ouvert Ω de 0 dans $\mathbb{R}^d$ et n fonctions $\varphi_i : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^1$ telles que l'application*

$$\varphi : u = (u_1, \dots, u_d) \mapsto x = (\varphi_1(u), \dots, \varphi_n(u))$$

soit un homéomorphisme de Ω sur $V \cap U$, avec $a = \varphi(0)$ et que la matrice jacobienne $D\varphi(0)$ soit injective (i.e. de rang d).

Si on note $f = (f_1, \dots, f_{n-d})$, $g = (g_1, \dots, g_{n-d})$, $\varphi = (\varphi_1, \dots, \varphi_n)$, l'espace vectoriel tangent en a à V est alors

(ii) *Le noyau de $Df(a)$.*

(iii) *Le graphe de $Dg(a_1, \dots, a_d)$.*

(iv) *L'image de $D\varphi(0)$.*

Calcul de l'aire délimitée par une courbe plane :

On considère une courbe patatoïdale (fermée). On appelle A son point le plus à gauche, d'abscisse a, et B son point le plus à droite, d'abscisse b. Alors on peut voir le patatoïde comme réunion de deux courbes $x \mapsto y_1 = f_1(x)$ et $x \mapsto y_2 = f_2(x)$ (celle "du bas" et celle "du haut"). L'aire délimitée par ce patatoïde est donc donnée par :

$$A = \int_a^b y_2 \,\mathrm{d}x - \int_a^b y_1 \,\mathrm{d}x.$$

Supposons maintenant que le patatoïde est défini paramétriquement par $x = f(t)$ et $y = g(t)$. On suppose que la "courbe du bas" est décrite pour $t \in [t_0, t_1]$ et celle "du haut" pour $t \in [t_1, t_2]$. Alors comme $\mathrm{d}x = f'(t)\mathrm{d}t$,

$$\int_a^b y_2 \,\mathrm{d}x = -\int_{t_1}^{t_2} g(t)f'(t)\,\mathrm{d}t \quad \text{et} \quad \int_a^b y_1 \,\mathrm{d}x = \int_{t_0}^{t_1} g(t)f'(t)\,\mathrm{d}t,$$

et on obtient l'aire de notre boucle par la formule :

$$A = \int_a^b y_2 \,\mathrm{d}x - \int_a^b y_1 \,\mathrm{d}x = -\int_{t_0}^{t_2} g(t)f'(t)\,\mathrm{d}t = -\int_C y \,\mathrm{d}x.$$

La dernière écriture est une intégrale curviligne : on intègre par rapport au paramètre t sur le pourtour C de la boucle, que l'on parcourt dans le sens trigonométrique (un mobile parcourant la courbe voit l'intérieur à sa gauche).

Si on effectue une rotation d'angle $\frac{\pi}{2}$ autour de l'origine, cela revient à changer y en x et x en $-y$. On obtient donc une nouvelle expression de l'aire :

$$A = \int_C x \,\mathrm{d}y.$$

On en déduit également l'expression suivante :

$$A = \frac{1}{2}\int_C (x\,\mathrm{d}y - y\,\mathrm{d}x).$$

Références :

– Rouvière - *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation* - Page 237 et page 269 (pour l'aspect sous-variété).

3.9 Formule des compléments

Lemme. *Pour tout $\alpha \in]0,1[$, on a :*

$$I_\alpha = \int_0^\infty \frac{1}{t^\alpha(1+t)}\,\mathrm{d}t = \frac{\pi}{\sin(\pi\alpha)}.$$

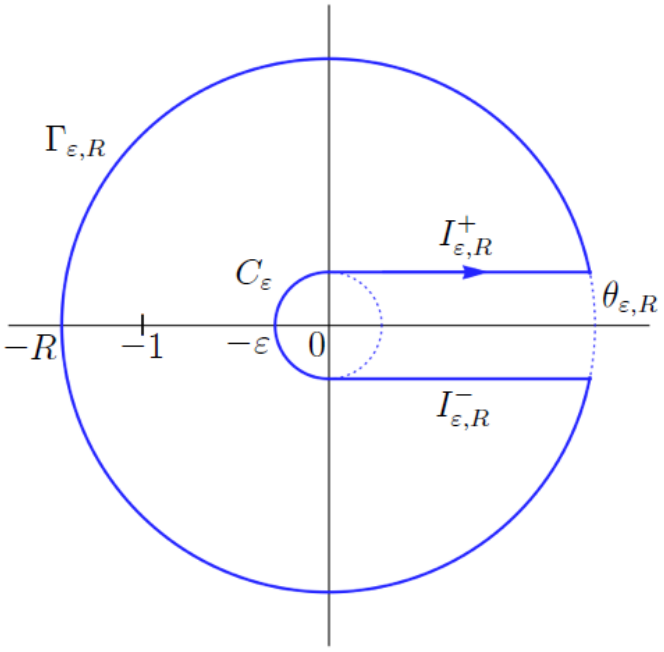
Démonstration. Soit $\alpha \in]0,1[$. On note $\Omega = \mathbb{C}\backslash\mathbb{R}_+$ et on choisit la détermination du logarithme $\log(z) = \ln(r) + it$ pour $z = re^{it} \in \Omega$, $t \in]0,2\pi[$. On définit alors z^α par $z^\alpha = e^{\alpha\log(z)} = r^\alpha e^{i\alpha t}$ pour $z = re^{it} \in \Omega$. La fonction $z \mapsto z^\alpha$ est ainsi holomorphe sur Ω. On introduit ensuite la fonction $f : z \mapsto \frac{1}{z^\alpha(1+z)}$ définie sur $\Omega\backslash\{-1\}$. La fonction f est holomorphe sur $\Omega\backslash\{-1\}$ et possède un pôle simple en -1 dont on calcule le résidu :

$$\mathrm{Res}_{-1}(f) = \lim_{z\to-1}(z+1)f(z) = \lim_{z\to-1}\frac{1}{z^\alpha} = \lim_{\substack{r\to1\\ \theta\to\pi}}\frac{1}{r^\alpha e^{i\alpha\theta}} = e^{-i\pi\alpha}.$$

Pour $0 < \varepsilon < 1 < R$, on note $K_{\varepsilon,R}$ le compact délimité par :

$$\begin{aligned}
C_\varepsilon &= \{z \in \mathbb{C} \,/\, |z| = \varepsilon \text{ et } \mathrm{Re}(z) \leqslant 0\},\\
I^+_{\varepsilon,R} &= \left\{t + i\varepsilon \,/\, 0 \leqslant t \leqslant \sqrt{R^2-\varepsilon^2}\right\},\\
I^-_{\varepsilon,R} &= \left\{t - i\varepsilon \,/\, 0 \leqslant t \leqslant \sqrt{R^2-\varepsilon^2}\right\},\\
\Gamma_{\varepsilon,R} &= \left\{Re^{i\theta} \,/\, \theta \in [-\pi,\pi], |\theta| \geqslant \theta_{\varepsilon,R}\right\},
\end{aligned}$$

où $\theta_{\varepsilon,R} = \arctan\left(\frac{\varepsilon}{\sqrt{R^2-\varepsilon^2}}\right)$.



Le théorème des résidus nous donne :

$$\int_{\partial K_{\varepsilon,R}} f(z)\,\mathrm{d}z = 2\pi i e^{-i\pi\alpha}$$

pour tout ε, R.

Nous allons maintenant calculer la limite de l'intégrale sur chaque morceau de $\partial K_{\varepsilon,R}$ quand $\varepsilon \to 0$. Sur C_ε, on a :

$$\begin{aligned}\left|\int_{C_\varepsilon} f(z)\,\mathrm{d}z\right| &= \left|\int_{\frac{3\pi}{2}}^{\frac{\pi}{2}} f(\varepsilon e^{i\theta}) i\varepsilon\,\mathrm{d}\theta\right| \leqslant \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \frac{\varepsilon}{\varepsilon^\alpha |1+\varepsilon e^{i\theta}|}\,\mathrm{d}\theta \\ &\leqslant \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \frac{\varepsilon}{\varepsilon^\alpha |1-\varepsilon|}\,\mathrm{d}\theta = \frac{\pi\varepsilon}{\varepsilon^\alpha(1-\varepsilon)} = \frac{\pi\varepsilon^{1-\alpha}}{1-\varepsilon} \underset{\varepsilon\to 0}{\longrightarrow} 0.\end{aligned}$$

Sur $\Gamma_{\varepsilon,R}$, on a :

$$\int_{\Gamma_{\varepsilon,R}} f(z)\,\mathrm{d}z = \int_{\theta_{\varepsilon,R}}^{2\pi-\theta_{\varepsilon,R}} iR^{1-\alpha}\frac{e^{i(1-\alpha)\theta}}{1+Re^{i\theta}}\,\mathrm{d}\theta \underset{\varepsilon\to 0}{\longrightarrow} \int_0^{2\pi} iR^{1-\alpha}\frac{e^{i(1-\alpha)\theta}}{1+Re^{i\theta}}\,\mathrm{d}\theta.$$

On étudie maintenant l'intégrale sur $I^+_{\varepsilon,R}$ et sur $I^-_{\varepsilon,R}$. Pour $t \in]0,\infty[$, on a :

$$(t+i\varepsilon)^\alpha \underset{\varepsilon\to 0}{\longrightarrow} t^\alpha \quad \text{et} \quad (t-i\varepsilon)^\alpha \underset{\varepsilon\to 0}{\longrightarrow} t^\alpha e^{2i\pi\alpha},$$

car $\arg((t-i\varepsilon)^\alpha) = \alpha\arg(t-i\varepsilon) = \alpha\left(2\pi - \arctan\left(\frac{\varepsilon}{t}\right)\right) \underset{\varepsilon\to 0}{\longrightarrow} 2\pi\alpha$. D'autre part, on a

$$|f(t+i\varepsilon)| = \left|\frac{1}{(t+i\varepsilon)^\alpha(1+t+i\varepsilon)}\right| \leqslant \frac{1}{t^\alpha(1+t)}.$$

Ce majorant est intégrable sur $I^+_{\varepsilon,R}$ et sur $I^-_{\varepsilon,R}$ et ne dépend pas de ε. On peut appliquer le théorème de convergence dominée :

$$\int_{I^+_{\varepsilon,R}} f(z)\,\mathrm{d}z \underset{\varepsilon\to 0}{\longrightarrow} \int_0^R \frac{1}{t^\alpha(1+t)}\,\mathrm{d}t \quad \text{et} \quad \int_{I^-_{\varepsilon,R}} f(z)\,\mathrm{d}z \underset{\varepsilon\to 0}{\longrightarrow} \int_R^0 \frac{e^{-2i\pi\alpha}}{t^\alpha(1+t)}\,\mathrm{d}t.$$

On fait maintenant tendre ε vers 0 dans la formule donnée par le théorème des résidus. On obtient :

$$(1-e^{-2i\pi\alpha})\int_0^R \frac{1}{t^\alpha(1+t)}\,\mathrm{d}t + \int_0^{2\pi} iR^{1-\alpha}\frac{e^{i(1-\alpha)\theta}}{1+Re^{i\theta}}\,\mathrm{d}\theta = 2i\pi e^{-i\pi\alpha},$$

et cela pour tout $R > 1$. Mais

$$\left|\int_0^{2\pi} iR^{1-\alpha}\frac{e^{i(1-\alpha)\theta}}{1+Re^{i\theta}}\,\mathrm{d}\theta\right| \leqslant \frac{R^{1-\alpha}}{R-1}\int_0^{2\pi}\mathrm{d}\theta \underset{R\to\infty}{\longrightarrow} 0.$$

En faisant maintenant tendre R vers l'infini dans la formule précédente, on trouve :

$$(1-e^{-2i\pi\alpha})I_\alpha = 2i\pi e^{-i\pi\alpha}.$$

On en conclut alors :

$$I_\alpha = \pi\frac{2i}{e^{i\pi\alpha}(1-e^{-2i\pi\alpha})} = \pi\frac{2i}{e^{i\pi\alpha}-e^{-i\pi\alpha}} = \frac{\pi}{\sin(\pi\alpha)}.$$

□

Proposition (Formule des compléments). *Pour tout $z \in \mathbb{C}$ tel que* $\mathrm{Re}(z) \in]0,1[$, *on a :*

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}.$$

Démonstration. On doit prouver l'égalité de fonctions holomorphes sur le domaine $\{z \in \mathbb{C} \,/\, \mathrm{Re}(z) \in]0,1[\}$, il suffit donc de prouver l'égalité pour $z = \alpha \in]0,1[$. En effet, leur fonction différence est une fonction analytique qui aura ainsi un zéro non isolé et sera donc nulle sur tout le domaine (principe des zéros isolés et principe du prolongement analytique).

Soit $\alpha \in]0,1[$. On note $U = \{(s,t) \in \mathbb{R}^2 \,/\, s,t > 0\}$. On a :

$$\begin{aligned}\Gamma(\alpha)\Gamma(1-\alpha) &= \left(\int_0^\infty t^{\alpha-1}e^{-t}\,\mathrm{d}t\right)\left(\int_0^\infty s^{-\alpha}e^{-s}\,\mathrm{d}s\right)\\ &= \int_U \frac{1}{t}\left(\frac{t}{s}\right)^\alpha e^{-(t+s)}\,\mathrm{d}s\mathrm{d}t\end{aligned}$$

par le théorème de Fubini-Tonelli (tout est positif et intégrable). On effectue le changement de variable $\varphi : (s,t) \mapsto \left(s+t, \frac{t}{s}\right)$ d'inverse $\psi : (u,v) \mapsto \left(\frac{u}{1+v}, \frac{uv}{1+v}\right)$ qui définit un $\mathcal{C}^1$-difféomorphisme de U sur U (φ et ψ sont clairement différentiables sur U). On calcule le jacobien de ψ en (u,v) :

$$\begin{aligned}J_\psi\big((u,v)\big) &= \det\begin{pmatrix}\frac{\partial\psi_1}{\partial u}(u,v) & \frac{\partial\psi_1}{\partial v}(u,v)\\ \frac{\partial\psi_2}{\partial u}(u,v) & \frac{\partial\psi_2}{\partial v}(u,v)\end{pmatrix} = \det\begin{pmatrix}\frac{1}{1+v} & \frac{-u}{(1+v)^2}\\ \frac{v}{1+v} & \frac{u}{(1+v)^2}\end{pmatrix}\\ &= \frac{u}{(1+v)^3} - \frac{-uv}{(1+v)^3} = \frac{u}{(1+v)^2}.\end{aligned}$$

En posant $f(s,t) = \left(\frac{t}{s}\right)^\alpha e^{-(t+s)}\frac{1}{t}$, on a :

$$\begin{aligned}\int_U \frac{1}{t}\left(\frac{t}{s}\right)^\alpha e^{-(t+s)}\,\mathrm{d}s\mathrm{d}t &= \int_U f(s,t)\,\mathrm{d}s\mathrm{d}t\\ &= \int_U f\big(\psi(u,v)\big)|J_\psi(u,v)|\,\mathrm{d}u\mathrm{d}v\\ &= \int_U v^\alpha e^{-u}\frac{1}{\frac{uv}{1+v}}\frac{u}{(1+v)^2}\,\mathrm{d}u\mathrm{d}v\\ &= \int_U v^\alpha e^{-u}\frac{1}{v(1+v)}\,\mathrm{d}u\mathrm{d}v\\ &= \underbrace{\left(\int_0^\infty e^{-u}\,\mathrm{d}u\right)}_{=1}\underbrace{\left(\int_0^\infty \frac{1}{v^{1-\alpha}(1+v)}\,\mathrm{d}v\right)}_{=I_{1-\alpha}}.\end{aligned}$$

D'où enfin :

$$\Gamma(\alpha)\Gamma(1-\alpha) = I_{1-\alpha} = \frac{\pi}{\sin\big(\pi(1-\alpha)\big)} = \frac{\pi}{\sin(\pi\alpha)}.$$

□

Définition. Pour tout $z \in \mathbb{C}$ tel que $\mathrm{Re}(z) > 0$, on définit la fonction Γ, appelée fonction gamma d'Euler, par :

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \, \mathrm{d}t.$$

Cette intégrale converge absolument pour $\mathrm{Re}(z) > 0$. La fonction Γ peut être prolongée analytiquement en une fonction méromorphe sur $\mathbb{C}\backslash(-\mathbb{N})$, les nombres 0, -1, -2, ... étant des pôles de Γ.

Définition. Soient U et V deux ouverts non vides de $\mathbb{R}^n$ et $\varphi : U \to V$ différentiable sur U. Si on écrit $\varphi = (\varphi_1, \ldots, \varphi_n)$, le jacobien $J_\varphi(x)$ de φ en $x \in U$ est le déterminant de la matrice jacobienne de φ en x, *i.e.*

$$J_\varphi(x) = \det \begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1}(x) & \ldots & \frac{\partial \varphi_1}{\partial x_n}(x) \\ \vdots & \ldots & \vdots \\ \frac{\partial \varphi_n}{\partial x_1}(x) & \ldots & \frac{\partial \varphi_n}{\partial x_n}(x) \end{pmatrix}.$$

Théorème (Théorème de changement de variables)**.** *Soient U et V deux ouverts non vides de $\mathbb{R}^n$ et $\varphi : U \to V$ un $\mathcal{C}^1$-difféomorphisme. On note λ la mesure de Lebesgue sur $\mathbb{R}^n$. Alors pour toute fonction $f : V \to \mathbb{R}_+$ mesurable positive, on a :*

$$\int_V f(t) \, \mathrm{d}\lambda(t) = \int_U f(\varphi(x)) |J_\varphi(x)| \, \mathrm{d}\lambda(x).$$

De plus, si $f : V \to \mathbb{R}$ ou $\mathbb{C}$, alors f est intégrable sur V si et seulement si $(f \circ \varphi)|J_\varphi|$ est intégrable sur U, et dans ce cas l'égalité précédente est encore valable.

Références :

- Amar et Matheron - *Analyse complexe* - Page 249.

3.10 Formule sommatoire de Poisson et application

Proposition (Formule sommatoire de Poisson)**.** *Soit $F \in L^1(\mathbb{R}) \cap \mathcal{C}^0(\mathbb{R})$. On suppose :*

(i) Il existe $M > 0$ et $\alpha > 1$ tels que pour tout $x \in \mathbb{R}$, $|F(x)| \leqslant \frac{M}{(1+|x|)^\alpha}$.

(ii) $\sum_{n=-\infty}^{\infty} |\hat{F}(n)| < \infty$.

Alors pour tout $x \in \mathbb{R}$, les séries (à double sens) suivantes convergent et on a :

$$\sum_{n=-\infty}^{\infty} F(x+n) = \sum_{n=-\infty}^{\infty} \hat{F}(n) e^{2i\pi nx}.$$

Démonstration. Introduisons la fonction $f : x \mapsto \sum_{n=-\infty}^{\infty} F(x+n)$. La série la définissant est normalement convergente sur tout compact de $\mathbb{R}$: en effet,

$$|F(x+n)| \leqslant \frac{M}{(1+|x+n|)^\alpha}$$

et si $A > 0$, $|x| \leqslant A$, et $|n| \geqslant 2A$, alors $|x+n| \geqslant |n| - |x| \geqslant |n| - A \geqslant |n| - \frac{|n|}{2} = \frac{|n|}{2}$. Donc

$$|F(x+n)| \leqslant \frac{M}{\left(1+\frac{|n|}{2}\right)^\alpha}$$

qui est le terme général d'une série (à double sens) convergente. Comme F est par hypothèse continue, f l'est aussi. On effectue le changement d'indice $p = n+1$:

$$f(x+1) = \sum_{n\in\mathbb{Z}} F(x+n+1) = \sum_{p\in\mathbb{Z}} F(x+p) = f(x),$$

qui nous donne que f est 1-périodique. On calcule son coefficient de Fourier d'indice m :

$$c_m(f) = \int_0^1 f(t) e^{-2i\pi mt}\,\mathrm{d}t = \sum_{n\in\mathbb{Z}} \int_0^1 F(t+n) e^{-2i\pi mt}\,\mathrm{d}t,$$

où l'on a pu intervertir somme et intégrale par convergence normale de $\sum F(x+n)$ sur le compact $[0,1]$. Puis :

$$\begin{aligned} c_m(f) &= \sum_{n\in\mathbb{Z}} \int_0^1 F(t+n) e^{-2i\pi m(t+n)}\,\mathrm{d}t \\ &= \sum_{n\in\mathbb{Z}} \int_n^{n+1} F(u) e^{-2i\pi mu}\,\mathrm{d}u \\ &= \int_{-\infty}^{\infty} F(u) e^{-2i\pi mu}\,\mathrm{d}u = \hat{F}(m). \end{aligned}$$

Par hypothèse, $\sum_{m\in\mathbb{Z}} |\hat{F}(m)|$ converge, donc $\sum_{m\in\mathbb{Z}} |c_m(f)|$ converge. On en déduit que la série de Fourier de f converge normalement, donc uniformément, et ainsi f est somme de sa série de Fourier :

$$f(x) = \sum_{n=-\infty}^{\infty} F(x+n) = \sum_{n=-\infty}^{\infty} \hat{F}(n) e^{2i\pi nx}.$$

□

Selon la leçon, on pourra donner une des applications suivantes.

Application (Inversion de Fourier)**.** *Soit $f \in L^1(\mathbb{R}) \cap \mathcal{C}^0(\mathbb{R})$. On suppose que f et $\hat{f}$ vérifient l'hypothèse (i) de la formule sommatoire de Poisson. Alors*

$$f(x) = \int_{\mathbb{R}} \hat{f}(t) e^{2i\pi xt} \, \mathrm{d}t$$

pour tout $x \in \mathbb{R}$.

Démonstration. Soit $g : x \mapsto e^{-2i\pi x\tau}$. On a :

$$\widehat{fg}(x) = \int_{\mathbb{R}} f(t) g(t) e^{-2i\pi xt} \, \mathrm{d}t = \int_{\mathbb{R}} f(t) e^{-2i\pi(x+\tau)t} \, \mathrm{d}t = \hat{f}(x+\tau).$$

On applique la formule sommatoire de Poisson à la fonction fg :

$$\sum_{n\in\mathbb{Z}} (fg)(x+n) = \sum_{n\in\mathbb{Z}} \widehat{fg}(n) e^{2i\pi nx},$$

soit

$$\sum_{n\in\mathbb{Z}} f(x+n) e^{-2i\pi(x+n)\tau} = \sum_{n\in\mathbb{Z}} \hat{f}(n+\tau) e^{2i\pi nx}.$$

En multipliant des deux côtés par $e^{2i\pi x\tau}$, on obtient

$$\sum_{n\in\mathbb{Z}} f(x+n) e^{-2i\pi n\tau} = \sum_{n\in\mathbb{Z}} \hat{f}(n+\tau) e^{2i\pi(n+\tau)x}.$$

On peut intégrer termes à termes par rapport à la variable τ (grâce aux hypothèses de convergence) les deux membres de l'égalité : d'une part

$$\int_0^1 \sum_{n\in\mathbb{Z}} f(x+n) e^{-2i\pi n\tau} \, \mathrm{d}\tau = f(x),$$

et d'autre part

$$\int_0^1 \sum_{n\in\mathbb{Z}} \hat{f}(n+\tau) e^{2i\pi(n+\tau)x} \, \mathrm{d}\tau = \sum_{n\in\mathbb{Z}} \int_n^{n+1} \hat{f}(t) e^{2i\pi tx} \, \mathrm{d}t = \int_{\mathbb{R}} \hat{f}(t) e^{2i\pi tx} \, \mathrm{d}t,$$

d'où le résultat. □

Remarque (Source : Wikipedia). La formule de transformation de Fourier inverse est en fait valable si f et $\hat{f}$ sont intégrables. Pour le prouver, on utilise cette formule établie sous ces hypothèses plus fortes et une méthode de densité. On approche f par une suite de fonctions (f_p) vérifiant les hypothèses de l'application qu'on vient de donner et telle que (f_p) et $(\widehat{f_p})$ convergent respectivement vers f et $\hat{f}$ en norme $L^1(\mathbb{R})$. Pour obtenir une telle suite, on tronque f en l'annulant en dehors de $[-p,p]$ et en la régularisant par convolution. Soit alors φ une fonction de classe $\mathcal{C}^2$, d'intégrale 1 et à support borné. On pose $\varphi_p(x) = p\varphi(px)$ et on convole $f_{|[-p,p]}$ par φ_p. A vérifier si ça marche bien...

Application. *Soit $a > 0$. On a alors*

$$\sum_{n\in\mathbb{Z}} \frac{1}{a^2+n^2} = \frac{\pi}{a}\coth(\pi a).$$

Démonstration. Soit $f : x \mapsto e^{-2\pi a|x|}$. On a :

$$\begin{aligned}\hat{f}(x) &= \int_{\mathbb{R}} e^{-2\pi a|t|}e^{-2i\pi tx}\,\mathrm{d}t \\ &= \int_{-\infty}^{0} e^{2\pi at}e^{-2i\pi tx}\,\mathrm{d}t + \int_0^{\infty} e^{-2\pi at}e^{-2i\pi tx}\,\mathrm{d}t \\ &= \frac{1}{\pi}\frac{a}{a^2+x^2}.\end{aligned}$$

On peut appliquer la formule sommatoire de Poisson qui nous donne en $x = 0$:

$$\sum_{n\in\mathbb{Z}} e^{-2\pi a|n|} = \sum_{n\in\mathbb{Z}} \frac{1}{\pi}\frac{a}{a^2+n^2}.$$

On calcule enfin :

$$\sum_{n\in\mathbb{Z}} e^{-2\pi a|n|} = 2\sum_{n=0}^{\infty} e^{-2\pi an} - 1 = 2\frac{1}{1-e^{-2\pi a}} - 1 = \frac{1+e^{-2\pi a}}{1-e^{-2\pi a}} = \coth(\pi a),$$

d'où le résultat. □

Si $F \in \mathcal{S}(\mathbb{R})$ (l'espace de Schwartz), $\hat{F} \in \mathcal{S}(\mathbb{R})$ (voir [ZQ] page 321). Donc les hypothèses de la formule sommatoire de Poisson sont vérifiées.

Application (Distributions)**.** *Pour $N \geqslant 0$, on pose $T_N = \sum_{n=-N}^{N} \delta_n \in \mathcal{S}'(\mathbb{R})$. Alors la suite $(T_N)_{N\in\mathbb{N}}$ converge dans $\mathcal{S}'(\mathbb{R})$ vers une distribution $\delta_{\mathbb{Z}}$ qui vérifie la relation $\hat{\delta_{\mathbb{Z}}} = \delta_{\mathbb{Z}}$.*

Démonstration. Soit $\varphi \in \mathcal{S}(\mathbb{R})$. D'après la formule sommatoire de Poisson en $x = 0$, $\sum_{n\in\mathbb{Z}} \varphi(n)$ converge, donc

$$(T_N, \varphi) = \sum_{n=-N}^{N} (\delta_n, \varphi) = \sum_{n=-N}^{N} \varphi(n) \underset{N\to\infty}{\longrightarrow} \sum_{n=-\infty}^{\infty} \varphi(n).$$

On définit alors $\delta_\mathbb{Z}$ par $(\delta_\mathbb{Z}, f) = \sum_{n=-\infty}^{\infty} f(n)$ pour $f \in \mathcal{S}(\mathbb{R})$.

Vérifions que $\delta_\mathbb{Z}$ est bien une distribution tempérée. Pour $f \in \mathcal{S}(\mathbb{R})$, on a

$$|(\delta_\mathbb{Z}, f)| \leqslant \sum_{n\in\mathbb{Z}} |f(n)| = |f(0)| + \sum_{n\in\mathbb{Z}^*} \frac{1}{n^2} |n^2 f(n)| \leqslant \|f\|_{0,0} + \left(\sum_{n\in\mathbb{Z}^*} \frac{1}{n^2}\right) \|f\|_{2,0},$$

où $\|f\|_{n,p} = \sup_{x\in\mathbb{R}} |x^n f^{(p)}(x)|$ (ce sont les semi-normes définissant la topologie de $\mathcal{S}(\mathbb{R})$). Donc $\delta_\mathbb{Z}$ est une distribution tempérée.

Pour $f \in \mathcal{S}(\mathbb{R})$, on a par définition de la transformée de Fourier sur $\mathcal{S}'(\mathbb{R})$: $(\hat{\delta_\mathbb{Z}}, f) = (\delta_\mathbb{Z}, \hat{f}) = \sum_{n\in\mathbb{Z}} \hat{f}(n)$. La formule sommatoire de Poisson en $x = 0$ nous donne ensuite $\sum_{n\in\mathbb{Z}} \hat{f}(n) = \sum_{n\in\mathbb{Z}} f(n)$, et on en déduit :

$$(\hat{\delta_\mathbb{Z}}, f) = \sum_{n\in\mathbb{Z}} f(n) = (\delta_\mathbb{Z}, f),$$

soit $\hat{\delta_\mathbb{Z}} = \delta_\mathbb{Z}$. □

Conventions :

- Pour $f \in L^1(\mathbb{R})$, on définit la transformée de Fourier de f sur $\mathbb{R}$ par

$$\hat{f} : x \mapsto \int_\mathbb{R} f(t) e^{-2i\pi xt} \, \mathrm{d}t.$$

  On a rajouté le facteur 2π dans l'exponentielle pour que la formule finale de la formule de Poisson soit plus agréable.
- Pour f continue et 1-périodique sur $\mathbb{R}$, on note $(c_n(f))_{n\in\mathbb{Z}}$ la suite des coefficients de Fourier de f, définis par

$$c_n(f) = \int_0^1 f(t) e^{-2i\pi nt} \, \mathrm{d}t.$$

Références :

- Zuily et Queffélec - *Eléments d'analyse, 2ème édition* - Page 93.

3.11 Lemme de Morse

Lemme. *Soit $S_0 \in \mathcal{S}_n(\mathbb{R})$ une matrice symétrique inversible. Alors il existe un voisinage U de S_0 dans $\mathcal{S}_n(\mathbb{R})$ et une application $f : U \to \mathrm{GL}_n(\mathbb{R})$ de classe $\mathcal{C}^1$ tels que $f(S_0) = I_n$ et pour tout $S \in U$, $S = {}^t(f(S))S_0 f(S)$.*

Démonstration. Soit

$$\begin{array}{rccc} \varphi : & \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathcal{S}_n(\mathbb{R}) \\ & M & \longmapsto & {}^tMS_0M. \end{array}$$

D'abord, φ est de classe $\mathcal{C}^1$ (car polynomiale). En développant $\varphi(I_n + H)$, on trouve $D\varphi(I_n).H = {}^t(S_0H)+(S_0H)$. Donc le noyau $\mathrm{Ker}(D\varphi(I_n))$ est l'ensemble des matrices H telles que S_0H soit antisymétrique.

On note maintenant

$$F = \{M \in \mathcal{M}_n(\mathbb{R}) \,/\, S_0M \in \mathcal{S}_n(\mathbb{R})\},$$

qui est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ de même dimension que $\mathcal{S}_n(\mathbb{R})$ car S_0 est inversible. Soit alors $\psi = \varphi_{|F}$. On a :

$$\mathrm{Ker}(D\psi(I_n)) = \mathrm{Ker}(D\varphi(I_n)) \cap F = \{0\}$$

car $S_0M \in \mathcal{A}_n(\mathbb{R}) \cap \mathcal{S}_n(\mathbb{R})$ implique $S_0M = 0$ et donc $M = 0$ puisque S_0 est inversible. La différentielle $D\psi(I_n)$ est donc inversible (injective, donc surjective par égalité des dimensions). On peut appliquer le théorème d'inversion locale : il existe un voisinage ouvert V de I_n dans F et un voisinage ouvert U de $S_0 = \psi(I_n)$ dans $\mathcal{S}_n(\mathbb{R})$ tels que $\psi_{|V} : V \to U$ est un $\mathcal{C}^1$-difféomorphisme. Comme $\mathrm{GL}_n(\mathbb{R})$ est ouvert, quitte à restreindre, on peut supposer $V \subset \mathrm{GL}_n(\mathbb{R})$.

Ainsi, en posant $f = {\psi_{|V}}^{-1} : U \to V$, on obtient

$$\psi \circ f(S) = S = {}^tf(S)S_0f(S)$$

pour tout $S \in U$, d'où le résultat. □

Théorème (Lemme de Morse)**.** *Soit $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^3$ sur un ouvert U de $\mathbb{R}^n$ contenant l'origine. On suppose que 0 est un point critique non dégénéré de f, i.e. $Df(0) = 0$ et la forme quadratique hessienne $D^2f(0)$ est non dégénérée, de signature $(p, n-p)$. Alors il existe un changement de coordonnées local $\varphi : x \mapsto u$ de classe $\mathcal{C}^1$ entre deux voisinages de l'origine dans $\mathbb{R}^n$ tel que $\varphi(0) = 0$ et*

$$f(x) - f(0) = u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \cdots - u_n^2.$$

Démonstration. On écrit la formule de Taylor avec reste intégral à l'ordre 1 en 0 :

$$f(x) - f(0) = \int_0^1 (1-t)D^2f(tx).(x,x)\,\mathrm{d}t = {}^t xQ(x)x,$$

avec $Q(x) = \int_0^1 (1-t)D^2f(tx)\,\mathrm{d}t$ matrice symétrique, Q est une fonction $\mathcal{C}^1$ de x car f est $\mathcal{C}^3$. On a $Q(0) = \frac{1}{2}D^2f(0)$ matrice symétrique inversible car $D^2f(0)$ non dégénérée.

On applique le lemme : il existe U voisinage ouvert de $Q(0)$ dans $\mathcal{S}_n(\mathbb{R})$ et $g : U \to \mathrm{GL}_n(\mathbb{R})$ de classe $\mathcal{C}^1$ tels que $g(Q(0)) = I_n$ et pour tout $S \in U$, $S = {}^t(g(S))Q(0)g(S)$. Comme Q est $\mathcal{C}^1$, en particulier continue, il existe un voisinage W de 0 dans $\mathbb{R}^n$ tel que $Q(x) \in U$ pour tout $x \in W$. Donc pour tout $x \in W$, $Q(x) = {}^t(M(x))Q(0)M(x)$ avec $M(x) = g(Q(x))$ fonction $\mathcal{C}^1$ de x comme composée de fonctions $\mathcal{C}^1$. D'où

$$f(x) - f(0) = {}^t(y(x))Q(0)y(x)$$

avec $y(x) = M(x)x$. Or $Q(0) = \frac{1}{2}D^2f(0)$ est de signature $(p, n-p)$: il existe donc un changement linéaire de coordonnées $y(x) = Au(x)$ avec A inversible tel que

$${}^t yQ(0)y = {}^t u\,{}^t AQ(0)A\,u = u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \cdots - u_n^2.$$

On a bien la forme demandée pour f. Enfin, l'application $\varphi : x \mapsto u = A^{-1}M(x)x$ est $\mathcal{C}^1$. On calcule sa différentielle à l'origine :

$$\begin{aligned}
\varphi(x) - \varphi(0) &= A^{-1}M(x)x - 0 \\
&= A^{-1}\big(M(0) + DM(0).x + o\left(\|x\|\right)\big)x \\
&= A^{-1}M(0)x + (A^{-1}DM(0).x)x + o\left(\|x\|^2\right) \\
&= A^{-1}M(0)x + o\left(\|x\|\right).
\end{aligned}$$

On en déduit que $D\varphi(0) = A^{-1}M(0) = A^{-1}$ qui est inversible. D'après le théorème d'inversion locale, φ est un $\mathcal{C}^1$-difféomorphisme entre deux voisinages de l'origine dans $\mathbb{R}^n$. □

Remarque. La fonction f devient par changement de coordonnées une simple forme quadratique (formule de Taylor sans reste). Le résultat s'étend immédiatement au cas où $Df(0) \neq 0$ en remplaçant $f(x)$ par $g(x) = f(x) - f(0) - Df(0)x$.

Remarque. Dans la démonstration du lemme, la différentielle $D\varphi(I_n)$ est surjective : pour un S donné dans $\mathcal{S}_n(\mathbb{R})$, on a $D\varphi(I_n).H = S$ pour la matrice $H = \frac{1}{2}S_0^{-1}S \in \mathcal{M}_n(\mathbb{R})$. Mais cela ne sert pas.

On peut proposer l'application suivante qui consiste à étudier la position d'une surface par rapport à son plan tangent :

Application. *Soit S la surface d'équation $z = f(x, y)$ où f est une fonction de classe $\mathcal{C}^3$ au voisinage de $a \in \mathbb{R}^2$. On suppose la forme quadratique $D^2f(a)$ non dégénérée et on note Π le plan tangent à S au point $(a, f(a))$. Alors :*

(i) Si $D^2f(a)$ est de signature $(2, 0)$, alors S est au-dessus de Π au voisinage de a.

(ii) Si $D^2f(a)$ est de signature $(0, 2)$, alors S est en-dessous de Π au voisinage de a.

(iii) Si $D^2f(a)$ est de signature $(1, 1)$, alors S traverse Π selon deux courbes qui se coupent en a.

Démonstration. Comme le plan Π a pour équation $z = f(a) + Df(a).h$, le problème revient à étudier la différence d'altitude

$$\delta(h) = f(a + h) - \big(f(a) + Df(a).h\big).$$

Quitte à translater, on peut supposer $a = 0$. Le lemme de Morse appliqué à δ nous donne alors l'existence d'un $\mathcal{C}^1$-difféomorphisme $h \mapsto (u(h), v(h))$ entre deux voisinages de 0 dans $\mathbb{R}^2$ tel que

$$\delta(h) = \varepsilon_1 u(h)^2 + \varepsilon_2 v(h)^2$$

où $\varepsilon_1, \varepsilon_2 \in \{-1, +1\}$. En particulier, u et v ne s'annulent simultanément que pour $h = 0$.

(i) On obtient immédiatement $\delta(h) > 0$, d'où le résultat.

(ii) De même, $\delta(h) < 0$ donne le résultat.

(iii) On a $\delta(h) = u^2 - v^2$ et la surface S traverse son plan tangent Π selon une courbe admettant un point double en $(a, f(a))$, dont la projection sur le plan xOy se décompose en les deux courbes $u = v$ et $u = -v$ (dans le plan xOy, $\delta(h)$ qui représente l'altitude est nul, d'où $u^2 = v^2$, ce qui donne l'équation des deux courbes).

□

Notation. Soit $f : U \to \mathbb{R}^p$ avec U ouvert de $\mathbb{R}^n$, k fois différentiable en $a \in U$. Pour simplifier les formules de Taylor, on introduit la notation abrégée suivante : pour $h = (h_1, \dots, h_n) \in \mathbb{R}^n$, on note

$$D^k f(a)(h)^k = D^k f(a)(h, h, \dots, h) = \sum_{1 \leqslant i_1, \dots, i_k \leqslant n} \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(a) h_{i_1} \dots h_{i_k}.$$

Théorème (Formule de Taylor-Young)**.** *Si $f : U \to \mathbb{R}^p$ est k fois différentiable en $a \in U$, U ouvert de $\mathbb{R}^n$, on a*

$$f(a+h) = f(a) + Df(a).h + \dots + \frac{1}{k!} D^k f(a)(h)^k + o\left(\|h\|^k\right)$$

lorsque h tend vers 0 dans $\mathbb{R}^n$.

Théorème (Formule de Taylor avec reste intégral)**.** *Si $f : U \to \mathbb{R}^p$, U ouvert de $\mathbb{R}^n$, est de classe $\mathcal{C}^{k+1}$ sur U, et si le segment $[a, a+h]$ est entièrement contenu dans U, on a*

$$f(a+h) = f(a) + Df(a).h + \dots + \frac{1}{k!} D^k f(a)(h)^k + \int_0^1 \frac{(1-t)^k}{k!} D^{k+1} f(a+th)(h)^{k+1} \, \mathrm{d}t.$$

Références :

- Rouvière - *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation* - Page 209 (le lemme), page 354 (le théorème) et page 341 (l'application).

3.12 Maximalité et globalité des solutions de $y' = f(t, y)$

Théorème. *Soient U un ouvert de $\mathbb{R} \times \mathbb{R}^n$ et $y : [t_0, b[\to \mathbb{R}^n$ une solution de l'équation $y' = f(t, y)$ où f est une fonction continue sur U. Alors y peut se prolonger au-delà de b si et seulement s'il existe un compact $K \subset U$ tel que la courbe $t \mapsto (t, y(t))$, $t \in [t_0, b[$, reste contenue dans K.*

Démonstration. Si y se prolonge en b, alors l'image du compact $[t_0, b]$ par l'application continue $t \mapsto (t, y(t))$ est un compact $K \subset U$.

Réciproquement, supposons qu'il existe un compact K de U tel que $(t, y(t)) \in K$ pour tout $t \in [t_0, b[$. Comme f est continue et K compact, on a

$$M = \sup_{(t,z)\in K} \|f(t, z)\| < \infty.$$

Par l'inégalité de la moyenne, on a

$$\|y(t_1) - y(t_2)\| \leqslant M|t_1 - t_2|$$

pour tout $t_1, t_2 \in [t_0, b[$. Par le critère de Cauchy, on en déduit que $y(t)$ admet une limite finie l lorsque t tend vers b à gauche. On peut prolonger y par continuité en b en posant $y(b) = l$ et on a $(b, y(b)) \in K \subset U$ car K est fermé. De plus, $y'(t) = f(t, y(t)) \to f(b, y(b))$ quand $t \to b$. Par le théorème de prolongement de la dérivée, cela prouve que $y'(b)$ existe et vaut $f(b, y(b))$.

Maintenant, le théorème d'existence locale des solutions implique qu'il existe une solution locale z au problème de Cauchy de donnée initiale $z(b) = l = y(b)$ sur un intervalle $[b - \varepsilon, b + \varepsilon]$. On obtient alors un prolongement $\tilde{y}$ de y sur $[t_0, b + \varepsilon]$ en posant $\tilde{y}(t) = z(t)$ pour $t \in [b, b + \varepsilon]$. □

Théorème. *Soit $f : U \to \mathbb{R}^n$ une application continue sur un ouvert $U = J \times \mathbb{R}^n$ où $J \subset \mathbb{R}$ est un intervalle ouvert. On fait l'une ou l'autre des deux hypothèses suivantes :*

(i) Il existe une fonction continue $k : J \to \mathbb{R}_+$ telle que pour tout $t \in J$ fixé, l'application $y \mapsto f(t, y)$ soit lipschitzienne de rapport $k(t)$ sur $\mathbb{R}^n$.

(ii) Il existe des fonctions $c, k : J \to \mathbb{R}_+$ continues telles que l'application $y \mapsto f(t, y)$ satisfasse une croissance linéaire à l'infini du type

$$\|f(t, y)\| \leqslant c(t) + k(t)\|y\|.$$

Alors toute solution maximale de l'équation différentielle $y' = f(t, y)$ est globale, i.e. définie sur J tout entier.

Démonstration. On remarque que l'hypothèse (i) entraîne l'hypothèse (ii) : en effet, pour $t \in J$, l'hypothèse $y \mapsto f(t, y)$ est lipschitzienne implique

$$\|f(t, y) - f(t, 0)\| \leqslant k(t)\|y - 0\|,$$

pour tout $y \in \mathbb{R}^n$. On obtient alors (ii) avec $c(t) = \|f(t, 0)\|$.

Nous allons donc pouvoir nous contenter de démontrer le résultat sous l'hypothèse (ii). Supposons qu'on ait une solution $y : [t_0, b[\to \mathbb{R}^n$ avec $t_0, b \in J$ (ainsi b n'est pas la borne supérieure de J). On pose $C = \sup_{t\in[t_0,b]} c(t)$ et $K = \sup_{t\in[t_0,b]} k(t)$. Alors

$$\|y'(t)\| = \|f(t, y(t))\| \leqslant C + K\|y(t)\|$$

pour tout $t \in [t_0, b[$. Majorons maintenant $\|y(t)\|$. On a $y(t) = y(t_0) + \int_{t_0}^t y'(u)\,\mathrm{d}u$, donc

$$\|y(t)\| \leqslant v(t) = \|y(t_0)\| + \int_{t_0}^t \|y'(u)\|\,\mathrm{d}u.$$

De plus, $v'(t) = \|y'(t)\| \leqslant C + K\|y(t)\| \leqslant C + Kv(t)$. On intègre cette inéquation différentielle :

$$\left(v(t)e^{-K(t-t_0)}\right)' = (v'(t) - Kv(t))e^{-K(t-t_0)} \leqslant Ce^{-K(t-t_0)},$$

d'où

$$v(t)e^{-K(t-t_0)} - v(t_0) \leqslant \frac{C}{K}\left(1 - e^{-K(t-t_0)}\right).$$

Comme $v(t_0) = \|y(t_0)\|$, il vient

$$\sup_{t\in[t_0,b[} \|y(t)\| \leqslant \sup_{t\in[t_0,b[} v(t) \leqslant R = \frac{C}{K}\left(e^{K(b-t_0)} - 1\right) + \|y(t_0)\|e^{K(b-t_0)}.$$

Par conséquent, $(t, y(t))$ reste dans le compact $K = [t_0, b] \times \overline{B}(0, R)$ de $U = J \times \mathbb{R}^n$. D'après le théorème précédent, y peut se prolonger au-delà de b.

D'autre part, si $y :]a, t_0] \to \mathbb{R}^n$ est solution, avec $a, t_0 \in J$, on pose $z(t) = y(t_0 + a - t)$. Alors $z : [a, t_0[\to \mathbb{R}^n$ et vérifie l'équation différentielle $z' = g(t, z)$ avec $g(t, z) = -f(t_0 + a - t, z)$. Cette application g vérifie l'hypothèse (ii), donc d'après le premier cas effectué, z peut se prolonger au-delà de t_0, donc y peut se prolonger au-delà de a. Finalement, toute solution maximale est nécessairement globale, sinon elle pourrait se prolonger. □

Application. *Toute solution maximale de l'équation différentielle* $y' = t\sqrt{t^2 + y^2}$, $(t, y) \in \mathbb{R} \times \mathbb{R}$, *est globale.*

Démonstration. On note $f(t, y) = t\sqrt{t^2 + y^2}$. Alors

$$\left|\frac{\partial f}{\partial y}(t, y)\right| = \left|\frac{2ty}{2\sqrt{t^2 + y^2}}\right| \leqslant |t|$$

car $\frac{|y|}{\sqrt{t^2+y^2}} \leqslant 1$. Par l'inégalité des accroissements finis, $y \mapsto f(t, y)$ est lipschitzienne de rapport $k(t) = |t|$. La fonction k étant continue, le théorème s'applique et les solutions maximales de l'équation différentielle sont globales. □

Remarque. La fonction $f : y \mapsto y\sin(y)$ vérifie l'hypothèse (ii) du théorème mais pas l'hypothèse (i). En effet, on a $f'(y) = y\cos(y) + \sin(y)$, donc par l'égalité des accroissements finis, pour tout $y_1, y_2 \in \mathbb{R}$, il existe $y_3 \in]y_1, y_2[$ tel que

$$|f(y_1) - f(y_2)| = |y_3\cos(y_3) + \sin(y_3)||y_1 - y_2|.$$

Comme $|y_3\cos(y_3) + \sin(y_3)|$ n'est pas borné quand y_1 et y_2 varient dans $\mathbb{R}$, f n'est pas lipschitzienne sur $\mathbb{R}$.

Remarque. Les conditions du théorème ne sont pas nécessaires. En effet, considérons

$$\begin{array}{rccl} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & y & \longmapsto & \begin{cases} 0 \text{ si } y \leqslant 0 \\ y\ln(y) \text{ si } y > 0. \end{cases} \end{array}$$

Alors f n'est pas lipschitzienne : si elle l'était, on aurait $|f(y) - f(0)| \leqslant k|y - 0|$ pour tout $y > 0$, soit $|y\ln(y)| \leqslant k|y|$ pour tout $y > 0$, soit encore $|\ln(y)| \leqslant k$ pour tout $y > 0$ (faux au voisinage de 0 et au voisinage de l'infini). Donc f ne vérifie pas l'hypothèse (i) du théorème.

D'autre part, f ne vérifie pas non plus l'hypothèse (ii) : sinon il existerait des constantes $c, k \in \mathbb{R}$ telles que $|y\ln(y)| \leqslant c + k|y|$ au voisinage de l'infini, ce qui est absurde. On note au passage que comme (i) implique (ii), si f ne vérifie pas (ii), elle ne peut pas vérifier (i).

Résolvons maintenant l'équation différentielle $y' = f(y)$. Les fonctions constantes égales à $k \in]-\infty, 0] \cup \{1\}$ sont solutions sur $\mathbb{R}$. Soient y une solution maximale et I un intervalle sur lequel elle est strictement positive et différente de 1 (s'il n'en existe pas, alors $y' = 0$ sur tout intervalle où elle est définie, donc y est constante égale à $k \in]-\infty, 0] \cup \{1\}$ sur tout son intervalle de définition grâce à sa continuité, et cet intervalle vaut nécessairement $\mathbb{R}$ puisque y est une solution maximale). Alors sur I, on a

$$\frac{y'}{y\ln(y)} = 1.$$

Or $\ln(\ln(y))' = \frac{y'}{y\ln(y)}$, donc il existe $K \in \mathbb{R}$ tel que

$$y(t) = e^{e^{t+K}},$$

pour tout $t \in I$. Cette fonction est toujours strictement plus grande que 1, elle est donc solution sur $\mathbb{R}$. Toute solution maximale de l'équation différentielle $y' = f(y)$ est donc globale bien qu'aucune des hypothèses du théorème ne soit vérifiée.

Conséquence immédiate du théorème de prolongement :

Corollaire (Critère de maximalité)**.** *Une solution $y :]a, b[\to \mathbb{R}^n$ de $y' = f(t, y)$ est maximale si et seulement si $t \mapsto (t, y(t))$ s'échappe de tout compact K de U quand $t \to a^+$ ou quand $t \to b^-$.*

Puisque les compacts sont les fermés bornés, ceci signifie encore que $(t, y(t))$ s'approche du bord de U ou tend vers l'infini, i.e.

$$|t| + \|y(t)\| + \frac{1}{d\big((t, y(t)), \partial U\big)} \underset{t \to a^+ \text{ou } b^-}{\longrightarrow} \infty.$$

Exemple. *Voici un exemple de non-unicité d'une solution maximale à $y' = f(t, y)$ lorsque f est seulement continue. On considère l'équation $y' = 3|y|^{\frac{2}{3}}$. Le problème de Cauchy de condition initiale $y(0) = 0$ admet au moins deux solutions maximales : $y_1(t) = 0$ et $y_2(t) = t^3$ pour $t \in \mathbb{R}$.*

Références :

– Demailly - *Analyse numérique et équations différentielles* - Pages 138 (pour le critère de maximalité) et 144 (pour la globalité).

3.13 Méthode de Newton

Théorème. *Soit I un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{R}$. On suppose que f s'annule en un point $a \in \mathring{I}$, que f est deux fois dérivable sur un voisinage de a, que $f'(a) \neq 0$ et que f'' est bornée. Alors la suite (x_n) définie par*

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

converge quadratiquement vers a si x_0 est suffisamment proche de a.

Démonstration. Quitte à changer f en $-f$, on peut supposer que $f'(a) > 0$. Comme $a \in \mathring{I}$ et f' est continue (car dérivable) au voisinage de a, il existe $\alpha > 0$ tel que $J = [a - \alpha, a + \alpha] \subset I$ et $f' > 0$ sur J. Pour $x \in J$, on pose :

$$F(x) = x - \frac{f(x)}{f'(x)}.$$

On peut noter qu'ainsi, $F(a) = a$ et $F'(a) = 0$ (calcul immédiat). On a :

$$F(x) - a = x - a - \frac{f(x)}{f'(x)} = \frac{f(a) - f(x) - (a - x)f'(x)}{f'(x)},$$

en ayant artificiellement ajouté $f(a)$ qui est nul par hypothèse. On applique maintenant la formule de Taylor-Lagrange : pour $x \in J$, il existe c entre a et x tel que

$$F(x) - a = \frac{f''(c)(x - a)^2}{2f'(x)}.$$

Comme f'' est bornée sur un voisinage de a qu'on peut supposer contenant J (quitte à réduire J), il existe $M > 0$ tel que pour tout $x \in J$,

$$|F(x) - a| \leqslant \frac{M}{2f'(\gamma)}|x - a|^2 = C|x - a|^2,$$

car f' est continue sur le segment J donc atteint sa borne inférieure en un $\gamma \in J$, et on a ensuite noté $C = \frac{M}{2f'(\gamma)} > 0$.

On a donc pour tout $x \in J$,

$$|F(x) - a| \leqslant C|x - a|^2 \leqslant C\alpha^2 \leqslant \alpha,$$

quitte à réduire J en prenant α plus petit de sorte que $C\alpha < 1$. Ainsi, l'intervalle J est stable par F. Pour $x_0 \in J$, la suite (x_n) est donc bien définie. De plus, pour tout $n \geqslant 0$,

$$|x_{n+1} - a| = |F(x_n) - a| \leqslant C|x_n - a|^2,$$

soit encore, en multipliant par C : $C|x_{n+1} - a| \leqslant \big(C|x_n - a|\big)^2$. Par une récurrence immédiate, on obtient :

$$C|x_n - a| \leqslant \big(C|x_0 - a|\big)^{2^n}$$

avec $C|x_0 - a| \leqslant C\alpha < 1$, d'où la convergence quadratique. □

Proposition. *Sous les hypothèses du théorème précédent, en supposant de plus f deux fois dérivable sur tout l'intervalle I et convexe sur I, toujours avec la supposition $f'(a) > 0$, alors tout $x_0 \in I$ avec $x_0 \geqslant a$ convient.*

Démonstration. La fonction f étant convexe sur I, on a que f' est croissante (et $f'' \geqslant 0$), donc $f'(x) \geqslant f'(a) > 0$ pour tout $x \geqslant a$. La fonction F précédente est donc désormais définie pour tout $x \geqslant a$. De plus, la stricte positivité de $f'(x)$ pour $x > a$ implique que $f(x) > f(a) = 0$, donc pour $x > a$:

$$F(x) = x - \frac{f(x)}{f'(x)} < x,$$

De plus, en reprenant une égalité déjà établie plus haut :

$$F(x) - a = \frac{f''(c)(x-a)^2}{2f'(x)} \geqslant 0,$$

d'où $F(x) \geqslant a$ pour tout $x \geqslant a$ (ceci se voit aussi géométriquement par convexité de f). Ces deux inégalités montrent que l'intervalle $I \cap \{x \geqslant a\}$ est stable par F et donc que la suite (x_n) est bien définie pour tout n lorsque $x_0 \geqslant a$. De plus, $F(x) \leqslant x$ implique la décroissance de (x_n), et même la stricte décroissance lorsque $x_0 \neq a$. La suite (x_n) est décroissante minorée, elle converge donc vers une limite $l \geqslant a$ vérifiant $F(l) = l$, *i.e.* $f(l) = 0$, et le seul l possible est $l = a$ (car $f(x) > f(a) = 0$ pour $x > a$).

Enfin, la décroissance de (x_n) est quadratique comme démontrée dans le théorème précédent. □

Ici, si on suppose f'' continue (*i.e.* f de classe $\mathcal{C}^2$) et strictement positive, l'inégalité de décroissance quadratique est essentiellement optimale car si $x_0 \neq a$, alors $x_n > a$ pour tout n et

$$x_{n+1} - a = \frac{f''(c_n)(x_n - a)^2}{2f'(x_n)}$$

avec $a < c_n < x_n$. Donc par continuité de f'' :

$$\frac{x_{n+1} - a}{(x_n - a)^2} \underset{n\to\infty}{\sim} \frac{f''(a)}{2f'(a)} > 0.$$

On peut éventuellement donner l'exemple suivant :

Exemple. *Soient $y > 0$ et la fonction f définie par $f(x) = x^2 - y$. On veut approcher $a = \sqrt{y}$.*

On choisit comme intervalle d'étude $[c, d]$ avec $0 < c < d$ et $c^2 < y < d^2$ (pour que $a = \sqrt{y}$ soit compris entre c et d). Toutes les hypothèses du théorème et de la

proposition sont alors vérifiées pour f ($f'' = 2 \geqslant 0$, donc f est convexe). On doit alors itérer la fonction

$$F(x) = x - \frac{x^2 - y}{2x} = \frac{1}{2}\left(x + \frac{y}{x}\right),$$

en partant d'un $x_0 \geqslant a$, par exemple $x_0 \geqslant \max(y, 1)$. L'erreur $e_n = |x_n - a|$ commise à la n-ième itération est majorée par l'inégalité déjà établie :

$$C|x_n - a| \leqslant \big(C|x_0 - a|\big)^{2^n}.$$

Ici, $C = \frac{M}{2f'(\gamma)} = \frac{\|f''\|_\infty}{2f'(a)} = \frac{2}{4a} = \frac{1}{2a}$. D'où :

$$e_n \leqslant 2a\left(\frac{x_0 - a}{2a}\right)^{2^n}.$$

Bien sûr, si on a pris x_0 trop grand, cette majoration ne nous donne rien. Mais la convergence quadratique est quand même juste car d'après la démonstration faite, la suite converge nécessairement vers a, donc pour n assez grand, la majoration quadratique de l'erreur sera valable.

Sur l'origine de cette méthode : pour résoudre $f(x) = 0$, on cherche à transformer l'équation en un problème équivalent de point fixe, de la forme $F(x) = x$. Cela peut se faire de plusieurs manières, par exemple en prenant $F(x) = x + \lambda(x)f(x)$ où λ est une fonction ne s'annulant pas. On sait que la convergence de la suite $x_{n+1} = F(x_n)$ vers la solution a recherchée (point fixe de F et zéro de f) sera très rapide si ce point a est superattractif, *i.e.* si $F'(a) = 0$. Or

$$F'(a) = 1 + \lambda'(a)\underbrace{f(a)}_{=0} + \lambda(a)f'(a) = 1 + \lambda(a)f'(a).$$

On est donc incité à choisir $\lambda(x) = -\frac{1}{f'(x)}$ et à considérer la suite récurrente

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Application pour les polynômes :

Soit $P(x) = (x - \alpha_1)^{m_1} \dots (x - \alpha_r)^{m_r}$ où $\alpha_1 < \dots < \alpha_r$ sont des réels et les m_i sont des entiers non nuls. D'après le théorème de Gauss-Lucas, les racines de P' sont dans l'enveloppe convexe (dans $\mathbb{C}$) des racines de P. Donc les racines de P' et celles de P'' sont dans $[\alpha_1, \alpha_r]$, et donc P' et P'' ne s'annulent pas sur $]\alpha_r, \infty[$.

On a les deux résultats suivants :

Proposition. *Si $x_0 > \alpha_r$, alors la suite (x_n) décroît strictement et converge vers α_r.*

Démonstration. On a pour tout x, $\frac{P'(x)}{P(x)} = \sum_{i=1}^{r} \frac{m_i}{x-\alpha_i}$. Donc pour tout $n \geqslant 0$,

$$x_{n+1} = x_n - \left(\sum_{i=1}^{r} \frac{m_i}{x_n - \alpha_i} \right)^{-1}.$$

En particulier si $x_n > \alpha_r$, alors $x_{n+1} < x_n$. En dérivant $F(x) = x - \frac{P(x)}{P'(x)}$, on trouve $F'(x) = \frac{P(x)P''(x)}{P'(x)^2}$. Donc $F' > 0$ sur $]\alpha_r, \infty[$ car P est unitaire et on a dit plus haut que P, P', P'' ne s'annulent pas sur $]\alpha_r, \infty[$. Donc F est strictement croissante sur $]\alpha_r, \infty[$. Si $\alpha_r < x_n$, alors $\alpha_r = F(\alpha_r) < F(x_n) = x_{n+1}$. La condition $x_0 > \alpha_r$ implique donc que $x_n > \alpha_r$ pour tout $n \geqslant 0$.

Finalement, la suite (x_n) est décroissante minorée, elle converge donc vers un élément qui annule $\frac{P}{P'}$, *i.e.* vers α_r. □

Proposition. *En ayant choisit $x_0 > \alpha_r$, si $m_r \geqslant 2$, alors il existe $c > 0$ tel que $|x_n - \alpha_r| \sim c \left(1 - \frac{1}{m_r}\right)^n$.*

Démonstration. Comme $\frac{P'(x)}{P(x)} = \sum_{i=1}^{r} \frac{m_i}{x-\alpha_i}$, on a en dérivant cette égalité :

$$\frac{P''(x)P(x)}{P(x)^2} - \frac{P'(x)^2}{P(x)^2} = -\sum_{i=1}^{r} \frac{m_i}{(x-\alpha_i)^2}.$$

On dérive ensuite $F(x) = x - \frac{P(x)}{P'(x)}$:

$$\begin{aligned} F'(x) &= \frac{P(x)P''(x)}{P'(x)^2} \\ &= \frac{P(x)P''(x)}{P(x)^2} \left(\frac{P(x)}{P'(x)} \right)^2 \\ &= \left(\left(\sum_{i=1}^{r} \frac{m_i}{x-\alpha_i} \right)^2 - \sum_{i=1}^{r} \frac{m_i}{(x-\alpha_i)^2} \right) \left(\sum_{i=1}^{r} \frac{m_i}{x-\alpha_i} \right)^{-2} \\ &= 1 - \left(\sum_{i=1}^{r} \frac{m_i}{(x-\alpha_i)^2} \right) \left(\sum_{i=1}^{r} \frac{m_i}{x-\alpha_i} \right)^{-2}. \end{aligned}$$

En multipliant le numérateur et le dénominateur par $(x-\alpha_r)^2$, on a :

$$\left(\sum_{i=1}^{r} \frac{m_i}{(x-\alpha_i)^2} \right) \left(\sum_{i=1}^{r} \frac{m_i}{x-\alpha_i} \right)^{-2} = \frac{\left(\sum_{i=1}^{r} m_i \frac{(x-\alpha_r)^2}{(x-\alpha_i)^2} \right)}{\left(\sum_{i=1}^{r} m_i \frac{x-\alpha_r}{x-\alpha_i} \right)^2} \underset{x \to \alpha_r}{\longrightarrow} \frac{m_r}{m_r^2} = \frac{1}{m_r}.$$

D'où :

$$F'(\alpha_r) = \lim_{x \to \alpha_r} F'(x) = 1 - \frac{1}{m_r} < 1.$$

D'après la proposition précédente, la suite (x_n) est décroissante et converge vers α_r. Ainsi par l'inégalité des accroissements finis, si on choisit $d \in]F'(\alpha_r), 1[$, alors $|x_n - \alpha_r| = O(d^n)$.

D'après la formule de Taylor-Lagrange, il existe $z_n \in]\alpha_r, x_n[$ tel que

$$x_{n+1} - \alpha_r = F'(\alpha_r)(x_n - \alpha_r) + \frac{F''(z_n)}{2}(x_n - \alpha_r)^2.$$

D'où :

$$\varepsilon_n = \frac{x_{n+1} - \alpha_r}{F'(\alpha_r)(x_n - \alpha_r)} - 1 = O(x_n - \alpha_r) = O(d^n).$$

On passe au logarithme :

$$\underbrace{\ln(1+\varepsilon_n)}_{=O(\varepsilon_n)=O(d^n)} = \ln(x_{n+1} - \alpha_r) - \ln(x_n - \alpha_r) - \ln(F'(\alpha_r)).$$

Par conséquent, la série définie par ce terme général converge, ce qui signifie que $\ln(x_n - \alpha_r) - n\ln(F'(\alpha_r))$ converge vers un réel λ. Finalement :

$$x_n - \alpha_r \sim e^{\lambda} F'(\alpha_r)^n = e^{\lambda}\left(1 - \frac{1}{m_r}\right)^n.$$

□

Donc tout $x_0 > \alpha_r$ convient pour avoir la convergence de (x_n) vers α_r. Pour trouver un tel x_0, on écrit $P = x^n + a_{n-1}x^{n-1} + \cdots + a_0$, et alors si $P(\alpha) = 0$, on a

$$|\alpha|^n = \left|\sum_{i=0}^{n-1} a_i \alpha^i\right| \leqslant \sum_{i=0}^{n-1} |a_i||\alpha|^i.$$

Si $|\alpha| \geqslant 1$, en divisant cette inégalité par $|\alpha|^{n-1}$, on obtient $|\alpha| \leqslant \sum_{i=0}^{n-1} |a_i|$. D'où :

$$|\alpha_r| \leqslant \max\left(1, \sum_{i=0}^{n-1} |a_i|\right).$$

Il nous suffit donc ne prendre n'importe quel x_0 plus grand que $\max\left(1, \sum_{i=0}^{n-1} |a_i|\right)$.

Par ailleurs, il est conseillé de diviser P par le PGCD de P et P' de sorte que α_r soit racine simple, ce qui donne une convergence plus rapide.

Pour trouver les autres racines de P, on peut appliquer la méthode de Newton à $\frac{P(x)}{(x-\alpha_r)^{m_r}}$, mais cela pose des problèmes de stabilité car on n'a qu'une valeur approchée de α_r.

Théorème (Gauss-Lucas)**.** *Les racines de P' sont dans l'enveloppe convexe des racines de P.*

Démonstration. Ecrivons $P = (X-\alpha_1)^{m_1}\dots(X-\alpha_r)^{m_r}$. Alors

$$\frac{P'(x)}{P(x)} = \sum_{i=1}^{r} \frac{m_i}{x-\alpha_i}.$$

Soit β une racine de P'. S'il s'agit aussi d'une racine de P, alors le résultat est clair. Sinon on a :

$$\frac{P'(\beta)}{P(\beta)} = \sum_{i=1}^{r} \frac{m_i}{\beta-\alpha_i} = \sum_{i=1}^{r} m_i \frac{\overline{\beta-\alpha_i}}{|\beta-\alpha_i|^2} = 0.$$

D'où :

$$\beta \sum_{i=1}^{r} \frac{m_i}{|\beta-\alpha_i|^2} = \sum_{i=1}^{r} \frac{m_i}{|\beta-\alpha_i|^2}\alpha_i.$$

En posant $a_i = \frac{m_i}{|\beta-\alpha_i|^2}$, on a donc :

$$\beta = \sum_{i=1}^{r} \frac{a_i}{\sum_{j=1}^{r} a_j}\alpha_i,$$

d'où le résultat.

Remarque : Dans le cas où les racines sont toutes réelles, le résultat s'obtient aussi en remarquant que chaque α_i est une racine de P' de multiplicité $m_i - 1$ et en appliquant le théorème de Rolle sur chaque intervalle $]\alpha_i, \alpha_{i+1}[$, on obtient $r-1$ autres racines. On a alors bien $n-1$ racines, toutes comprises entre les α_i. □

Références :

- Rouvière - *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation* - Page 152.
- Chambert-Loir - *Exercices d'analyse 2* (pour l'application aux polynômes).

3.14 Nombres de Bell

Proposition. *Pour $n \in \mathbb{N}^*$, on note B_n le nombre de partitions de l'ensemble $[\![1, n]\!]$, avec $B_0 = 1$ par convention. Alors :*

(i) La série entière $\sum \frac{B_n}{n!} z^n$ a un rayon de convergence $R > 0$ et en notant f sa somme, on a

$$f(z) = e^{e^z - 1}$$

pour tout $|z| < R$.

(ii) Pour tout $k \in \mathbb{N}$, on a

$$B_k = \frac{1}{e} \sum_{n=0}^{\infty} \frac{n^k}{n!}.$$

Démonstration. Pour $k \in [\![1, n]\!]$, on note E_k l'ensemble des partitions de $[\![1, n+1]\!]$ pour lesquelles la partie de $[\![1, n+1]\!]$ contenant $n+1$ est de cardinal $k+1$. Alors on a :

$$\mathrm{Card}(E_k) = \binom{n}{k} B_{n-k}.$$

En effet, il y a $\binom{n}{k}$ choix possibles pour la partie à $k+1$ éléments contenant l'élément $n+1$ (on a k éléments à choisir dans $[\![1, n]\!]$), puis il faut réaliser une partition des $n-k$ éléments restants.

Les ensembles $E_0, E_1, \ldots, E_n$ forment clairement une partition de l'ensemble des partitions de $[\![1, n+1]\!]$. On a donc

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_{n-k} = \sum_{j=0}^{n} \binom{n}{j} B_j$$

car $\binom{n}{n-k} = \binom{n}{k}$. On montre maintenant par récurrence sur n que $B_n \leqslant n!$. Pour $n = 0$ et $n = 1$, c'est clair : $B_0 = 1$, $B_1 = 1$. Supposons que $B_k \leqslant k!$ pour tout $k \leqslant n$. Alors

$$B_{n+1} \leqslant \sum_{k=0}^{n} \binom{n}{k} k! = n! \sum_{k=0}^{n} \underbrace{\frac{1}{(n-k)!}}_{\leqslant 1} \leqslant (n+1)!.$$

Donc pour tout $n \in \mathbb{N}^*$, $\frac{B_n}{n!} \leqslant 1$, ce qui prouve que le rayon de convergence R est $\geqslant 1$.

Pour $|z| < R$, on a

$$f(z) = 1 + \sum_{n=0}^{\infty} \frac{B_{n+1}}{(n+1)!} z^{n+1}.$$

La fonction f est dérivable sur son disque ouvert de convergence et

$$f'(z) = \sum_{n=0}^{\infty} \frac{B_{n+1}}{n!} z^n.$$

Donc pour tout $|z| < R$,

$$f'(z) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^{n} \binom{n}{k} B_k \right) z^n = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \frac{B_k}{k!} \frac{1}{(n-k)!} \right) z^n.$$

On reconnaît dans le terme de droite le produit de Cauchy des séries $\sum \frac{B_n}{n!} z^n$ et $\sum \frac{z^n}{n!}$ qui ont toutes deux un rayon de convergence $\geqslant R$. Donc pour $|z| < R$, on a

$$f'(z) = f(z)e^z.$$

On en déduit qu'il existe $C \in \mathbb{C}$ tel que $f(z) = Ce^{e^z}$. En évaluant en 0, comme $f(0) = B_0 = 1$, on trouve $C = e^{-1}$. D'où

$$f(z) = e^{e^z - 1}.$$

On a donc démontré le point (i), et on peut remarquer que l'expression de f donne $R = \infty$.

Pour tout $z \in \mathbb{C}$, on a :

$$e^{e^z} = \sum_{n=0}^{\infty} \frac{e^{nz}}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{\infty} \frac{(nz)^k}{k!}.$$

En notant $u_{n,k} = \frac{(nz)^k}{n!k!}$, on a, pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{\infty} |u_{n,k}| = \sum_{k=0}^{\infty} \frac{|nz|^k}{n!k!} = \frac{e^{|nz|}}{n!}.$$

Puis,

$$\sum_{n=0}^{\infty} \frac{e^{|nz|}}{n!} = \sum_{n=0}^{\infty} \frac{\left(e^{|z|}\right)^n}{n!} = e^{e^{|z|}}.$$

Donc la série $\sum_n \sum_{k=0}^{\infty} |u_{n,k}|$ est convergente. Par le théorème de Fubini pour les séries doubles, on peut intervertir les symboles sommes :

$$f(z) = e^{-1} \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} u_{n,k} = e^{-1} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} u_{n,k} = e^{-1} \sum_{k=0}^{\infty} \left(\sum_{n=0}^{\infty} \frac{n^k}{n!} \right) \frac{z^k}{k!}.$$

Par unicité du développement en série entière de f, on obtient

$$B_k = e^{-1} \sum_{n=0}^{\infty} \frac{n^k}{n!}$$

pour tout $k \in \mathbb{N}$. □

Formulation adaptée pour la leçon *"Anneau des séries formelles, applications"* :

Proposition. *Pour $n \in \mathbb{N}^*$, on note B_n le nombre de partitions de l'ensemble $[\![1, n]\!]$, avec $B_0 = 1$ par convention. Alors pour tout $k \in \mathbb{N}$, on a*

$$B_k = \frac{1}{e} \sum_{n=0}^{\infty} \frac{n^k}{n!}.$$

Démonstration. Pour $k \in [\![1, n]\!]$, on note E_k l'ensemble des partitions de $[\![1, n+1]\!]$ pour lesquelles la partie de $[\![1, n+1]\!]$ contenant $n+1$ est de cardinal $k+1$. Alors on a :

$$\mathrm{Card}(E_k) = \binom{n}{k} B_{n-k}.$$

En effet, il y a $\binom{n}{k}$ choix possibles pour la partie à $k+1$ éléments contenant l'élément $n+1$ (on a k éléments à choisir dans $[\![1, n]\!]$), puis il faut réaliser une partition des $n-k$ éléments restants.

Les ensembles $E_0, E_1, \ldots, E_n$ forment clairement une partition de l'ensemble des partitions de $[\![1, n+1]\!]$. On a donc

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_{n-k} = \sum_{j=0}^{n} \binom{n}{j} B_j$$

car $\binom{n}{n-k} = \binom{n}{k}$.

On note $F = \sum_{n=0}^{\infty} \frac{B_n}{n!} X^n$ la série génératrice exponentielle de (B_n). On dérive F :

$$\begin{aligned} F' &= \sum_{n=0}^{\infty} \frac{B_{n+1}}{n!} X^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^{n} \binom{n}{k} B_k \right) X^n \\ &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \frac{B_k}{k!} \frac{1}{(n-k)!} \right) X^n. \end{aligned}$$

On reconnaît dans le terme de droite le produit des séries formelles $\sum \frac{B_n}{n!} X^n$ et $\sum \frac{X^n}{n!}$. On a donc

$$F' = E_1 F,$$

où $E_1 = \sum_{n=0}^{\infty} \frac{X^n}{n!}$.

On sait que l'équation $A' = E_1 A$ admet une unique solution dans $\mathbb{C}[[X]]$ lorsque $A(0)$ est donné. En effet, en écrivant $A = \sum_{n=0}^{\infty} a_n X^n$, on obtient $(n+1)a_{n+1} = \sum_{k=0}^{n} \frac{a_k}{(n-k)!}$, ce qui détermine uniquement la suite (a_n) en fonction de a_0.

Comme on a $F(0) = B_0 = 1$, on va chercher une série entière $f = \sum a_n z^n$ vérifiant $f'(z) = E_1(z) f(z)$ et $f(0) = 1$. La série formelle $\sum_{n=0}^{\infty} a_n X^n$ sera alors égale à F.

L'équation $f'(z) = E_1(z)f(z) = e^z f(z)$ s'intègre en $f(z) = Ke^{e^z}$, et $f(0) = 1$ donne $K = e^{-1}$. Pour tout $z \in \mathbb{C}$, on a :

$$e^{e^z} = \sum_{n=0}^{\infty} \frac{e^{nz}}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{\infty} \frac{(nz)^k}{k!}.$$

En notant $u_{n,k} = \frac{(nz)^k}{n!k!}$, on a, pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{\infty} |u_{n,k}| = \sum_{k=0}^{\infty} \frac{|nz|^k}{n!k!} = \frac{e^{|nz|}}{n!}.$$

Puis,

$$\sum_{n=0}^{\infty} \frac{e^{|nz|}}{n!} = \sum_{n=0}^{\infty} \frac{\left(e^{|z|}\right)^n}{n!} = e^{e^{|z|}}.$$

Donc la série $\sum_n \sum_{k=0}^{\infty} |u_{n,k}|$ est convergente. Par le théorème de Fubini pour les séries doubles, on peut intervertir les symboles sommes :

$$f(z) = e^{-1} \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} u_{n,k} = e^{-1} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} u_{n,k} = e^{-1} \sum_{k=0}^{\infty} \left(\sum_{n=0}^{\infty} \frac{n^k}{n!} \right) \frac{z^k}{k!}.$$

Finalement,

$$F = \sum_{k=0}^{\infty} \frac{B_k}{k!} X^k = e^{-1} \sum_{k=0}^{\infty} \left(\sum_{n=0}^{\infty} \frac{n^k}{n!} \right) \frac{X^k}{k!},$$

d'où l'on déduit

$$B_k = e^{-1} \sum_{n=0}^{\infty} \frac{n^k}{n!}$$

pour tout $k \in \mathbb{N}$. □

Références :

- Francinou, Gianella, Nicolas - *Oraux X-ENS, Algèbre 1* - Page 14.
- Saux-Picart - *Cours de calcul formel, Algorithmes fondamentaux* - Page 132 (exercice 14, pour une idée dans le cadre des séries formelles).

3.15 Nombres normaux

On considère l'espace de probabilité $(\Omega, \mathcal{A}, P)$ constitué des irrationnels de $[0,1]$ avec leur tribu borélienne $\mathcal{A}$ et la probabilité induite par la mesure de Lebesgue : $P(B) = \int_0^1 \mathbb{1}_B(x)\,\mathrm{d}x$ si $B \in \mathcal{A}$.

Théorème. *Presque tout irrationnel est normal.*

Démonstration. Soient $r \geqslant 2$ et $A = \{0, 1, \dots, r-1\}$. Pour $x \in \Omega$, on note $(\varepsilon_n(x))_{n\geqslant 1}$ le développement propre de x en base r. On commence par montrer que les $v.a.$ ε_n sont indépendantes sous P et de même loi uniforme sur A.

Soient $a_1, \dots, a_N \in A$. On montre que :

$$\{\varepsilon_1 = a_1, \dots, \varepsilon_N = a_N\} = \left]\sum_{n=1}^{N} \frac{a_n}{r^n}, \sum_{n=1}^{N} \frac{a_n}{r^n} + \frac{1}{r^N}\right[\cap\, \Omega.$$

En effet, si $x = \sum_{n=1}^{\infty} \frac{\varepsilon_n(x)}{r^n}$ avec $\varepsilon_n(x) = a_n$ pour tout $n \in [\![1, N]\!]$, alors

$$x \leqslant \sum_{n=1}^{N} \frac{a_n}{r^n} + \sum_{n=N+1}^{\infty} \frac{r-1}{r^n} = \sum_{n=1}^{N} \frac{a_n}{r^n} + \frac{1}{r^N},$$

et l'inégalité est stricte puisque x est irrationnel. On a donc que l'ensemble de gauche est inclu dans celui de droite. Pour l'autre inclusion, voir ((i)) des compléments.

Par conséquent, $P(\varepsilon_1 = a_1, \dots, \varepsilon_N = a_N) = \frac{1}{r^N}$ (la mesure de l'ensemble de droite). Par ailleurs, on a

$$P(\varepsilon_1 = a_1) = \sum_{a_2,\dots,a_N \in A} P(\varepsilon_1 = a_1, \dots, \varepsilon_N = a_N) = \sum_{a_2,\dots,a_N \in A} \frac{1}{r^N} = \frac{r^{N-1}}{r^N} = \frac{1}{r}.$$

De même, $P(\varepsilon_i = a_i) = \frac{1}{r}$, et on trouve ainsi :

$$P(\varepsilon_1 = a_1, \dots, \varepsilon_N = a_N) = \frac{1}{r^N} = P(\varepsilon_1 = a_1) \dots P(\varepsilon_N = a_N),$$

ce qui prouve l'indépendance des ε_i (voir ((ii)) des compléments) qui suivent une loi uniforme sur A.

Soit $b \in A$. On note $X_j = \mathbb{1}_{\{\varepsilon_j = b\}}$ et

$$N_{b,n}(x) = \mathrm{Card}(\{i \in [\![1, n]\!] \,/\, \varepsilon_i(x) = b\}).$$

Alors par la loi forte des grands nombres ($X_1 \in L^1$), on a :

$$\frac{N_{b,n}}{n} = \frac{X_1 + \dots + X_n}{n} \xrightarrow[n\to\infty]{p.s.} E[X_1] = P(\varepsilon_1 = b) = \frac{1}{r}.$$

Ceci étant vrai pour tout $b \in A$, on en déduit que presque tout $x \in \Omega$ est simplement normal en base r.

Soit maintenant $b = (u, v) \in A^2$. On note $Y_i = \mathbb{1}_{\{\varepsilon_i = u,\, \varepsilon_{i+1} = v\}}$. Alors en notant $N_{b,n}(x) = \text{Card}(\{i \in [\![1, n-1]\!] \,/\, \varepsilon_i(x) = u, \varepsilon_{i+1}(x) = v\})$, on a

$$\frac{N_{b,n}}{n} = \frac{Y_1 + \cdots + Y_{n-1}}{n}.$$

On ne peut pas conclure directement car les Y_i ne sont pas indépendantes. Mais les Y_i sont identiquement distribuées, et $(Y_1, Y_3, \dots)$ d'une part et $(Y_2, Y_4, \dots)$ d'autre part sont des suites indépendantes. Par la loi forte des grands nombres, on obtient :

$$\begin{cases} \frac{Y_1+Y_3+\cdots+Y_{2n-1}}{n} \xrightarrow[n\to\infty]{p.s.} E[Y_1] = P(\varepsilon_1 = u, \varepsilon_2 = v) = \frac{1}{r^2} \\ \frac{Y_2+Y_4+\cdots+Y_{2n}}{n} \xrightarrow[n\to\infty]{p.s.} E[Y_2] = P(\varepsilon_2 = u, \varepsilon_3 = v) = \frac{1}{r^2}. \end{cases}$$

Donc

$$\begin{cases} \frac{Y_1+Y_2+\cdots+Y_{2n}}{2n} = \frac{1}{2}\left(\frac{Y_1+Y_3+\cdots+Y_{2n-1}}{n} + \frac{Y_2+Y_4+\cdots+Y_{2n}}{n}\right) \xrightarrow[n\to\infty]{p.s.} \frac{1}{2}\left(\frac{1}{r^2} + \frac{1}{r^2}\right) = \frac{1}{r^2} \\ \frac{Y_1+Y_2+\cdots+Y_{2n-1}}{2n-1} = \frac{n}{2n-1}\left(\frac{Y_1+Y_3+\cdots+Y_{2n-1}}{n} + \frac{Y_2+Y_4+\cdots+Y_{2n-2}}{n}\right) \xrightarrow[n\to\infty]{p.s.} \frac{1}{r^2}. \end{cases}$$

On en déduit que

$$\frac{N_{b,n}}{n} \xrightarrow[n\to\infty]{p.s.} \frac{1}{r^2}.$$

On peut faire de même avec les mots $b \in A^k$ de longueur $k \geqslant 1$.

On note à présent $E_r = \{x \in \Omega \,/\, x \text{ est normal en base } r\}$. Alors

$$E_r = \bigcap_{k=1}^{\infty} \bigcap_{b \in A^k} \left\{ x \in \Omega \,/\, \frac{N_{b,n}(x)}{n} \xrightarrow[n\to\infty]{} \frac{1}{r^k} \right\}$$

est de probabilité 1 comme intersection dénombrable d'événements de probabilité 1. Enfin, on note $E = \{x \in \Omega \,/\, x \text{ est normal}\}$. Alors $E = \bigcap_{r \geqslant 2} E_r$ et on a de même $P(E) = 1$ (voir ((iii)) des compléments), ce qui termine la démonstration. □

Compléments de la démonstration :

(i) Montrons que

$$\{\varepsilon_1 = a_1, \dots, \varepsilon_N = a_N\} \supset \left]\sum_{n=1}^{N} \frac{a_n}{r^n}, \sum_{n=1}^{N} \frac{a_n}{r^n} + \frac{1}{r^N}\right[\cap\, \Omega.$$

Fixons $x = \sum_{n=1}^{\infty} \frac{\varepsilon_n(x)}{r^n}$ dans l'ensemble de droite. Supposons qu'il existe $i \leqslant N$ tel que $\varepsilon_i(x) \neq a_i$ et considérons le plus petit i vérifiant cette propriété. On note $S_N = \sum_{n=1}^{N} \frac{a_n}{r^n}$. Alors

$$x - S_N = \frac{\varepsilon_i(x) - a_i}{r^i} + \underbrace{\sum_{n=i+1}^{N} \frac{\varepsilon_n(x) - a_n}{r^n} + \sum_{n=N+1}^{\infty} \frac{\varepsilon_n(x)}{r^n}}_{=R_i}.$$

On a $|R_i| \leqslant \sum_{n=i+1}^{\infty} \frac{r-1}{r^n} \leqslant \frac{1}{r^i}$, et l'inégalité est stricte puisque x est irrationnel. D'autre part, $R_i \geqslant \sum_{n=i+1}^{N} \frac{\varepsilon_n(x)-a_n}{r^n}$ et l'inégalité est aussi stricte. Il vient alors en calculant :

$$R_i > -\sum_{n=i+1}^{N} \frac{r-1}{r^n} = -(r-1)\frac{\frac{1}{r^{i+1}} - \frac{1}{r^{N+1}}}{1-\frac{1}{r}} = -\frac{1}{r^i} + \frac{1}{r^N}.$$

Finalement, si $\varepsilon_i(x) - a_i \leqslant -1$, on obtient $x - S_N \leqslant -\frac{1}{r^i} + R_i < -\frac{1}{r^i} + \frac{1}{r^i} = 0$. Et si $\varepsilon_i(x) - a_i \geqslant 1$, on obtient $x - S_N \geqslant \frac{1}{r^i} + R_i > \frac{1}{r^i} - \frac{1}{r^i} + \frac{1}{r^N} = \frac{1}{r^N}$. Dans tous les cas, c'est absurde. Donc pour tout $i \leqslant N$, $\varepsilon_i(x) = a_i$, et on obtient l'inclusion voulue.

(ii) Pour l'indépendance, on a en effet que si $I \subset \mathbb{N}^*$ est une partie finie de $\mathbb{N}^*$, alors il existe N tel que $I \subset [\![1, N]\!]$ et ainsi

$$P\left(\bigcap_{i\in I}\{\varepsilon_i = a_i\}\right) = \sum_{\substack{a_{i_1},\dots,a_{i_k}\in A \\ i_j\notin I}} P\left(\bigcap_{i=1}^{N}\{\varepsilon_i = a_i\}\right) = \sum_{\substack{a_{i_1},\dots,a_{i_k}\in A \\ i_j\notin I}} \frac{1}{r^N},$$

où $k = N - \mathrm{Card}(I)$. Enfin, on obtient :

$$P\left(\bigcap_{i\in I}\{\varepsilon_i = a_i\}\right) = \frac{r^{N-\mathrm{Card}(I)}}{r^N} = \frac{1}{r^{\mathrm{Card}(I)}} = \prod_{i\in I} P(\varepsilon_i = a_i).$$

(iii) Justifions que $P(E) = 1$. On a :

$$P(\Omega\backslash E) = P\left(\Omega\backslash\bigcap_{r\geqslant 2} E_r\right) = P\left(\bigcup_{r\geqslant 2}(\Omega\backslash E_r)\right) \leqslant \sum_{r=2}^{\infty} \underbrace{P(\Omega\backslash E_r)}_{=0} = 0.$$

Notation. Si $x \in [0,1]$ et $r \in \mathbb{N}$, $r \geqslant 2$. On sait que x possède un unique développement propre en base r, *i.e.* qu'il existe des $\varepsilon_n(x) \in A = \{0, 1, \dots, r-1\}$ tels que

$$x = \sum_{n=1}^{\infty} \frac{\varepsilon_n(x)}{r^n},$$

et tels qu'il n'existe pas de rang n_0 à partir duquel $\varepsilon_n(x) = r-1$ pour tout n (le développement est propre).

Soient $k \geqslant 1$ et $b \in A^k$. On pose

$$N_{b,n}(x) = \text{Card}(\{i \in [\![1, n-k+1]\!] \,/\, \varepsilon_i(x) = b_1, \ldots, \varepsilon_{i+k-1}(x) = b_k\}),$$

qui correspond au nombre d'occurrences du mot b dans les n premières positions de l'écriture de x en base r.

Définition. On dit que x est simplement normal en base r si, pour tout $b \in A$,

$$\frac{N_{b,n}(x)}{n} \underset{n\to\infty}{\longrightarrow} \frac{1}{r}.$$

Définition. On dit que x est normal en base r si, pour tout $k \geqslant 1$ et tout $b \in A^k$,

$$\frac{N_{b,n}(x)}{n} \underset{n\to\infty}{\longrightarrow} \frac{1}{r^k}.$$

Définition. On dit que x est normal s'il est normal en toute base $r \geqslant 2$.

Il est difficile de répondre à la question "existe-t-il des nombres normaux ?" car on ne connaît la loi de développement en base r d'aucun irrationnel de $[0,1]$ donné à l'avance. Les probabilités nous aident donc beaucoup.

Références :

– Zuily et Queffélec - *Eléments d'analyse (2ème édition)* - Page 539.

3.16 Polynômes de Bernstein

Proposition. *Soit $f : [0,1] \to \mathbb{C}$ continue. Soit ω son module de continuité uniforme : $\omega(h) = \sup_{|u-v|\leqslant h} |f(u)-f(v)|$. Pour $n \geqslant 1$, on note B_n le n-ième polynôme de Bernstein de f défini sur $[0,1]$ par $B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right)$.*

Alors il existe $C > 0$ tel que $\|f - B_n\|_\infty \leqslant C\omega\left(\frac{1}{\sqrt{n}}\right)$, en particulier (B_n) converge vers f uniformément. De plus, cette estimation est optimale.

Démonstration. Soit $x \in [0,1]$. Soit $(X_n)_{n\geqslant 1}$ une suite de variables de Bernoulli de paramètre x indépendantes et identiquement distribuées. On note $S_n = \sum_{k=1}^{n} X_k$. Alors S_n suit une loi binomiale de paramètres (n,x) et

$$E\left[f\left(\frac{S_n}{n}\right)\right] = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k} = B_n(x).$$

Le module de continuité ω est sous-additif et vérifie donc $\omega(\lambda h) \leqslant (\lambda+1)\omega(h)$ pour tout $\lambda, h \geqslant 0$. D'où :

$$\omega\left(\frac{1}{\sqrt{n}}\left|x - \frac{S_n}{n}\right|\sqrt{n}\right) \leqslant \left(\left|x-\frac{S_n}{n}\right|\sqrt{n}+1\right)\omega\left(\frac{1}{\sqrt{n}}\right).$$

Donc

$$\begin{aligned}
|f(x) - B_n(x)| &= \left|E[f(x)] - E\left[f\left(\frac{S_n}{n}\right)\right]\right| \\
&\leqslant E\left[\left|f(x) - f\left(\frac{S_n}{n}\right)\right|\right] \\
&\leqslant E\left[\omega\left(\left|x-\frac{S_n}{n}\right|\right)\right] \\
&\leqslant \omega\left(\frac{1}{\sqrt{n}}\right)\left(E\left[\left|x-\frac{S_n}{n}\right|\right]\sqrt{n}+1\right) \\
&\leqslant \omega\left(\frac{1}{\sqrt{n}}\right)\left(\sqrt{n}\left\|x-\frac{S_n}{n}\right\|_2+1\right),
\end{aligned}$$

par l'inégalité de Hölder, où $\|X\|_2 = E[X^2]^{\frac{1}{2}}$. Or $E\left[x - \frac{S_n}{n}\right] = 0$, donc $\left\|x-\frac{S_n}{n}\right\|_2 = \mathrm{Var}\left(x-\frac{S_n}{n}\right)^{\frac{1}{2}} = \left(\frac{1}{n^2}\mathrm{Var}(S_n)\right)^{\frac{1}{2}}$. Et $\mathrm{Var}(S_n) = \sum_{i=1}^{n}\mathrm{Var}(X_i) = nx(1-x)$ car les X_i sont indépendantes. D'où :

$$|f(x)-B_n(x)| \leqslant \omega\left(\frac{1}{\sqrt{n}}\right)\left(\sqrt{x(1-x)}+1\right) \leqslant \omega\left(\frac{1}{\sqrt{n}}\right)\left(\sqrt{\frac{1}{4}}+1\right) = \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right),$$

en notant que $x(1-x) \leqslant \frac{1}{4}$ (par une étude variationnelle par exemple). En fin de compte, $\|f - B_n\|_\infty \leqslant \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right)$.

Pour montrer que l'estimation est optimale, on considère la fonction $f : x \mapsto \left|x - \frac{1}{2}\right|$. Alors pour tout $u, v \in [0,1]$, on a $|f(u) - f(v)| = \left|\left|u - \frac{1}{2}\right| - \left|v - \frac{1}{2}\right|\right| \leqslant |u - v|$. Donc $\omega(h) \leqslant h$. D'autre part,

$$\|f - B_n\|_\infty \geqslant \left|f\left(\frac{1}{2}\right) - B_n\left(\frac{1}{2}\right)\right| = \left|B_n\left(\frac{1}{2}\right)\right| = E\left[\left|f\left(\frac{S_n}{n}\right)\right|\right] = E\left[\left|\frac{S_n}{n} - \frac{1}{2}\right|\right]$$

en choisissant les X_k *i.i.d.* suivant une loi de Bernoulli de paramètre $\frac{1}{2}$, et $S_n = X_1 + \cdots + X_n$. On pose ensuite $Y_k = 2X_k - 1$: les Y_k sont *i.i.d.* de loi $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. On obtient :

$$E\left[\left|\frac{S_n}{n} - \frac{1}{2}\right|\right] = \frac{1}{2n}E[|2S_n - n|] = \frac{1}{2n}E\left[\left|\sum_{k=1}^n Y_k\right|\right] \geqslant \frac{1}{2n\sqrt{e}}\left\|\sum_{k=1}^n Y_k\right\|_2$$

par l'inégalité de Khintchine. Or

$$\left\|\sum_{k=1}^n Y_k\right\|_2^2 = E\left[\left(\sum_{k=1}^n Y_k\right)^2\right] = \sum_{i\neq j} E[Y_i]\,E[Y_j] + \sum_{k=1}^n E\big[Y_k^2\big] = n$$

car $E[Y_k] = 0$ et $E[Y_k^2] = 1$. Finalement :

$$\|f - B_n\|_\infty \geqslant \frac{1}{2n\sqrt{e}}\sqrt{n} = \frac{1}{2\sqrt{n}\sqrt{e}} \geqslant \frac{1}{2\sqrt{e}}\omega\left(\frac{1}{\sqrt{n}}\right).$$

□

On peut aussi démontrer que (B_n) converge vers f uniformément ainsi :

$$\begin{aligned}
|f(x) - B_n(x)| &= \left|E\left[f(x) - f\left(\frac{S_n}{n}\right)\right]\right| \\
&\leqslant E\left[\left|f(x) - f\left(\frac{S_n}{n}\right)\right|\right] \\
&\leqslant \omega(\delta)\underbrace{E\left[\mathbb{1}_{\left|x-\frac{S_n}{n}\right|\leqslant\delta}\right]}_{\leqslant 1} + 2\|f\|_\infty \underbrace{E\left[\mathbb{1}_{\left|x-\frac{S_n}{n}\right|>\delta}\right]}_{=P\left(\left|x-\frac{S_n}{n}\right|>\delta\right)} \\
&\leqslant \omega(\delta) + 2\|f\|_\infty \frac{\mathrm{Var}(X_1)}{n\delta^2},
\end{aligned}$$

par l'inégalité de Bienaymé-Tchebychev, en notant que $E\left[\frac{S_n}{n}\right] = E[X_1] = x$. D'autre part, $\mathrm{Var}(X_1) = x(1-x) \leqslant \frac{1}{4}$ (par une étude variationnelle par exemple). Finalement,

$$\|f - B_n\|_\infty \leqslant \omega(\delta) + \frac{\|f\|_\infty}{2n\delta^2},$$

ce qui permet de conclure facilement quant à la convergence uniforme de (B_n) vers f.

Définition. On définit L^p comme étant l'espace vectoriel des variables aléatoires X telles que

$$\|X\|_p = E[|X|^p]^{\frac{1}{p}} < \infty.$$

On dit alors que X a un moment d'ordre p. On remarque que $\|\cdot\|_p$ est une semi-norme sur L^p.

Théorème (Inégalité de Hölder)**.** *Soit $1 < p, q < \infty$ avec $\frac{1}{p} + \frac{1}{q} = 1$. Soient $X \in L^p$ et $Y \in L^q$. On a alors :*

$$E[|XY|] \leqslant \|X\|_p \|Y\|_q.$$

Démonstration. Pour éviter de noter les valeurs absolues, on peut supposer $X, Y \geqslant 0$. Si $E[X^p] = 0$, alors $E[XY] = 0$ et l'inégalité est vraie. On suppose maintenant $E[X^p] > 0$. Quitte à considérer $X \wedge n$ et $Y \wedge n$ puis à passer à la limite quand $n \to \infty$ par le théorème de convergence monotone, on peut supposer X et Y bornées. On définit une probabilité Q par

$$Q(\mathrm{d}\omega) = \frac{X^p}{E[X^p]} P(\mathrm{d}\omega)$$

et une variable aléatoire Z par

$$Z = \mathbb{1}_{X>0} \frac{Y}{X^{p-1}}.$$

On applique l'inégalité de Jensen à Z :

$$E_Q[Z]^q \leqslant E_Q[Z^q].$$

Or

$$E_Q[Z] = \int_\Omega \frac{Y(\omega)}{X(\omega)^{p-1}} \, Q(\mathrm{d}\omega) = \int_\Omega \frac{Y(\omega)X(\omega)}{E[X^p]} \, P(\mathrm{d}\omega) = \frac{E[XY]}{E[X^p]}.$$

De même, $E_Q[Z^q] = \frac{E[Y^q]}{E[X^p]}$. D'où le résultat. □

Proposition. *Soit $\omega(h) = \sup_{|u-v| \leqslant h} |f(u) - f(v)|$ le module de continuité uniforme de $f : \mathbb{R} \to \mathbb{C}$. Alors ω est sous-additif, i.e. $\omega(s+t) \leqslant \omega(s) + \omega(t)$. En conséquence, pour tout $\lambda, h \geqslant 0$, $\omega(\lambda h) \leqslant (\lambda + 1)\omega(h)$.*

Démonstration. Soient $s,t>0$, et u,v tels que $|u-v|\leqslant s+t$. On a :

$$\begin{aligned}|f(u)-f(v)| &\leqslant |f(u)-f(w)|+|f(w)-f(v)|\\ &\leqslant \sup_{|u-v|\leqslant s+t}|f(u)-f(w)|+\sup_{|u-v|\leqslant s+t}|f(w)-f(v)|.\end{aligned}$$

On choisit $w=u+s\frac{v-u}{|v-u|}$. Alors $|w-u|=s\leqslant s$, donc

$$\sup_{|u-v|\leqslant s+t}|f(u)-f(w)|\leqslant\omega(s).$$

D'autre part,

$$|w-v|=\left|u+s\frac{v-u}{|v-u|}-v\right|=\left|(u-v)\left(\frac{|v-u|-s}{|v-u|}\right)\right|=\big||v-u|-s\big|.$$

Or si $|u-v|\leqslant s+t$, alors

$$-s\leqslant|u-v|-s\leqslant t.$$

Quitte à inverser s et t, on peut supposer $s\leqslant t$. Alors $-t\leqslant -s$ et on en déduit :

$$-t\leqslant|u-v|-s\leqslant t,$$

i.e. $|w-v|\leqslant t$. Il vient donc :

$$\sup_{|u-v|\leqslant s+t}|f(w)-f(v)|\leqslant\omega(t).$$

Finalement, pour tout u,v avec $|u-v|\leqslant s+t$, on a

$$|f(u)-f(v)|\leqslant\omega(s)+\omega(t),$$

d'où $\omega(s+t)\leqslant\omega(s)+\omega(t)$.

Pour montrer la conséquence, on commence par montrer que $\omega(nh)\leqslant n\omega(h)$ pour tout $n\in\mathbb{N}$ et tout $h\geqslant 0$. C'est immédiat avec la sous-additivité de ω. Ensuite on fixe $\lambda\geqslant 0$ et on l'encadre par $n\leqslant\lambda<n+1$ où $n=\lfloor\lambda\rfloor$. Alors :

$$\omega(\lambda h)\leqslant\omega((n+1)h)\leqslant(n+1)\omega(h)\leqslant(\lambda+1)\omega(h).$$

□

Théorème (Inégalité de Khintchine)**.** *Soient $X_1,\dots,X_n$ des variables i.i.d. de loi $\frac{1}{2}\delta_{-1}+\frac{1}{2}\delta_1$. Soit $Y\in\mathrm{Vect}_{\mathbb{R}}(X_1,\dots,X_n)$. Alors*

$$\|Y\|_2\leqslant\sqrt{e}\|Y\|_1.$$

Démonstration. On écrit $Y = \sum_{k=1}^n a_k X_k$. Comme les X_k sont $i.i.d.$ avec $E[X_k] = 0$ et $E[X_k^2] = 1$, on a $\|Y\|_2^2 = \sum_{k=1}^n a_k^2$. On remarque qu'on peut supposer $\|Y\|_2 = 1$, quitte à diviser Y par $\|Y\|_2$.

On pose $Z = \prod_{k=1}^n (1 + ia_k X_k)$ variable aléatoire complexe. Alors

$$|Z| = \prod_{k=1}^n \sqrt{1 + a_k^2 X_k^2} = \prod_{k=1}^n \sqrt{1 + a_k^2} \leqslant \prod_{k=1}^n \sqrt{e^{a_k^2}} = e^{\frac{1}{2}\|Y\|_2^2} = \sqrt{e}.$$

Donc $\|Z\|_\infty \leqslant \sqrt{e}$. De plus, pour $j \in [\![1, n]\!]$, on a :

$$\begin{aligned} E[X_j Z] &= E\left[X_j(1 + ia_j X_j) \prod_{k \neq j} (1 + ia_k X_k)\right] \\ &= \underbrace{E\big[X_j + ia_j X_j^2\big]}_{=ia_j} \prod_{k \neq j} \underbrace{E[1 + ia_k X_k]}_{=1} \\ &= ia_j, \end{aligned}$$

en utilisant $E[X_k] = 0$ et $E[X_k^2] = 1$. Donc

$$E[YZ] = \sum_{k=1}^n a_k E[X_k Z] = i \sum_{k=1}^n a_k^2 = i.$$

On en déduit enfin :

$$|E[YZ]\,| = 1 = \|Y\|_2 \leqslant \|Z\|_\infty \|Y\|_1 \leqslant \sqrt{e}\|Y\|_1.$$

□

Références :

– Zuily et Queffélec - *Eléments d'analyse pour l'agrégation* - Page 518.

3.17 Preuve probabiliste de la formule de Stirling

Théorème (Formule de Stirling)**.** *On a l'équivalent suivant :*

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

Démonstration. Soit $(X_n)_{n\geqslant 1}$ une suite de $v.a.i.i.d.$ de loi de Poisson de paramètre 1. On pose

$$S_n = \sum_{k=1}^{n} X_k \quad \text{et} \quad Z_n = \frac{S_n - n}{\sqrt{n}}.$$

Une somme de $v.a.i.$ de Poisson étant encore une $v.a.$ de Poisson, on a

$$S_n \overset{\text{loi}}{=} \mathcal{P}(n)\,,\ E[S_n] = n,\ \text{et } \mathrm{Var}(S_n) = n.$$

La première étape est d'appliquer le théorème central limite pour avoir la convergence en loi de (Z_n) vers une gaussienne centrée réduite. Les X_n sont bien $i.i.d.$ de carré intégrable, avec $E[X_1] = 1$ et $\sigma = \sqrt{\mathrm{Var}(X_1)} = 1$, donc le théorème s'applique et nous donne :

$$P(Z_n > t) = P\left(\frac{S_n - n}{\sqrt{n}} > t\right) \underset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-\frac{u^2}{2}}\,\mathrm{d}u = P(Z > t),$$

pour tout $t \in \mathbb{R}$, où Z est une $v.a.$ de loi gaussienne centrée réduite.

Soit maintenant $t > 0$. Comme $\{Z_n > t\} \subset \{Z_n^2 > t^2\}$, on a

$$P(Z_n > t) \leqslant P(Z_n^2 > t^2) \leqslant \frac{E[Z_n^2]}{t^2}$$

par l'inégalité de Markov (qui s'applique car $Z_n^2 \geqslant 0$). Ensuite

$$\frac{E[Z_n^2]}{t^2} = \frac{\mathrm{Var}(Z_n) + E[Z_n]^2}{t^2} = \frac{1}{t^2}$$

car $\mathrm{Var}(Z_n) = \mathrm{Var}\left(\frac{S_n - n}{\sqrt{n}}\right) = \left(\frac{1}{\sqrt{n}}\right)^2 \mathrm{Var}(S_n) = 1$ et $E[Z_n] = \frac{1}{\sqrt{n}}(E[S_n] - n) = 0$. En fin de compte,

$$P(Z_n > t) \leqslant \begin{cases} 1 \text{ si } t \leqslant 1 \\ \frac{1}{t^2} \text{ si } t > 1, \end{cases}$$

et on a ainsi majoré $P(Z_n > t)$ par une fonction intégrable sur $\mathbb{R}_+^*$ indépendante de n. Par le théorème de convergence dominée, on obtient

$$\int_0^\infty P(Z_n > t)\,\mathrm{d}t \underset{n\to\infty}{\longrightarrow} \int_0^\infty P(Z > t)\,\mathrm{d}t.$$

On calcule d'une part

$$\begin{aligned}\int_0^\infty P(Z>t)\,\mathrm{d}t &= \int_0^\infty \left(\frac{1}{\sqrt{2\pi}}\int_t^\infty e^{-\frac{u^2}{2}}\,\mathrm{d}u\right)\mathrm{d}t\\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty\int_0^u e^{-\frac{u^2}{2}}\,\mathrm{d}t\,\mathrm{d}u\\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty u e^{-\frac{u^2}{2}}\,\mathrm{d}u\\ &= \frac{1}{\sqrt{2\pi}}\left[-e^{-\frac{u^2}{2}}\right]_0^\infty = \frac{1}{\sqrt{2\pi}},\end{aligned}$$

où le théorème de Fubini-Tonelli nous a permis d'intervertir les intégrales (tout est positif). On calcule ensuite

$$\begin{aligned}\int_0^\infty P(Z_n>t)\,\mathrm{d}t &= \int_0^\infty P(S_n>t\sqrt{n}+n)\,\mathrm{d}t\\ &= \int_0^\infty \sum_{k=0}^\infty P(S_n=k)\mathbb{1}_{\{k>t\sqrt{n}+n\}}\,\mathrm{d}t\\ &= \sum_{k=0}^\infty \int_0^\infty P(S_n=k)\mathbb{1}_{\{k>t\sqrt{n}+n\}}\,\mathrm{d}t\\ &= \sum_{k=n+1}^\infty \int_0^{\frac{k-n}{\sqrt{n}}} P(S_n=k)\,\mathrm{d}t\\ &= \sum_{k=n+1}^\infty \frac{k-n}{\sqrt{n}}P(S_n=k)\\ &= \frac{1}{\sqrt{n}}\sum_{k=n+1}^\infty (k-n)e^{-n}\frac{n^k}{k!}\\ &= \frac{1}{\sqrt{n}e^n}\left(\sum_{k=n+1}^\infty \frac{n^k}{(k-1)!}-\sum_{k=n+1}^\infty \frac{n^{k+1}}{k!}\right)\\ &= \frac{1}{\sqrt{n}e^n}\frac{n^{n+1}}{n!}.\end{aligned}$$

D'où enfin

$$\int_0^\infty P(Z_n>t)\,\mathrm{d}t = \left(\frac{n}{e}\right)^n\sqrt{n}\frac{1}{n!}\underset{n\to\infty}{\longrightarrow}\frac{1}{\sqrt{2\pi}},$$

ce qui nous donne la formule de Stirling. □

Proposition. *Soient X et Y deux v.a.i. de loi de Poisson de paramètres respectifs $\lambda>0$ et $\mu>0$. Alors $X+Y$ suit une loi de Poisson de paramètre $\lambda+\mu$.*

Démonstration. On a

$$G_X(t) = E\big[t^X\big] = \sum_{k=0}^{\infty} t^k e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}e^{t\lambda} = e^{\lambda(t-1)}.$$

Par indépendance, $G_{X+Y}(t) = G_X(t)G_Y(t) = e^{(\lambda+\mu)(t-1)}$. Donc G_{X+Y} est la fonction génératrice d'une $v.a.$ de loi $\mathcal{P}(\lambda+\mu)$. Donc $X+Y \overset{\text{loi}}{=} \mathcal{P}(\lambda+\mu)$.

Rappel : On s'est servi du fait que la fonction génératrice caractérise entièrement la loi d'une $v.a.$. En effet, si X et Y sont deux $v.a.$ à valeurs entières telles que $G_X(t) = G_Y(t)$ pour tout $|t| < 1$ (la série définissant G_X converge au moins pour $|t| < 1$ puisque $G_X(1) = 1 < \infty$), l'unicité du développement d'une fonction en série entière montre que X et Y ont la même loi. □

Théorème (Théorème central limite)**.** *Soit (X_i) une suite de v.a.i.i.d. de carré intégrable (et non constantes). On note $\mu = E[X_1]$ et $\sigma^2 = \mathrm{Var}(X_1)$ avec $\sigma > 0$. Alors en notant $S_n = X_1 + \cdots + X_n$, on a*

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{loi} \mathcal{N}(0,1),$$

ce qui s'écrit encore

$$P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leqslant x\right) \underset{n\to\infty}{\longrightarrow} \phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,\mathrm{d}t,$$

pour tout $x \in \mathbb{R}$.

Références :

- Aucune.

3.18 Projection sur un convexe fermé et représentation de Riesz

Soient H un espace de Hilbert et C un convexe fermé non vide de H. Tout est fait ici dans le cas d'un espace de Hilbert réel mais fonctionne exactement de la même manière en adaptant avec un produit scalaire hermitien.

Théorème (Projection sur un convexe fermé)**.** *Pour tout $x \in H$, il existe un unique vecteur $y \in C$ tel que $d(x,C) = \|x-y\|$. On appelle y le projeté orthogonal de x sur C.*

Démonstration. On pose $d = d(x,C) = \inf_{h\in C} \|x-h\|$. Par définition de la borne inférieure, il existe une suite (h_n) d'éléments de C telle que pour tout $n \in \mathbb{N}$, $\|h_n - x\|^2 \leqslant d^2 + \frac{1}{n}$. On utilise l'identité du parallélogramme

$$\|z-z'\|^2 = 2\|z\|^2 + 2\|z'\|^2 - \|z+z'\|^2$$

avec $z = h_n - x$ et $z' = h_p - x$:

$$\begin{aligned}
\|h_n - h_p\|^2 &= 2\|h_n - x\|^2 + 2\|h_p - x\|^2 - \|h_n + h_p - 2x\|^2 \\
&= 2\|h_n - x\|^2 + 2\|h_p - x\|^2 - 4\left\|\frac{h_n+h_p}{2} - x\right\|^2 \\
&\leqslant 2d^2 + \frac{2}{n} + 2d^2 + \frac{2}{p} - 4d^2 \leqslant \frac{2}{n} + \frac{2}{p},
\end{aligned}$$

car C étant convexe, $\frac{h_n+h_p}{2} \in C$, donc $\left\|\frac{h_n+h_p}{2} - x\right\| \geqslant d$. Donc (h_n) est de Cauchy, donc convergente dans H qui est complet. On note y sa limite. Comme C est fermé, $y \in C$. Par continuité de la norme, $\|y-x\|^2 \leqslant d^2$, donc $\|y-x\| = d$ et on a prouvé l'existence du projeté orthogonal.

Unicité : supposons qu'il existe $y, y' \in C$ tels que $y \neq y'$ et $d = \|y-x\| = \|y'-x\|$. Toujours par l'identité du parallélogramme, on a :

$$\begin{aligned}
\left\|\frac{1}{2}(y+y') - x\right\|^2 &= \left\|\frac{1}{2}(y-x) + \frac{1}{2}(y'-x)\right\|^2 \\
&= \frac{1}{2}(\|y-x\|^2 + \|y'-x\|^2) - \frac{1}{4}\|y-y'\|^2 \\
&= d^2 - \frac{1}{4}\|y-y'\|^2 < d^2,
\end{aligned}$$

ce qui est absurde car C étant convexe, $\frac{1}{2}(y+y') \in C$, donc $\left\|\frac{1}{2}(y+y') - x\right\| \geqslant d$. □

Proposition. *Soient $x \in H$ et $y \in C$. Alors y est le projeté orthogonal de x sur C si et seulement si $(y-x, y-z) \leqslant 0$ pour tout $z \in C$.*

Démonstration. Supposons que y est le projeté orthogonal de x sur C. Soit $z \in C$. Pour tout $\lambda \in [0,1]$, $(1-\lambda)y + \lambda z \in C$, donc par définition du projeté,

$$\|x - (1-\lambda)y - \lambda z\| = \|x - y + \lambda(y-z)\| \geqslant \|x-y\|.$$

En développant, on en déduit

$$\|x-y\|^2 + 2\lambda(x-y, y-z) + \lambda^2\|y-z\|^2 \geqslant \|x-y\|^2,$$

soit encore, pour $\lambda \neq 0$,

$$2(x-y, y-z) + \lambda\|y-z\|^2 \geqslant 0.$$

En faisant tendre λ vers 0, on obtient $(y-x, y-z) \leqslant 0$.

Réciproquement, si cette condition est satisfaite, on a

$$\|z-x\|^2 = \|(y-x)-(y-z)\|^2 = \|y-x\|^2 + \|y-z\|^2 - 2(y-x, y-z) \geqslant \|y-x\|^2$$

pour tout $z \in C$. Donc y est le projeté orthogonal de x sur C. □

On note p l'application de H sur C qui à $x \in H$ associe son projeté orthogonal sur C.

Théorème. *On suppose que C est un sous-espace vectoriel fermé de H. Alors $H = C \oplus C^\perp$ et p est une application linéaire continue.*

Démonstration. Soient $x \in H$, $y = p(x)$, $z \in C$. Alors $t = y - z \in C$ (sous-espace vectoriel), donc d'après la proposition, $(y-x, y-t) = (y-x, z) \leqslant 0$. Comme $-z \in C$, on a aussi $(y-x, -z) \leqslant 0$. Donc $(y-x, z) = 0$. Comme ceci est vrai pour tout $z \in C$, alors $x - y = x - p(x) \in C^\perp$. On écrit $x = p(x) + (x - p(x)) \in C + C^\perp$. La somme est évidemment directe, donc $H = C \oplus C^\perp$.

On en déduit ensuite que p est linéaire car c'est la projection orthogonale sur C au sens habituel. On peut vérifier : si $x = y + z$ et $x' = y' + z'$ avec $y, y' \in C$, $z, z' \in C^\perp$, et si $\lambda \in \mathbb{R}$, alors

$$p(\lambda x + x') = p(\underbrace{\lambda y + y'}_{=Y \in C} + \underbrace{\lambda z + z'}_{=Z \in C^\perp}) = \lambda y + y' = \lambda p(x) + p(x'),$$

car $Y + Z = p(Y+Z) + (Y + Z - p(Y+Z))$, et par unicité de la décomposition, $Y = p(Y+Z) \in C$ et $Z = Y + Z - p(Y+Z) \in C^\perp$.

Montrons enfin la continuité de p. Soit $z \in H$ qu'on écrit $z = x + y$ avec $x \in C$ et $y \in C^\perp$. Alors

$$\|p(z)\|^2 = \|x\|^2 \leqslant \|x\|^2 + \|y\|^2 = \|x+y\|^2 = \|z\|^2.$$

Donc p est bien continue avec de plus $\|p\| \leqslant 1$. □

Théorème (Représentation de Riesz-Fréchet)**.** *Soit f une forme linéaire continue sur H. Alors il existe un unique vecteur $a \in H$ tel que $f(x) = (a, x)$ pour tout $x \in H$.*

Démonstration. Si f est l'application nulle, $a = 0$ convient. Sinon, $\mathrm{Ker}(f)$ est un sous-espace vectoriel de H, fermé car f est continue. D'après le théorème précédent, $H = \mathrm{Ker}(f) \oplus \mathrm{Ker}(f)^\perp$. Or $\mathrm{Ker}(f)^\perp \neq \{0\}$ car $f \neq 0$. Soit alors $h \in \mathrm{Ker}(f)^\perp$, $h \neq 0$. Pour $x \in H$, on a $x - \frac{f(x)}{f(h)}h \in \mathrm{Ker}(f)$, d'où

$$\left(h, x - \frac{f(x)}{f(h)}h\right) = 0.$$

En développant, on obtient $(h, x) = \frac{\|h\|^2}{f(h)} f(x)$, soit encore

$$f(x) = \left(\frac{f(h)}{\|h\|^2}h, x\right).$$

Le vecteur $a = \frac{f(h)}{\|h\|^2}h$ convient donc.

Montrons qu'il est unique : soit $a' \in H$ tel que pour tout $x \in H$, $f(x) = (a, x) = (a', x)$. Alors $(a - a', x) = 0$ pour tout $x \in H$, donc $a - a' = 0$. □

On pourrait admettre le premier théorème (projection sur un convexe fermé) pour aboutir au théorème de représentation de Riesz-Fréchet (leçon sur les espaces vectoriels normés et applications linéaires continues), ou inversement s'arrêter avant le dernier théorème (représentation de Riesz-Fréchet) (leçon sur les espaces de Hilbert).

On peut utiliser le théorème de représentation de Riesz-Fréchet pour montrer l'existence et l'unicité de l'adjoint d'une application linéaire continue entre espaces de Hilbert.

Théorème. *Soient E, F deux espaces de Hilbert et $T \in \mathcal{L}(E, F)$. Alors il existe une unique application $T^* \in \mathcal{L}(F, E)$ telle que pour tout $x \in E$ et tout $y \in F$,*

$$(T(x), y) = (x, T^*(y)).$$

De plus, on a $\|T\| = \|T^*\|$.

Démonstration. Pour $y \in F$, on considère l'application

$$\phi_y : x \mapsto (T(x), y),$$

qui est clairement linéaire. Elle est continue par l'inégalité de Cauchy-Schwarz :

$$|\phi_y(x)| \leqslant \|T(x)\|\|y\| \leqslant \|T\|\|x\|\|y\| = C\|x\|$$

en posant $C = \|T\|\|y\|$. Donc d'après le théorème de Riesz-Fréchet, il existe un unique vecteur, qu'on note $T^*(y)$, tel que

$$\phi_y(x) = (T(x), y) = (x, T^*(y))$$

pour tout $x \in E$. On a ainsi défini une application $T^* : F \to E$.

Montrons que T^* est linéaire. Soient $y_1, y_2 \in F$ et $\lambda \in \mathbb{R}$. Alors pour tout $x \in E$, on a :

$$\begin{aligned}(x, T^*(\lambda y_1 + y_2)) &= (T(x), \lambda y_1 + y_2) \\ &= \lambda(T(x), y_1) + (T(x), y_2) \\ &= \lambda(x, T^*(y_1)) + (x, T^*(y_2)) \\ &= (x, \lambda T^*(y_1) + T^*(y_2)).\end{aligned}$$

On en déduit immédiatement la linéarité de T^*.

Montrons que T^* est continue. Pour $x \in E$ et $y \in F$, on a

$$|(x, T^*(y))| = |(T(x), y)| \leqslant \|T\|\|x\|\|y\|.$$

Mais on a

$$\|T^*(y)\| = \sup_{\|x\|=1} |(x, T^*(y))|,$$

ce qui se montre facilement avec l'inégalité de Cauchy-Schwarz en sachant que c'est une égalité lorsque les deux vecteurs sont colinéaires. Ainsi

$$\|T^*(y)\| \leqslant \|T\|\|y\|,$$

i.e. T^* est continue et $\|T^*\| \leqslant \|T\|$. Par unicité de l'adjoint, on a $(T^*)^* = T$, d'où :

$$\|T\| = \|(T^*)^*\| \leqslant \|T^*\|,$$

et enfin $\|T\| = \|T^*\|$. □

Références :

- Francinou, Gianella, Nicolas - *Oraux X-ENS, Analyse 3* - Page 152.
- Skandalis - *Topologie et analyse, 3ème année* - Page 252 (raccourci pour montrer que la projection est continue), page 261 (pour l'application à l'existence d'un adjoint).

3.19 Sous-espaces vectoriels fermés de $L^p(\mu)$

Théorème. *Soit $(X, \mathcal{F}, \mu)$ un espace mesuré de mesure positive finie, et soit $1 < p < \infty$. Si F est un sous-espace vectoriel fermé de $L^p(\mu)$ et $F \subset L^\infty(\mu)$, alors $\dim_{\mathbb{C}}(F) < \infty$.*

Démonstration. Soit F un sous-espace vectoriel fermé de $L^p(\mu)$ avec $F \subset L^\infty(\mu)$. Par emboîtement décroissant (X est de mesure finie), on a $F \subset L^\infty \subset L^2$. On va montrer le résultat préliminaire suivant : il existe $\alpha > 0$ tel que $\|f\|_\infty \leqslant \alpha \|f\|_2$, pour tout $f \in F$.

Montrons que les normes $\|\cdot\|_p$ et $\|\cdot\|_\infty$ sont équivalentes sur F. D'abord, l'injection $L^\infty \hookrightarrow L^p$ est continue car X est de mesure finie. On montre ensuite que l'injection $(F, \|\cdot\|_p) \hookrightarrow L^\infty$ est aussi continue grâce au théorème du graphe fermé :

(*i*) L^∞ est un Banach et $(F, \|\cdot\|_p)$ aussi car F est fermé dans L^p qui est complet, donc complet.

(*ii*) Ensuite si (f_n) est une suite de F (donc aussi une suite de L^∞) telle que $f_n \underset{\|\cdot\|_p}{\longrightarrow} f \in F$ et $f_n \underset{\|\cdot\|_\infty}{\longrightarrow} g \in L^\infty$, on sait qu'il existe une sous-suite de (f_n) qui converge vers f presque partout (voir le développement sur la complétude de $L^p(\mu)$). Donc $f = g$ presque partout, *i.e.* $f = g$ dans L^∞.

Le graphe de l'injection est donc fermé et celle-ci est continue. La continuité des deux injections prouve que les normes $\|\cdot\|_p$ et $\|\cdot\|_\infty$ sont équivalentes sur F.

L'injection $(F, \|\cdot\|_2) \hookrightarrow L^p$ est continue : si $p \leqslant 2$, c'est clair car X est de mesure finie (emboîtement décroissant). Si $p > 2$, pour $f \in F$, on a $\|f\|_p^p = \int_X |f|^{p-2}|f|^2 \,\mathrm{d}\mu$, et donc $\|f\|_p^p \leqslant \|f\|_\infty^{p-2}\|f\|_2^2$. Mais les normes $\|\cdot\|_p$ et $\|\cdot\|_\infty$ sont équivalentes sur F : il existe $C > 0$ tel que $\|\cdot\|_\infty \leqslant C\|\cdot\|_p$. Donc

$$\|f\|_p^p \leqslant C^{p-2}\|f\|_p^{p-2}\|f\|_2^2,$$

et enfin : $\|f\|_p \leqslant C^{\frac{p-2}{2}}\|f\|_2$.

Finalement, comme $\|\cdot\|_p$ et $\|\cdot\|_\infty$ sont équivalentes sur F, il existe $\alpha > 0$ tel que $\|f\|_\infty \leqslant \alpha\|f\|_2$, pour tout $f \in F$.

Soit $(f_1, \ldots, f_n)$ une famille libre de fonctions de F, qu'on peut supposer orthonormale dans L^2 muni du produit scalaire induit par la norme $\|\cdot\|_2$ (quitte à utiliser le procédé de Gram-Schmidt). On note B la boule unité fermée de $\mathbb{C}^n$ pour la norme $\|\cdot\|_2$. Pour $c = (c_1, \ldots, c_n) \in B$, on pose

$$f_c = \sum_{i=1}^{n} c_i f_i.$$

On peut noter que $\|f_c\|_2 \leqslant 1$:

$$\|f_c\|_2^2 = \sum_{i=1}^{n} \|c_i f_i\|_2^2 = \sum_{i=1}^{n} |c_i|^2 = \|c\|_2^2 \leqslant 1.$$

Sachant que B est séparable (car inclu dans $\mathbb{C}^n$ séparable), il existe une suite $(c(k))_{k\in\mathbb{N}}$ dense dans B. Pour tout k, il existe une partie mesurable Ω_k de X, avec $\mu(\Omega_k) = \mu(X)$, telle que $\|f_{c(k)}\|_\infty = \sup_{x\in\Omega_k} |f_{c(k)}(x)|$. Alors $\Omega = \bigcap_{k=1}^{\infty} \Omega_k$ est encore une partie mesurable avec $\mu(\Omega) = \mu(X)$ (le complémentaire de Ω est une réunion dénombrable d'ensembles de mesure nulle, donc de mesure nulle).

Pour tout $x \in \Omega$ et tout $k \in \mathbb{N}$, $|f_{c(k)}(x)| \leqslant \|f_{c(k)}\|_\infty$. D'après le résultat préliminaire,

$$|f_{c(k)}(x)| \leqslant \|f_{c(k)}\|_\infty \leqslant \alpha \|f_{c(k)}\|_2 \leqslant \alpha.$$

Par densité, il s'ensuit immédiatement que $|f_c(x)| \leqslant \alpha$ pour tout $c \in B$ et tout $x \in \Omega$ (la convergence d'une suite $(c(k_i))_{i\in\mathbb{N}}$ de $\mathbb{C}^n$ pour la norme $\|\cdot\|_2$ implique la convergence composante par composante).

Pour $x \in \Omega$, on pose :

$$c(x) = \begin{cases} \dfrac{\left(\overline{f_1(x)},\ldots,\overline{f_n(x)}\right)}{\left\|\left(\overline{f_1(x)},\ldots,\overline{f_n(x)}\right)\right\|_2} & \text{s'il existe } i \text{ tel que } f_i(x) \neq 0 \\ 0 \text{ sinon.} \end{cases}$$

Il vient alors

$$f_{c(x)}(x) = \frac{1}{\left\|\left(\overline{f_1(x)},\ldots,\overline{f_n(x)}\right)\right\|_2} \sum_{i=1}^{n} \overline{f_i(x)} f_i(x) = \|(f_1(x),\ldots,f_n(x))\|_2 .$$

En élevant au carré, et sachant que $|f_{c(x)}(x)|^2 \leqslant \alpha^2$, on obtient :

$$\sum_{i=1}^{n} |f_i(x)|^2 \leqslant \alpha^2.$$

L'inégalité ainsi obtenue étant vraie pour tout $x \in \Omega$, on intégre sur Ω :

$$\sum_{i=1}^{n} \underbrace{\int_\Omega |f_i|^2 \,\mathrm{d}\mu}_{=\int_X |f_i|^2 \,\mathrm{d}\mu = \|f_i\|_2^2 = 1} \leqslant \alpha^2 \mu(\Omega) = \alpha^2 \mu(X).$$

D'où finalement :

$$n \leqslant \alpha^2 \mu(X),$$

et par conséquent F est de dimension finie. □

Remarque. L'intérêt de s'être ramené à L^2 est que c'est le seul des espaces L^p à être muni d'un produit scalaire hermitien :

$$(f,g) = \int_X \overline{f} g \, \mathrm{d}\mu$$

qui en fait un espace de Hilbert, *i.e.* un espace vectoriel complet pour sa norme $\|\cdot\|_2$ induite par ce produit scalaire hermitien. On a pu ainsi définir une famille orthonormée dans L^2. On note également que ce produit scalaire est bien défini car si $f, g \in L^2$, $fg \in L^1$ grâce à Cauchy-Schwarz (ou l'inégalité de Hölder avec $p = 2$ et $q = 2$).

Remarque. Les hypothèses du théorème ne peuvent pas être diminuées. En effet, si $F \not\subset L^\infty$ ou si $p = +\infty$, alors si on prend $F = L^p$, F est de dimension infinie. Pour voir que L^p est de dimension infinie, par exemple lorsque le domaine X d'intégration est $X = [0,1]$ (de mesure finie), on peut prendre la famille des indicatrices $\mathbb{1}_{\left[\frac{1}{n}, \frac{1}{n+1}\right[}$ pour $n \in \mathbb{N}^*$ qui est libre dans L^p pour $p \leqslant \infty$.

Si on suppose maintenant $F \subset L^\infty$ et $p < \infty$, mais $\mu(X) = \infty$, le théorème est encore faux. On prend par exemple $X = [0, \infty[$ et F l'adhérence dans L^p de l'espace vectoriel de L^p engendré par les fonctions indicatrices $\mathbb{1}_{[n,n+1[}$. On vérifie que $F \subset L^\infty$: si $f \in F$, il existe une suite de complexes $(\lambda_i)_{i \geqslant 1}$ telle que

$$f = \sum_{i=0}^{\infty} \lambda_i \mathbb{1}_{[i,i+1[}.$$

Comme $f \in L^p$, on a $\sum_{i=0}^{\infty} |\lambda_i|^p < \infty$, donc $\lambda_i \to 0$, donc la suite (λ_i) est bornée, et donc f est bornée, *i.e.* $f \in L^\infty$. Pourtant, les $\mathbb{1}_{[n,n+1[}$ forment une famille libre de F, qui est donc de dimension infinie.

Dans ce qui suit, on désigne par X un espace mesuré quelconque muni d'une mesure positive μ.

On rappelle que $L^p(\mu)$ est un espace vectoriel complexe. C'est immédiat avec l'inégalité de Minkowski qui montre que si $f, g \in L^p(\mu)$, alors $f + g \in L^p(\mu)$, et le fait que $\|\alpha f\|_p = |\alpha| \|f\|_p$. On peut donc bien sûr parler de la dimension de F sous-espace vectoriel fermé de $L^p(\mu)$ comme $\mathbb{C}$-espace vectoriel.

Théorème (Emboîtement décroissant)**.** *On suppose X de mesure fini : $\mu(X) < \infty$. Alors pour $1 \leqslant p \leqslant q$, on a $L^q \subset L^p$ et l'injection est continue. On a de plus, pour $f \in L^p$:*

$$\|f\|_p \leqslant \mu(X)^{\frac{1}{p} - \frac{1}{q}} \|f\|_q.$$

Démonstration. On applique l'inégalité de Hölder avec $\frac{1}{p'} = \frac{p}{q}$ et $\frac{1}{q'} = 1 - \frac{p}{q}$:

$$\int_X |f|^p \, \mathrm{d}\mu \leqslant \left(\int_X (|f|^p)^{p'} \, \mathrm{d}\mu \right)^{\frac{1}{p'}} \left(\int_X 1 \, \mathrm{d}\mu \right)^{\frac{1}{q'}}.$$

On élève ensuite les deux membres à la puissance $\frac{1}{p}$ et on a le résultat. □

Remarque. On a donc $L^\infty \subset L^p$. Par exemple si $X = [0,1]$, X est de mesure finie, mais $L^p \not\subset L^\infty$: prendre la fonction $x \mapsto \left(\frac{1}{\sqrt{x}}\right)^{\frac{1}{p}}$.

Théorème (Théorème du graphe fermé)**.** *Soient E et F deux espaces de Banach, et f une application linéaire de E dans F. On suppose que pour toute suite (x_n) de E telle que $x_n \to x \in E$ et $f(x_n) \to y \in F$, on a $y = T(x)$ (i.e. le graphe de f est fermé). Alors f est continue.*

Démonstration. On note

$$\Gamma_f = \{(x, f(x)),\, x \in E\}$$

le graphe de f. Comme Γ_f est supposé fermé dans $E \times F$ complet, on a que Γ_f est complet. Soient

$$p_1 : \begin{array}{ccc} \Gamma_f & \longrightarrow & E \\ (x, f(x)) & \longmapsto & x \end{array} \quad \text{et} \quad p_2 : \begin{array}{ccc} \Gamma_f & \longrightarrow & F \\ (x, f(x)) & \longmapsto & f(x). \end{array}$$

Ces deux applications sont continues. De plus, p_1 est bijective, donc par le théorème d'isomorphisme de Banach, p_1^{-1} est continue. Finalement, $f = p_2 \circ p_1^{-1}$ est continue. □

Théorème (Théorème d'isomorphisme de Banach)**.** *Soient E et F deux espaces de Banach, et f une application linéaire continue bijective de E dans F. Alors f^{-1} est continue.*

Démonstration. f^{-1} est continue car f est ouverte par le théorème de l'application ouverte. □

Théorème (Théorème de l'application ouverte)**.** *Soient E et F deux espaces de Banach, et f une application linéaire continue surjective de E dans F. Alors il existe $c > 0$ tel que*

$$B_F(0, c) \subset f\big(B_E(0, 1)\big).$$

En particulier, f est une application ouverte.

Démonstration. Résulte du théorème de Baire. □

Définition. On dit qu'un espace topologique $(X, \mathcal{T})$ est *séparable* s'il contient une partie A au plus dénombrable dense dans X.

Exemple. *L'ensemble $\mathbb{R}$ des réels est séparable car il contient $\mathbb{Q}$ qui y est dense et dénombrable. Par conséquent, $\mathbb{R}^n$ est séparable, et $\mathbb{C}^n = \mathbb{R}^{2n}$ aussi.*

Références :

– Rudin - *Analyse fonctionnelle.*

3.20 Théorème de Borel

Proposition. *Soient a, b, c, d quatre réels avec $a < b < c < d$. Alors il existe une fonction $f \in \mathcal{C}^\infty(\mathbb{R})$ comprise entre 0 et 1, égale à 1 sur $[b, c]$ et nulle en dehors de $[a, d]$.*

Démonstration. Soit ψ la fonction définie par

$$\psi : t \mapsto \begin{cases} e^{-\frac{1}{t}} \text{ si } t > 0 \\ 0 \text{ si } t \leqslant 0. \end{cases}$$

Alors ψ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$. En effet, elle est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^*$ et on montre par une récurrence facile que pour tout $t > 0$ et tout $k \in \mathbb{N}$, il existe un polynôme P_k tel que

$$\psi^{(k)}(t) = P_k\left(\frac{1}{t}\right) e^{-\frac{1}{t}}.$$

On montre ensuite par récurrence sur k que ψ est k fois dérivable sur $\mathbb{R}$ avec $\psi^{(k)}(t) = 0$ pour tout $t \leqslant 0$: c'est vrai pour $k = 0$ et si c'est vrai au rang k, alors

$$\frac{\psi^{(k)}(t) - \psi^{(k)}(0)}{t} = \begin{cases} \frac{1}{t} P_k\left(\frac{1}{t}\right) e^{-\frac{1}{t}} \text{ si } t > 0 \\ 0 \text{ si } t < 0, \end{cases}$$

qui tend vers 0 quand $t \to 0$, et donc $\psi^{(k)}$ est dérivable en 0 et $\psi^{(k+1)}(0) = 0$. Ainsi ψ est indéfiniment dérivable sur $\mathbb{R}$, donc de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.

Soit maintenant la fonction $\alpha : t \mapsto \psi(t)\psi(1-t)$ définie sur $\mathbb{R}$, qui est $\mathcal{C}^\infty$ sur $\mathbb{R}$, positive et nulle en dehors de $[0, 1]$. On définit ensuite la fonction

$$\beta : t \mapsto \frac{\int_0^t \alpha(s)\,\mathrm{d}s}{\int_0^1 \alpha(s)\,\mathrm{d}s},$$

qui est $\mathcal{C}^\infty$ sur $\mathbb{R}$, strictement croissante, nulle pour $t \leqslant 0$ et égale à 1 pour $t \geqslant 1$.

Enfin, la fonction

$$f : x \mapsto \beta\left(\frac{x-a}{b-a}\right) \beta\left(\frac{d-x}{d-c}\right)$$

est $\mathcal{C}^\infty$ sur $\mathbb{R}$, comprise entre 0 et 1, égale à 1 si $x \in [b, c]$, nulle si $x \leqslant a$ et $x \geqslant d$. □

Théorème (Théorème de Borel)**.** *Soit $(a_n)_{n\in\mathbb{N}}$ une suite quelconque de réels. Alors il existe une fonction $u \in \mathcal{C}^\infty(\mathbb{R})$ telle que $u^{(n)}(0) = a_n$ pour tout $n \in \mathbb{N}$.*

Démonstration. D'après la proposition précédente, il existe une fonction $f \in \mathcal{C}^\infty(\mathbb{R})$ égale à 1 pour $|x| \leqslant \frac{1}{2}$ et nulle pour $|x| \geqslant 1$. Notons $f_k(x) = f(\lambda_k x) a_k \frac{x^k}{k!}$ et montrons

que l'on peut choisir les $\lambda_k > 0$ pour assurer la convergence uniforme sur $\mathbb{R}$ de la série $\sum f_k$ et de chaque série dérivée $\sum f_k^{(m)}$ pour $m \in \mathbb{N}$.

Soient $m \in \mathbb{N}$ et $k > m$. Par la formule de Leibniz de dérivation d'un produit, on a :

$$f_k^{(m)}(x) = a_k \sum_{p=0}^{m} \binom{m}{p} \left(\frac{x^{k-p}}{(k-p)!}\right) \left(\lambda_k^{m-p} f^{(m-p)}(\lambda_k x)\right).$$

Comme f et toutes ses dérivées $f^{(p)}$ sont continues et nulles en dehors de $[-1,1]$, on peut trouver un majorant K_p sur $\mathbb{R}$ pour chacune de ces fonctions. Soit M_m le maximum des K_p pour $p \leqslant m$. Pour $|x| \geqslant \frac{1}{\lambda_k}$, $f_k^{(m)}(x) = 0$ car f est nulle en dehors de $[-1,1]$. Pour $|x| \leqslant \frac{1}{\lambda_k}$, on a la majoration :

$$\begin{aligned}
\left|f_k^{(m)}(x)\right| &\leqslant |a_k| M_m \sum_{p=0}^{m} \binom{m}{p} \lambda_k^{m-p} \frac{1}{(k-p)!} \frac{1}{\lambda_k^{k-p}} \\
&\leqslant |a_k| M_m \frac{1}{(k-m)!} \frac{1}{\lambda_k^{k-m}} \sum_{p=0}^{m} \binom{m}{p} \\
&= \frac{2^m M_m}{(k-m)!} \frac{|a_k|}{\lambda_k^{k-m}}.
\end{aligned}$$

On choisit alors

$$\lambda_k = \max(1, |a_k|)$$

de sorte que $\lambda_k^{k-m} \geqslant \lambda_k \geqslant |a_k|$ car $k - m \geqslant 1$. Ainsi :

$$\left|f_k^{(m)}(x)\right| \leqslant \frac{2^m M_m}{(k-m)!},$$

pour tout $x \in \mathbb{R}$ (évidemment vrai aussi pour $|x| \geqslant \frac{1}{\lambda_k}$ où $f_k^{(m)}(x) = 0$). La série de fonctions $\sum_{k \geqslant 0} f_k^{(m)}$ est donc normalement convergente (la série $\sum_{k>m} \frac{2^m M_m}{(k-m)!}$ est convergente et indépendante de x), donc uniformément convergente sur $\mathbb{R}$. On en déduit que $u(x) = \sum_{k=0}^{\infty} f_k(x)$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et que pour tout $x \in \mathbb{R}$ et tout $m \in \mathbb{N}$,

$$u^{(m)}(x) = \sum_{k=0}^{\infty} f_k^{(m)}(x).$$

En particulier, pour tout $m \in \mathbb{N}$,

$$u^{(m)}(0) = \sum_{k=0}^{\infty} f_k^{(m)}(0) = a_m,$$

car $f_k(x) = a_k \frac{x^k}{k!}$ pour tout x avec $|x| \leqslant \frac{1}{2\lambda_k}$, et ainsi

$$f_k^{(m)}(0) = \left(a_k \frac{x^k}{k!}\right)^{(m)}(0) = \begin{cases} 0 \text{ si } k \neq m \\ a_m \text{ si } k = m. \end{cases}$$

□

Remarque. Si la série entière $\sum a_n x^n$ avait un rayon de convergence ∞, le théorème serait évident en prenant $u(x) = \sum_{n=0}^{\infty} a_n x^n$. Si cette série entière avait un rayon de convergence $R > 0$, on pourrait prendre

$$u(x) = \begin{cases} \sum_{n=0}^{\infty} a_n f(x) x^n \text{ si } |x| \leqslant \frac{3R}{4} \\ 0 \text{ si } |x| > \frac{3R}{4}, \end{cases}$$

où f est $\mathcal{C}^\infty$ sur $\mathbb{R}$, égale à 1 pour $|x| \leqslant \frac{R}{4}$, nulle pour $x \geqslant \frac{R}{2}$.

Remarque. Une conséquence du théorème de Borel est l'existence de fonctions différentes de la somme de leur série de Taylor sur tout voisinage de 0. En effet, soit la suite (a_n) définie par $a_n = (n!)^2$ pour tout $n \in \mathbb{N}$. D'après le théorème de Borel, il existe une fonction $u \in \mathcal{C}^\infty(\mathbb{R})$ telle que $u^{(n)}(0) = (n!)^2$ pour tout $n \in \mathbb{N}$. S'il existait un voisinage V de 0 tel que u soit égal à la somme de sa série de Taylor sur V, alors on aurait

$$u(x) = \sum_{n=0}^{\infty} u^{(n)}(0) \frac{x^n}{n!} = \sum_{n=0}^{\infty} n! x^n,$$

pour tout $x \in V$. Ceci est absurde car la série entière $\sum n! x^n$ a un rayon de convergence égal à 0.

Remarque. Le théorème de Borel se généralise aux fonctions de n variables : si $(a_\alpha)_{\alpha \in \mathbb{N}^n}$ est une suite quelconque de réels, alors il existe une fonction $u \in \mathcal{C}^\infty(\mathbb{R}^n)$ telle que $\partial^\alpha u(0) = a_\alpha$ pour tout α (en notation multi-indice). Pour cela, on peut prendre

$$u(x) = \sum_{\alpha} f(\lambda_\alpha x) a_\alpha \frac{x^\alpha}{\alpha!},$$

avec $f \in \mathcal{C}^\infty(\mathbb{R}^n)$, $f(x) = 1$ pour $\|x\| \leqslant \frac{1}{2}$, $f(x) = 0$ pour $\|x\| \geqslant 1$, et les $\lambda_\alpha > 0$ choisis convenablement par des calculs similaires aux précédents.

Références :

- Rouvière - *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation* - Page 359.

3.21 Théorème de Brouwer

Version partielle à présenter :

Théorème (Version faible)**.** *On note B la boule unité fermée de $\mathbb{R}^n$ pour une norme $\|\cdot\|$ quelconque. Alors toute application $f : B \to B$ de classe $\mathcal{C}^1$ admet un point fixe.*

Démonstration. On peut supposer que la norme $\|\cdot\|$ est la norme euclidienne usuelle. Soit $f : B \to B$ de classe $\mathcal{C}^1$. Raisonnons par l'absurde en supposant que f n'a pas de point fixe.

On note S la sphère unité de $\mathbb{R}^n$. Pour $x \in B$, on définit $r(x)$ comme étant le point d'intersection de S avec la demi-droite $[f(x), x\big)$. En écrivant la formule définissant r, on obtient que r est de classe $\mathcal{C}^1$.

On pose maintenant $F_t(x) = (1-t)x + t\,r(x)$ pour $t \in [0,1]$. Il est clair que F_t est une application de classe $\mathcal{C}^1$ de B dans B. En notant λ la mesure de Lebesgue sur $\mathbb{R}^n$, on définit pour $t \in [0,1]$,

$$P(t) = \int_{\mathring{B}} \det\left(DF_t(x)\right) \mathrm{d}\lambda(x),$$

qui est polynomiale en t (car à x fixé, $t \mapsto F_t(x)$ est polynomiale en t).

Nous allons montrer que P est constant et strictement positif pour t suffisamment petit. Comme P est un polynôme, on aura alors que P est constant non nul et on verra ensuite pourquoi c'est absurde.

Si on montre que, pour t suffisamment petit, F_t est un $\mathcal{C}^1$-difféomorphisme de $\mathring{B}$ sur $\mathring{B}$, de jacobien J_{F_t} partout > 0, alors par le théorème de changement de variable, on aura :

$$\begin{aligned} P(t) &= \int_{\mathring{B}} \det\left(DF_t(x)\right) \mathrm{d}\lambda(x) = \int_{\mathring{B}} J_{F_t}(x)\,\mathrm{d}\lambda(x) \\ &= \int_{\mathring{B}} 1\,|J_{F_t}(x)|\,\mathrm{d}\lambda(x) = \int_{F_t(\mathring{B})} 1\,\mathrm{d}\lambda(x) \\ &= \int_{\mathring{B}} 1\,\mathrm{d}\lambda(x) = K, \end{aligned}$$

où $K > 0$ est le volume de $\mathring{B}$, et on aura montré ce qu'on voulait.

Appliquons le théorème d'inversion globale :

(i) Puisque $F_t(x) = (1-t)x + t\,r(x)$, on a

$$J_{F_t}(x) = a_n(x)t^n + \cdots + a_1(x)t + a_0(x)$$

avec $a_0, \ldots, a_n$ des fonctions continues de B dans $\mathbb{R}$. Or $F_0(x) = x$, donc $DF_0(x) = \mathrm{Id}_{\mathbb{R}^n}$ et $J_{F_0}(x) = 1 = a_0(x)$. On pose

$$m = \sup_{\substack{x \in B \\ t \in [0,1]}} \left|a_n(x)t^{n-1} + \cdots + a_1(x)\right|.$$

Alors pour tout $t < \frac{1}{m}$ et tout $x \in B$, il vient :

$$t|a_n(x)t^{n-1} + \cdots + a_1(x)| < 1,$$

d'où

$$J_{F_t}(x) = t\big(a_n(x)t^{n-1} + \cdots + a_1(x)\big) + 1 > 0.$$

(ii) On montre que F_t est injective sur $\mathring{B}$ pour t suffisamment petit. Soient $x, y \in \mathring{B}$ tels que $F_t(x) = F_t(y)$. On a donc :

$$(1-t)(x-y) = t\big(r(y) - r(x)\big).$$

Notons $M = \sup_{x\in B} \|Dr(x)\|$. L'inégalité de la moyenne donne :

$$\|r(x) - r(y)\| \leqslant M\|x-y\|,$$

d'où $(1-t)\|x-y\| \leqslant tM\|x-y\|$. Pour t suffisamment petit, disons $t < \eta$, cela implique $x = y$.

Le théorème d'inversion globale nous dit alors que pour $t < \min\big(\eta, \frac{1}{m}\big)$, F_t est un $\mathcal{C}^1$-difféomorphisme de $\mathring{B}$ sur $F_t(\mathring{B})$ qui est ouvert.

Il reste à montrer que $F_t(\mathring{B}) = \mathring{B}$. On a déjà clairement $F_t(\mathring{B}) \subset \mathring{B}$ (car $F_t(x) = (1-t)x + t\,r(x)$). On va montrer que $F_t(\mathring{B})$ est à la fois ouvert et fermé dans $\mathring{B}$, ce qui impliquera l'égalité par connexité de $\mathring{B}$. L'ouverture est déjà acquise. Pour la fermeture, prenons (y_n) suite de $F_t(\mathring{B})$ convergeant vers $y \in \mathring{B}$. On écrit $y_n = F_t(x_n)$ avec $x_n \in \mathring{B}$ et on extrait de (x_n) une sous-suite convergeant vers un élément $x \in B$ par compacité de B. On a alors $y = F_t(x)$. Si $x \in S$, alors $F_t(x) = x$, et donc $y = x \in S$, ce qui est absurde. Donc $x \in \mathring{B}$, et $y \in F_t(\mathring{B})$. On a donc bien $F_t(\mathring{B}) = \mathring{B}$.

Montrons enfin l'absurdité. Pour $t = 1$, on a $F_1 = r$, et donc $DF_1(x) = Dr(x)$, qui n'est inversible en aucun point de $\mathring{B}$. En effet, s'il existait $x \in \mathring{B}$ où $Dr(x)$ est inversible, par le théorème d'inversion locale, r induirait un $\mathcal{C}^1$-difféomorphisme d'un ouvert V de $\mathbb{R}^n$ inclu dans $\mathring{B}$ contenant x sur un ouvert W de $\mathbb{R}^n$ contenu dans S (car l'image de r est dans S), ce qui est absurde car $\mathring{S} = \varnothing$. Donc $P(1) = 0$, ce qui contredit $P(1) = K > 0$. □

Version complète :

Théorème. *On note B la boule unité fermée de $\mathbb{R}^n$ pour une norme $\|\cdot\|$ quelconque. Alors toute application continue $f : B \to B$ admet un point fixe.*

Démonstration. On peut supposer que la norme $\|\cdot\|$ est la norme euclidienne usuelle. Soit $f : B \to B$ continue. Raisonnons par l'absurde en supposant que f n'a pas de point fixe.

Montrons d'abord qu'on peut supposer f de classe $\mathcal{C}^1$. Soit $\alpha = \inf_{x\in B} \|f(x)-x\|$. Par compacité de B, cette borne inférieure est atteinte, et comme f n'a pas de point fixe, $\alpha > 0$. Soit g une fonction de classe $\mathcal{C}^1$ telle que

$$\|f - g\|_{\infty,B} = \sup_{x\in B} \|f(x) - g(x)\| < \frac{\alpha}{2},$$

que l'on obtient par le théorème de Stone-Weierstrass en raisonnant sur les composantes de f : $f = (f_1, \dots, f_n)$ et on pose $g = (g_1, \dots, g_n)$ avec g_i des fonctions de classe $\mathcal{C}^1$ telles que $\|f_i - g_i\|_\infty = \sup_{x\in B} |f_i(x) - g_i(x)| < \frac{\alpha}{2\sqrt{n}}$. On a

$$\|g(x)\| \leqslant \|g(x) - f(x)\| + \|f(x)\| \leqslant \|f - g\|_{\infty,B} + \|f(x)\| < \frac{\alpha}{2} + 1,$$

pour tout $x \in B$. On pose alors $h = \frac{1}{1+\frac{\alpha}{2}} g$, qui est $\mathcal{C}^1$ sur B et à valeurs dans B. Enfin, pour tout $x \in B$, on a

$$\begin{aligned} \|h(x) - f(x)\| &\leqslant \|h(x) - g(x)\| + \|g(x) - f(x)\| \\ &= \left(1 - \frac{1}{1+\frac{\alpha}{2}}\right) \|g(x)\| + \|g(x) - f(x)\| \\ &< \alpha. \end{aligned}$$

Donc

$$\|h(x) - x\| \geqslant \|f(x) - x\| - \|h(x) - f(x)\| > 0,$$

i.e. h n'a pas de point fixe. Ainsi, si on montre qu'il est absurde que toute fonction $\mathcal{C}^1$ de B dans B n'a pas de point fixe, il sera aussi absurde que f n'a pas de point fixe.

On suppose alors f de classe $\mathcal{C}^1$. On note S la sphère unité de $\mathbb{R}^n$. Pour $x \in B$, on définit $r(x)$ comme étant le point d'intersection de S avec la demi-droite $[f(x), x\big)$. Montrons que r est de classe $\mathcal{C}^1$ sur B. Pour $x \in B$, il existe $\lambda(x) \in \mathbb{R}_+$ tel que

$$r(x) = f(x) + \lambda(x)(x - f(x)).$$

Comme $r(x) \in S$, $\lambda(x)$ est solution de $\|f(x) + \lambda(x)(x - f(x))\|^2 = 1$, *i.e.* de

$$\|f(x)\|^2 + 2\lambda(x)\big(f(x), x - f(x)\big) + \lambda(x)^2\|x - f(x)\|^2 = 1.$$

Le discriminant de cette équation du second degré en $\lambda(x)$ est clairement positif :

$$\Delta(x) = \big(2\big(f(x), x - f(x)\big)\big)^2 + 4\|x - f(x)\|^2 \underbrace{\big(1 - \|f(x)\|^2\big)}_{\geqslant 0} \geqslant 0.$$

On a même $\Delta(x) > 0$ car sinon il n'y aurait qu'un seul point d'intersection entre la droite passant par x et $f(x)$ et la sphère S, ce qui est absurde. Donc, $\lambda(x)$ étant la racine positive de l'équation, on obtient :

$$\lambda(x) = \frac{-2\big(f(x), x - f(x)\big) + \sqrt{\Delta(x)}}{2\|x - f(x)\|^2},$$

et il est clair que λ est de classe $\mathcal{C}^1$ sur B. Donc r est de classe $\mathcal{C}^1$ sur B.

On pose maintenant $F_t(x) = (1-t)x + t\,r(x)$ pour $t \in [0,1]$. Il est clair que F_t est une application de classe $\mathcal{C}^1$ de B dans B. En notant λ la mesure de Lebesgue sur $\mathbb{R}^n$, on définit pour $t \in [0,1]$,

$$P(t) = \int_{\mathring{B}} \det\left(DF_t(x)\right) \,\mathrm{d}\lambda(x),$$

qui est polynomiale en t (car à x fixé, $t \mapsto F_t(x)$ est polynomiale en t).

Nous allons montrer que P est constant et strictement positif pour t suffisamment petit. Comme P est un polynôme, on aura alors que P est constant non nul et on verra ensuite pourquoi c'est absurde.

Si on montre que, pour t suffisamment petit, F_t est un $\mathcal{C}^1$-difféomorphisme de $\mathring{B}$ sur $\mathring{B}$, de jacobien J_{F_t} partout > 0, alors par le théorème de changement de variable, on aura :

$$\begin{aligned} P(t) &= \int_{\mathring{B}} \det\left(DF_t(x)\right) \,\mathrm{d}\lambda(x) = \int_{\mathring{B}} J_{F_t}(x)\,\mathrm{d}\lambda(x) \\ &= \int_{\mathring{B}} 1\,|J_{F_t}(x)|\,\mathrm{d}\lambda(x) = \int_{F_t(\mathring{B})} 1\,\mathrm{d}\lambda(x) \\ &= \int_{\mathring{B}} 1\,\mathrm{d}\lambda(x) = K, \end{aligned}$$

où $K > 0$ est le volume de $\mathring{B}$, et on aura montré ce qu'on voulait.

Appliquons le théorème d'inversion globale :

(i) Puisque $F_t(x) = (1-t)x + t\,r(x)$, on a

$$J_{F_t}(x) = a_n(x)t^n + \cdots + a_1(x)t + a_0(x)$$

avec $a_0, \ldots, a_n$ des fonctions continues de B dans $\mathbb{R}$. Or $F_0(x) = x$, donc $DF_0(x) = \mathrm{Id}_{\mathbb{R}^n}$ et $J_{F_0}(x) = 1 = a_0(x)$. On pose

$$m = \sup_{\substack{x \in B \\ t \in [0,1]}} \big|a_n(x)t^{n-1} + \cdots + a_1(x)\big|.$$

Alors pour tout $t < \frac{1}{m}$ et tout $x \in B$, il vient :

$$t|a_n(x)t^{n-1} + \cdots + a_1(x)| < 1,$$

d'où

$$J_{F_t}(x) = t\big(a_n(x)t^{n-1} + \cdots + a_1(x)\big) + 1 > 0.$$

(ii) On montre que F_t est injective sur $\mathring{B}$ pour t suffisamment petit. Soient $x, y \in \mathring{B}$ tels que $F_t(x) = F_t(y)$. On a donc :

$$(1-t)(x-y) = t\big(r(y) - r(x)\big).$$

Notons $M = \sup_{x\in B} \|Dr(x)\|$. L'inégalité de la moyenne donne :

$$\|r(x) - r(y)\| \leqslant M\|x-y\|,$$

d'où $(1-t)\|x-y\| \leqslant tM\|x-y\|$. Pour t suffisamment petit, disons $t < \eta$, cela implique $x = y$.

Le théorème d'inversion globale nous dit alors que pour $t < \min\big(\eta, \frac{1}{m}\big)$, F_t est un $\mathcal{C}^1$-difféomorphisme de $\mathring{B}$ sur $F_t(\mathring{B})$ qui est ouvert.

Il reste à montrer que $F_t(\mathring{B}) = \mathring{B}$. On a déjà clairement $F_t(\mathring{B}) \subset \mathring{B}$ (car $F_t(x) = (1-t)x + t\,r(x)$). On va montrer que $F_t(\mathring{B})$ est à la fois ouvert et fermé dans $\mathring{B}$, ce qui impliquera l'égalité par connexité de $\mathring{B}$. L'ouverture est déjà acquise. Pour la fermeture, prenons (y_n) suite de $F_t(\mathring{B})$ convergeant vers $y \in \mathring{B}$. On écrit $y_n = F_t(x_n)$ avec $x_n \in \mathring{B}$ et on extrait de (x_n) une sous-suite convergeant vers un élément $x \in B$ par compacité de B. On a alors $y = F_t(x)$. Si $x \in S$, alors $F_t(x) = x$, et donc $y = x \in S$, ce qui est absurde. Donc $x \in \mathring{B}$, et $y \in F_t(\mathring{B})$. On a donc bien $F_t(\mathring{B}) = \mathring{B}$.

Montrons enfin l'absurdité. Pour $t = 1$, on a $F_1 = r$, et donc $DF_1(x) = Dr(x)$, qui n'est inversible en aucun point de $\mathring{B}$. En effet, s'il existait $x \in \mathring{B}$ où $Dr(x)$ est inversible, par le théorème d'inversion locale, r induirait un $\mathcal{C}^1$-difféomorphisme d'un ouvert V de $\mathbb{R}^n$ inclu dans $\mathring{B}$ contenant x sur un ouvert W de $\mathbb{R}^n$ contenu dans S (car l'image de r est dans S), ce qui est absurde car $\mathring{S} = \varnothing$. Donc $P(1) = 0$, ce qui contredit $P(1) = K > 0$. □

Corollaire. *Soit $\overline{B}(0,R)$ la boule fermée centrée en 0 de rayon R de $\mathbb{R}^n$. Alors toute application $f : \overline{B}(0,R) \to \overline{B}(0,R)$ continue admet un point fixe.*

Démonstration. Pour $\lambda > 0$, soit h_λ l'homothétie de rapport λ :

$$\begin{array}{rccc} h_\lambda : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & x & \longmapsto & \lambda x. \end{array}$$

Alors $h_{\frac{1}{R}} \circ f \circ h_R : \overline{B}(0,1) \to \overline{B}(0,1)$ est continue. Elle admet donc un point fixe x :

$$h_{\frac{1}{R}} \circ f \circ h_R(x) = \frac{1}{R} f(Rx) = x.$$

Donc $Rx \in \overline{B}(0,R)$ est un point fixe de f. □

Corollaire. *Soit C un ensemble convexe et compact de $\mathbb{R}^n$. Alors toute application $f : C \to C$ continue admet un point fixe.*

Démonstration. Soit $f : C \to C$ continue. Soit $\Pi : \mathbb{R}^n \to C$ la projection de $\mathbb{R}^n$ sur C qui existe et est continue d'après le théorème de projection sur un convexe fermé dans un espace de Hilbert. Comme C est compact, il est borné, donc il existe $R > 0$ tel que $C \subset \overline{B}(0,R)$. On a donc une fonction $f \circ \Pi : \overline{B}(0,R) \to \overline{B}(0,R)$ continue. Il existe donc $x \in \overline{B}(0,R)$ tel que $f \circ \Pi(x) = x$. Mais $f \circ \Pi(x) \in C$ car $f : C \to C$. Donc $x \in C$ et $\Pi(x) = x$. Ainsi x est un point fixe de f. □

Théorème (Stone-Weierstrass)**.** *Soit X un espace compact et A une sous-algèbre de l'algèbre $\mathcal{C}(X,\mathbb{R})$ des fonctions continues à valeurs réelles, muni de la norme infini. On suppose que :*

(i) A sépare les points de X, i.e. pour tout $x, y \in X$ avec $x \neq y$, il existe $f \in A$ telle que $f(x) \neq f(y)$.

(ii) A est unitaire, i.e. 1 appartient à A.

Alors A est dense dans $\mathcal{C}(X,\mathbb{R})$.

Références :

– Aucune.

3.22 Théorème de Hardy-Littlewood

Théorème. *Soit (a_n) une suite de complexes. On suppose que :*

(i) $a_n = O\left(\frac{1}{n}\right)$.

(ii) La fonction S définie sur $]-1,1[$ par $S(x) = \sum_{n=0}^{\infty} a_n x^n$ admet une limite l lorsque $x \to 1^-$.

Alors la série $\sum a_n$ converge et sa somme vaut l.

Démonstration. On note que S est bien définie pour $|x| < 1$ car la série entière $\sum a_n z^n$ a un rayon de convergence $\geqslant 1$ grâce à $a_n = O\left(\frac{1}{n}\right)$.

Quitte à considérer la suite (b_n) définie par $b_0 = a_0 - l$ et $b_n = a_n$ pour $n \geqslant 1$, on peut supposer que $l = 0$.

On considère la fonction $f : [0,1] \to \mathbb{C}$ telle que $f(x) = 0$ pour $x < \frac{1}{2}$ et $f(x) = 1$ pour $x \geqslant \frac{1}{2}$. Alors si $x < 1$, on a $x^n < \frac{1}{2}$ dès que $n > N_x = \left\lfloor -\frac{\ln(2)}{\ln(x)} \right\rfloor$. Donc

$$S_{N_x} = \sum_{n=0}^{N_x} a_n = \sum_{n=0}^{\infty} a_n f(x^n).$$

Notre but est donc de montrer que $\lim_{x\to 1^-} \sum_{n=0}^{\infty} a_n f(x^n) = 0$, car $N_x \to \infty$ quand $x \to 1^-$.

On considère l'ensemble E des fonctions $\varphi : [0,1] \to \mathbb{R}$ telles que $\sum a_n \varphi(x^n)$ converge pour $0 \leqslant x < 1$ et $\lim_{x\to 1^-} \sum_{n=0}^{\infty} a_n \varphi(x^n) = 0$. Pour montrer que $f \in E$, on va l'approcher par des polynômes nuls en 0 et valant $f(1) = 1$ en 1 (nécessaire si on veut éviter des problèmes à la limite en 1^-). L'intérêt que les polynômes soient nuls en 0 est qu'ainsi, ils sont bien dans E : si $P = X^k$ est un monôme, $k \geqslant 1$, alors $\sum_{n=0}^{\infty} a_n P(x^n) = S(x^k)$ est bien définie pour $0 \leqslant x < 1$ et converge vers 0 par hypothèse du théorème. Par linéarité, tout polynôme nul en 0 est dans E.

Pour approcher f ainsi, on l'écrit $f(x) = x + x(1-x)g(x)$ pour une bonne fonction g, en réalité définie par $g(x) = \frac{f(x)-x}{x(1-x)}$ si $0 < x < 1$, et $g(0) = -1$, $g(1) = 1$. Il existe deux fonctions continues g_1 et g_2 telles que $g_1 \leqslant g \leqslant g_2$ et $\|g_2 - g_1\|_1 \leqslant \varepsilon$ (on prend g_1 égale à g partout sauf sur un petit voisinage à gauche de la discontinuité et on relie les deux extrémités par une droite). Ensuite, comme g_1 et g_2 sont continues sur le segment $[0,1]$, par le théorème de Stone-Weierstrass, il existe deux polynômes Q_1 et Q_2 tels que $-\varepsilon \leqslant g_i - Q_i \leqslant \varepsilon$ sur $[0,1]$. Alors les polynômes $R_1 = Q_1 - \varepsilon$ et $R_2 = Q_2 + \varepsilon$ vérifient

$$R_1 \leqslant g_1 \leqslant g \leqslant g_2 \leqslant R_2$$

et $R_2 - R_1 = Q_2 - Q_1 + 2\varepsilon \leqslant g_2 - g_1 + 4\varepsilon$. D'où :

$$\int_0^1 (R_2(x) - R_1(x))\,\mathrm{d}x \leqslant \int_0^1 (g_2(x) - g_1(x) + 4\varepsilon)\,\mathrm{d}x = \|g_2 - g_1\|_1 + 4\varepsilon \leqslant 5\varepsilon.$$

On pose enfin : $P_i = X + X(1-X)R_i$. Les polynômes P_1 et P_2 sont nuls en 0, valent 1 en 1 et vérifient $P_1 \leqslant f \leqslant P_2$. Si on note

$$Q(x) = \frac{P_2(x) - P_1(x)}{x(1-x)} = R_2(x) - R_1(x),$$

alors Q est un polynôme et vérifie $\|Q\|_1 \leqslant 5\varepsilon$.

On utilise maintenant l'hypothèse $a_n = O\left(\frac{1}{n}\right)$: il existe $M > 0$ tel que $|a_n| \leqslant \frac{M}{n}$. Pour tout $x \in [0,1[$, on a :

$$\begin{aligned}\left|\sum_{n=0}^{\infty} a_n f(x^n) - \sum_{n=0}^{\infty} a_n P_1(x^n)\right| &\leqslant \sum_{n=0}^{\infty} |a_n|(P_2 - P_1)(x^n) \\ &\leqslant M \sum_{n=1}^{\infty} \frac{x^n(1-x^n)}{n} Q(x^n) \\ &\leqslant M(1-x)\sum_{n=1}^{\infty} x^n Q(x^n),\end{aligned}$$

car $(1-x^n) = (1-x)(1+x+\cdots+x^{n-1}) \leqslant n(1-x)$. D'où :

$$\left|\sum_{n=0}^{\infty} a_n f(x^n)\right| \leqslant \underbrace{\left|\sum_{n=0}^{\infty} a_n P_1(x^n)\right|}_{=A(x)} + M \underbrace{(1-x)\sum_{n=1}^{\infty} x^n Q(x^n)}_{=B(x)},$$

pour tout $x \in [0,1[$. Comme $P_1 \in E$, on a déjà $A(x) \underset{x\to 1^-}{\longrightarrow} 0$. Pour montrer que $B(x) \to 0$, par linéarité, on calcule la limite lorsque Q est un monôme : pour tout $x \in [0,1[$,

$$(1-x)\sum_{n=0}^{\infty} x^n (x^n)^k = \frac{1-x}{1-x^{k+1}} = \frac{1}{1+x+\cdots+x^k} \underset{x\to 1^-}{\longrightarrow} \frac{1}{k+1} = \int_0^1 t^k \,\mathrm{d}t.$$

D'où par linéarité :

$$B(x) \underset{x\to 1^-}{\longrightarrow} \int_0^1 Q(t)\,\mathrm{d}t = \|Q\|_1.$$

Cela signifie que $B(x) = |B(x)| \leqslant \|Q\|_1 + \varepsilon$ pour x assez proche de 1^-, *i.e.* $B(x) \leqslant 6\varepsilon$. Enfin, il existe $\eta < 1$ tel que pour tout $x \in [\eta, 1[$,

$$\left|\sum_{n=0}^{\infty} a_n f(x^n)\right| \leqslant (6M+1)\varepsilon.$$

□

Application. *Soit* $a_n = \frac{(-1)^{n-1}}{n}$. *Alors* $a_n = O\left(\frac{1}{n}\right)$ *et pour* $x \in [0,1[$,

$$S(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n = \ln(1+x).$$

La fonction S *admet donc une limite quand* $x \to 1^-$, *qui vaut* $\ln(2)$. *Le théorème s'applique :* $\sum a_n$ *converge et*

$$\sum_{n=1}^{\infty} a_n = \ln(2).$$

Application. *Soit* $a_n = \frac{(-1)^n}{2n+1}$. *Alors* $a_n = O\left(\frac{1}{n}\right)$ *et pour* $x \in [0,1[$,

$$S(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^n.$$

On sait que $\arctan(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$ *(on peut dériver termes à termes pour le démontrer). Donc* $xS(x^2) = \arctan(x)$ *pour tout* $x \in [0,1[$. *On en déduit :*

$$S(x) = \frac{\arctan(\sqrt{x})}{\sqrt{x}}$$

pour tout $x \in [0,1[$. *La fonction* S *admet donc une limite quand* $x \to 1^-$, *qui vaut* $\arctan(1) = \frac{\pi}{4}$. *Le théorème s'applique :* $\sum a_n$ *converge et*

$$\sum_{n=0}^{\infty} a_n = \frac{\pi}{4}.$$

Références :

- Gourdon - *Analyse* - Page 289.

3.23 Théorème de Le Cam

Théorème. *Soit S_n une variable aléatoire telle que $S_n \overset{loi}{=} \sum_{i=1}^n X_i$ où les X_i sont des v.a.i. de Bernoulli de paramètres $p_i \in]0,1[$. Alors il existe une v.a. $Z \overset{loi}{=} \mathcal{P}(\sum_{i=1}^n p_i)$ telle que*

$$\|P^{S_n} - P^Z\| \leqslant \sum_{i=1}^n p_i^2,$$

où $\|P^{S_n} - P^Z\| = \sup_{B\in\mathcal{B}} |P^{S_n}(B) - P^Z(B)|$.

Démonstration. Soit $B \in \mathcal{B}$. On commence par majorer, pour X et Y deux *v.a.*, $|P^X(B) - P^Y(B)|$. On a :

$$\{X \in B\} = (\{X \in B\} \cap \{X = Y\}) \sqcup (\{X \in B\} \cap \{X \neq Y\}).$$

Donc $P(X \in B) = P(\{X \in B\} \cap \{X = Y\}) + P(\{X \in B\} \cap \{X \neq Y\})$, et on a pareil pour Y. Donc :

$$\begin{aligned} |P^X(B) - P^Y(B)| &= \big|P(\{X \in B\} \cap \{X \neq Y\}) - P(\{Y \in B\} \cap \{X \neq Y\})\big| \\ &\leqslant \max\left(P(\{X \in B\} \cap \{X \neq Y\}), P(\{Y \in B\} \cap \{X \neq Y\})\right) \\ &\leqslant P(X \neq Y), \end{aligned}$$

où pour la première inégalité, on a simplement remarqué que si $0 \leqslant a \leqslant b$, on a $|a - b| = b - a \leqslant b = \max(a, b)$. En passant au sup, il vient :

$$\|P^X - P^Y\| \leqslant P(X \neq Y).$$

On va d'abord prouver le résultat pour $n = 1$. Soient $p \in]0,1[$ et deux *v.a.i.* Y et W telles que $Y \overset{\text{loi}}{=} \mathcal{P}(p)$ et $W \overset{\text{loi}}{=} \text{Ber}(1 - (1-p)e^p)$. Soit $X = 1 - \mathbb{1}_{\{Y=W=0\}}$, *v.a.* à valeurs dans $\{0,1\}$. On cherche la loi de X :

$$P(X = 0) = P(Y = 0, W = 0) = P(Y = 0)P(W = 0) = (1-p)e^p e^{-p} = 1 - p.$$

Donc $X \overset{\text{loi}}{=} \text{Ber}(p)$.

Montrons que $P(X \neq Y) \leqslant p^2$. On décompose :

$$\{X \neq Y\} = \underbrace{(\{Y = 0\} \cap \{X = 1\})}_{(\{Y=0\}\cap\{W=1\})} \sqcup \underbrace{(\{Y = 1\} \cap \{X = 0\})}_{=\varnothing} \sqcup \{Y \geqslant 2\}.$$

Donc

$$\begin{aligned} P(X \neq Y) &= P(Y = 0)P(W = 1) + P(Y \geqslant 2) \\ &= e^{-p}(1 - (1-p)e^p) + (1 - e^{-p} - pe^{-p}) \end{aligned}$$

$$\begin{aligned} &= p(1-e^{-p}) \\ &\leqslant p^2, \end{aligned}$$

car $e^{-p} \geqslant 1-p$. Finalement, si $S_1 \overset{\text{loi}}{=} X$, en posant $Z = Y$, on a le résultat annoncé par le théorème :

$$\|P^{S_1} - P^Z\| = \|P^X - P^Y\| \leqslant P(X \neq Y) \leqslant p^2.$$

On va maintenant étendre ce résultat à n entier $\geqslant 1$. On considère les *v.a.* $Y_i \overset{\text{loi}}{=} \mathcal{P}(p_i)$, $W_i \overset{\text{loi}}{=} \text{Ber}(1-(1-p_i)e^{p_i})$ avec $Y_1, \ldots, Y_n, W_1, \ldots W_n$ indépendantes. Soit $X_i = 1 - \mathbb{1}_{\{Y_i = W_i = 0\}}$. On a donc $X_i \overset{\text{loi}}{=} \text{Ber}(p_i)$ et les X_i sont indépendantes. Enfin, on suppose que $S_n \overset{\text{loi}}{=} \sum_{i=1}^n X_i$ et on pose $Z = \sum_{i=1}^n Y_i$.

Comme les Y_i sont indépendantes, on a $Z \overset{\text{loi}}{=} \mathcal{P}(\sum_{i=1}^n p_i)$. En effet, si X et Y sont deux *v.a.i.* de loi de Poisson de paramètres respectifs $\lambda > 0$ et $\mu > 0$, alors

$$G_X(t) = E\big[t^X\big] = \sum_{k=0}^{\infty} t^k e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} e^{t\lambda} = e^{\lambda(t-1)}.$$

Par indépendance, $G_{X+Y}(t) = G_X(t) G_Y(t) = e^{(\lambda+\mu)(t-1)}$. Donc G_{X+Y} est la fonction génératrice d'une *v.a.* de loi $\mathcal{P}(\lambda+\mu)$. Donc $X + Y \overset{\text{loi}}{=} \mathcal{P}(\lambda+\mu)$.

Enfin, on a clairement $\bigcap_{i=1}^n \{X_i = Y_i\} \subset \{S_n = Z\}$. Donc $\{S_n \neq Z\} \subset \bigcup_{i=1}^n \{X_i \neq Y_i\}$, et ainsi :

$$P(S_n \neq Z) \leqslant P\left(\bigcup_{i=1}^n \{X_i \neq Y_i\}\right) \leqslant \sum_{i=1}^n P(X_i \neq Y_i) \leqslant \sum_{i=1}^n p_i^2.$$

On a donc trouvé la *v.a.* Z qu'on cherchait. □

Application. *Soient $\lambda > 0$ et $n \in \mathbb{N}^*$ avec $n \geqslant \lambda$. On pose $p_i = \frac{\lambda}{n}$ et $Z = \mathcal{P}(\lambda)$. Alors (S_n) converge en loi vers Z. On retrouve ainsi le résultat du théorème central limite poissonien dans ce cas particulier.*

Démonstration. Par le théorème de Le Cam, on a

$$|P^{S_n}(B) - P^Z(B)| \leqslant \|P^{S_n} - P^Z\| \leqslant \sum_{i=1}^n \left(\frac{\lambda}{n}\right)^2 = \frac{\lambda^2}{n},$$

pour tout $B \in \mathcal{B}$. En particulier avec $B = \{k\}$, on obtient :

$$\left| P(S_n = k) - \frac{\lambda^k}{k!} e^{-\lambda} \right| \leqslant \frac{\lambda^2}{n} \underset{n\to\infty}{\longrightarrow} 0,$$

et ceci prouve que (S_n) converge en loi vers Z (voir compléments plus bas). □

Définition. Pour X *v.a.* entière et positive, on appelle fonction génératrice de X la série entière :

$$G_X(t) = E\big[t^X\big] = \sum_{k=0}^{\infty} P(X=k)t^k,$$

dont le rayon de convergence est clairement $\geqslant 1$.

Définition. On appelle fonction de répartition de X, ou de sa loi P^X, et on note F^X, la fonction définie sur $\mathbb{R}$ par

$$F^X(t) = P^X(]-\infty,t]) = P(\{\omega \,/\, X(\omega) \leqslant t\}) = P(X \leqslant t).$$

Propriété. *Une fonction de répartition admet au plus un nombre dénombrable de points de discontinuité.*

Démonstration. Soit D_n l'ensemble des points de discontinuité de F avec un saut d'amplitude plus grande que $\frac{1}{n}$. En notant $F(t^-)$ la limite à gauche de F en t, on a :

$$D_n = \left\{ t \in \mathbb{R} \,/\, F(t) - F(t^-) \geqslant \frac{1}{n} \right\}.$$

Comme F est croissante et $0 \leqslant F \leqslant 1$, on a nécessairement $\mathrm{Card}(D_n) \leqslant n$. L'ensemble des points de discontinuité de F est $\bigcup_{n=1}^{\infty} D_n$ qui est au plus dénombrable. □

Définition. Soient X_n, $n \in \mathbb{N}$, des variables aléatoires réelles, définies sur $(\Omega, \mathcal{A}, P)$. On dit que (X_n) converge en loi vers X si l'une des conditions équivalentes suivantes est vérifiée :

(i) $\lim_{n\to\infty} F^{X_n}(t) = F^X(t)$ en tout point de continuité t de F^X.

(ii) $\lim_{n\to\infty} \int \phi(X_n)\,\mathrm{d}P = \int \phi(X)\,\mathrm{d}P$ pour toute fonction $\phi : \mathbb{R} \to \mathbb{R}$ continue bornée.

(iii) $\lim_{n\to\infty} \varphi^{X_n}(t) = \varphi^X(t)$ pour tout $t \in \mathbb{R}$.

Proposition. *Soient X_n, $n \in \mathbb{N}$, des variables aléatoires à valeurs entières. Alors (X_n) converge en loi vers X si et seulement si $\lim_{n\to\infty} P(X_n = k) = P(X = k)$, pour tout $k \in \mathbb{N}$.*

Démonstration. On suppose que (X_n) converge en loi vers X. Soit $k \in \mathbb{N}$. D'après la propriété précédente, toute fonction de répartition admet un nombre au plus dénombrable de points de discontinuité. Donc il existe s et t avec

$$k-1 < s < k < t < k+1$$

tels que F^X soit continue en s et en t. Alors par définition de la convergence en loi de (X_n) vers X, on a la limite suivante :

$$P(X_n = k) = F^{X_n}(t) - F^{X_n}(s) \underset{n\to\infty}{\longrightarrow} F^X(t) - F^X(s) = P(X = k).$$

Réciproquement, supposons que $\lim_{n\to\infty} P(X_n = k) = P(X = k)$, pour tout $k \in \mathbb{N}$. Soit $t \in [k, k+1[$. On a :

$$F^{X_n}(t) = \sum_{l=0}^{k} P(X_n = l) \underset{n\to\infty}{\longrightarrow} \sum_{l=0}^{k} P(X = l) = F^X(t),$$

donc (X_n) converge en loi vers X. □

Théorème (Théorème limite central poissonien)**.** *Soit S_n une variable aléatoire de loi $B(n, p_n)$. Si $\lim_{n\to\infty} np_n = \lambda > 0$, alors S_n converge en loi vers une variable aléatoire de Poisson de paramètre λ.*

Démonstration. D'après la proposition précédente, il suffit de prouver que pour tout $k \in \mathbb{N}$,

$$\lim_{n\to\infty} P(S_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

On calcule donc la limite de $P(S_n = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k}$, ce qui est facile. □

Références :

– Carrieu - *Probabilités, exercices corrigés* - Page 78.

3.24 Théorème de Liapounov

Lemme. *Soit $A \in \mathcal{M}_n(\mathbb{R})$ dont on note $\lambda_1, \dots, \lambda_r$ les valeurs propres distinctes. Alors il existe $C > 0$ tel que pour tout $t \in \mathbb{R}$ et tout $x \in \mathbb{R}^n$,*

$$\|e^{tA}x\| \leqslant C(1+|t|)^{n-1}\left(\sum_{i=1}^{r} e^{t\mathrm{Re}(\lambda_i)}\right)\|x\|.$$

En particulier, si $\mathrm{Re}(\lambda_i) < 0$ pour tout i et si a est tel que $0 < a < \min(-\mathrm{Re}(\lambda_i))$, alors $\|e^{tA}x\| \leqslant Ke^{-at}\|x\|$ pour tout $t \geqslant 0$.

Démonstration. On écrit $x = x_1 + \cdots + x_r$ avec $x_i \in E_i = \mathrm{Ker}((A-\lambda_i I_n)^{m_i})$ (sous-espace caractéristique de A associé à la valeur propre λ_i de multiplicité m_i). Chaque E_i est stable par A et

$$e^{tA}x_i = e^{t\lambda_i}e^{t(A-\lambda_i I_n)}x_i = e^{t\lambda_i}\left(\sum_{k=0}^{m_i-1}\frac{t^k}{k!}(A-\lambda_i I_n)^k\right)x_i.$$

On majore :

$$\left\|\sum_{k=0}^{m_i-1}\frac{t^k}{k!}(A-\lambda_i I_n)^k\right\| \leqslant C_i\left(\sum_{k=0}^{m_i-1}\frac{|t|^k}{k!}\right) \leqslant C_i(1+|t|)^{m_i-1} \leqslant C_i(1+|t|)^{n-1},$$

car $\frac{|t|^k}{k!} \leqslant |t|^k \leqslant \binom{m_i-1}{k}|t|^k$. Puis :

$$\begin{aligned}\|e^{tA}x\| &\leqslant \sum_{i=1}^{r}\|e^{tA}x_i\| \\ &\leqslant \max(C_i)(1+|t|)^{n-1}\left(\sum_{i=1}^{r}e^{t\mathrm{Re}(\lambda_i)}\right)\max(\|x_i\|).\end{aligned}$$

Or $N(x) = \max(\|x_i\|)$ est une norme et toutes les normes sur $\mathbb{R}^n$ sont équivalentes, donc il existe C' tel que $\max(\|x_i\|) \leqslant C'\|x\|$, ce qui donne le résultat.

Enfin, si $\mathrm{Re}(\lambda_i) < 0$ pour tout i et si a est tel que $0 < a < \min(-\mathrm{Re}(\lambda_i))$, on a

$$\frac{\|e^{tA}x\|}{e^{-at}\|x\|} \leqslant C'\max(C_i)(1+|t|)^{n-1}\left(\sum_{i=1}^{r}e^{t(\mathrm{Re}(\lambda_i)+a)}\right),$$

et ce majorant est borné pour tout $t \geqslant 0$ car il tend vers 0 quand $t \to \infty$, d'où la deuxième assertion. □

Théorème. *On considère le système différentiel*

$$y' = f(y),\ \ y(0) = x_0$$

avec $f : \mathbb{R}^n \to \mathbb{R}^n$ de classe $\mathcal{C}^1$ telle que $f(0) = 0$. On suppose que la matrice $A = Df(0)$ a toutes ses valeurs propres de partie réelle strictement négative. Alors pour x_0 assez proche de 0, la solution $y(t)$ tend exponentiellement vers 0 lorsque t tend vers l'infini.

Démonstration. On pose, pour $x, y \in \mathbb{R}^n$,

$$b(x,y) = \int_0^\infty (e^{tA}x, e^{tA}y)\,\mathrm{d}t.$$

Cette intégrale est absolument convergente par l'inégalité de Cauchy-Schwarz et le lemme précédent. On définit ainsi une forme bilinéaire symétrique. On pose ensuite $q(x) = b(x,x)$ qui définit une forme quadratique définie positive : $q(x) = \int_0^\infty \|e^{tA}x\|^2\,\mathrm{d}t \geqslant 0$ et si $q(x) = 0$, alors $e^{tA}x = 0$ pour tout t (par continuité), donc $x = 0$ en prenant $t = 0$.

On note $\|\cdot\|_q = \sqrt{q}$ la norme associée à q, qui est équivalente à $\|\cdot\|$ sur $\mathbb{R}^n$ de dimension finie. On va étudier le comportement de $\|y\|_q$. Par la formule de Taylor-Young, on a $f(y) = f(0) + Df(0).y + o(\|y\|) = Ay + o(\|y\|)$. Pour la suite, on note $r(y) = f(y) - Ay = o(\|y\|)$.

On a $q(y)' = 2b(y,y') = 2b(y,Ay) + 2b(y,r(y))$. Or pour $x \in \mathbb{R}^n$,

$$2b(x,Ax) = \int_0^\infty 2(e^{tA}x, e^{tA}Ax)\,\mathrm{d}t = \left[(e^{tA}x, e^{tA}x)\right]_0^\infty = -\|x\|^2,$$

encore d'après le lemme pour la limite du crochet à l'infini. On va maintenant majorer $b(y,r(y))$:

$$|b(y,r(y))| \leqslant \|y\|_q \|r(y)\|_q$$

par l'inégalité de Cauchy-Schwarz appliquée à la norme $\|\cdot\|_q$. Mais $\frac{\|r(y)\|_q}{\|y\|_q} \to 0$ quand $y \to 0$ par définition du reste. Soit $\varepsilon > 0$: il existe $\alpha > 0$ tel que $q(y) = \|y\|_q^2 \leqslant \alpha$ implique $\|r(y)\|_q \leqslant \varepsilon \|y\|_q$. Donc

$$|2b(y,r(y))| \leqslant 2\varepsilon \|y\|_q^2.$$

Par équivalence des normes, il existe $C > 0$ tel que $C\|y\|_q^2 \leqslant \|y\|^2$, d'où finalement :

$$q(y)' = -\|y\|^2 + 2b(y,r(y)) \leqslant -C\|y\|_q^2 + 2\varepsilon\|y\|_q^2 = -\beta q(y)$$

pour $q(y) \leqslant \alpha$, avec $\beta = C - 2\varepsilon > 0$ si on choisit $\varepsilon < \frac{C}{2}$.

On suppose maintenant que la condition initiale x_0 vérifie $q(x_0) = \|x_0\|_q^2 < \alpha$. Alors $q(y(t)) = \|y(t)\|_q^2 < \alpha$ pour tout $t > 0$. En effet, sinon il existerait un plus petit temps $t_0 > 0$ tel que $q(y(t_0)) \geqslant \alpha$. Par continuité de $t \mapsto q(y(t))$, on a en fait $q(y(t_0)) = \alpha$, et alors $q(y)'(t_0) \leqslant -\beta q(y(t_0)) < 0$. Donc $q(y(t))$ décroît et est strictement plus grand que α pour t légèrement inférieur à t_0, ce qui contredit la définition de t_0. Donc pour tout $t \geqslant 0$, $q(y(t)) < \alpha$, ce qui implique

$$q(y)'(t) \leqslant -\beta q(y(t)),$$

pour tout $t \geqslant 0$. On intègre cette inéquation différentielle :

$$\left(e^{\beta t}q(y)\right)' = e^{\beta t}(q(y)' + \beta q(y)) \leqslant 0,$$

et donc $t \mapsto e^{\beta t} q(y(t))$ est décroissante pour $t \geqslant 0$. Ainsi

$$q(y(t)) \leqslant e^{-\beta y} q(y(0)) = e^{-\beta t} q(x_0),$$

pour tout $t \geqslant 0$. Ceci prouve la décroissance exponentielle de $y(t)$ vers 0. □

Compléments :

On a ainsi prouvé que les solutions de $y' = f(y)$ se comportent de la même façon que celles de $z' = Az$ au voisinage de l'équilibre 0. L'origine est un point d'équilibre attractif, propriété qui se transmet du système différentiel linéaire au système perturbé par le petit terme correctif $r(y)$. La solution du système linéarisé est en effet $z(t) = e^{tA}x$ qui converge exponentiellement vers 0 d'après le lemme.

On n'a pas parlé de l'existence d'une solution y définie pour tout $t \geqslant 0$. Pour $x \in \mathbb{R}^n$, on note y l'unique solution maximale de $y' = f(y)$ telle que $y(0) = x$, dont l'existence et l'unicité sont données par le théorème de Cauchy-Lipschitz.

On a ensuite prouvé à la fin de la démonstration que lorsque $\|x\|_q < \alpha$, $y(t)$ reste dans la boule $B_q(0, \alpha)$. Cela empêche donc y d'exploser en temps fini, et donc y solution maximale est nécessairement définie pour tout $t > 0$. En effet, supposons par l'absurde que y soit définie sur $I =]b, c[$ intervalle ouvert maximal contenant 0 avec c fini. Alors comme f est continue, $y' = f(y)$ est bornée sur I, par un réel $M > 0$. Par l'inégalité de la moyenne, on a donc :

$$\|y(t_1) - y(t_2)\| \leqslant M|t_1 - t_2|$$

pour tout $t_1, t_2 \in I$. Par le critère de Cauchy, on en déduit que $y(t)$ admet une limite finie l lorsque t tend vers c à gauche. De plus, $y'(t) = f(y(t)) \to f(l)$ quand $t \to c$. Par le théorème de prolongement de la dérivée, cela prouve que $y'(c)$ existe et vaut $f(l)$. Si on pose

$$\begin{array}{rccl} \varphi : &]b, c] & \longrightarrow & \mathbb{R}^n \\ & t & \longmapsto & \begin{cases} y(t) \text{ si } t \neq c \\ l \text{ si } t = c, \end{cases} \end{array}$$

alors φ est solution du système différentiel sur $]b, c]$, ce qui contredit la maximalité de y.

On peut prouver de la même façon le théorème suivant qui montre le phénomène d'explosion en temps fini :

Théorème. *Soit l'équation différentielle $y' = F(t, y)$ avec $F : I \times \mathbb{R}^n \to \mathbb{R}^n$ continue, I étant un intervalle ouvert de $\mathbb{R}$. On suppose F localement lipschitzienne par rapport à sa deuxième variable. Soit y une solution maximale définie sur $J =]\alpha, \beta[$. Alors :*

(i) *Si* $\alpha \in I$, *alors* $\limsup_{t \to \alpha} \|y(t)\| = \infty$.

(ii) *Si* $\beta \in I$, *alors* $\limsup_{t \to \beta} \|y(t)\| = \infty$.

Enfin, on peut remarquer qu'on a eu seulement besoin que f soit localement lipschitzienne (pour le théorème de Cauchy-Lipschitz et avoir la continuité) au voisinage de 0 et différentiable en 0, et pas que f soit $\mathcal{C}^1$, cette hypothèse impliquant bien sûr les précédentes.

Références :

– Rouvière - *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation* - Page 138.

3.25 Théorème des extrema liés

Théorème. *Soient $f, g_1, \dots, g_r : U \to \mathbb{R}$ des fonctions de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$. On pose :*

$$\Gamma = \{x \in U \,/\, g_1(x) = \cdots = g_r(x) = 0\}.$$

Si $f_{|\Gamma}$ admet un extremum local en $a \in \Gamma$ et si les formes linéaires $Dg_1(a), \dots, Dg_r(a)$ sont linéairement indépendantes, alors il existe des réels $\lambda_1, \dots, \lambda_r$ (appelés multiplicateurs de Lagrange) tels que

$$Df(a) = \lambda_1 Dg_1(a) + \cdots + \lambda_r Dg_r(a).$$

Remarque. Les λ_i sont uniques car la famille $(Dg_i(a))$ est libre.

Démonstration. D'abord, on remarque que nécessairement, $r \leqslant n$ car la famille $(Dg_i(a))$ est libre dans $(\mathbb{R}^n)^*$ de dimension n. Ensuite, si $r = n$, le théorème est évident car la famille $(Dg_i(a))$ devient une base de $(\mathbb{R}^n)^*$ et donc $Df(a)$ s'exprime sur cette base. On suppose donc $r \leqslant n-1$.

Soit $s = n - r \geqslant 1$. On identifie $\mathbb{R}^n$ à $\mathbb{R}^s \times \mathbb{R}^r$ et on écrit les éléments de $\mathbb{R}^n$ sous la forme $(x, y) = (x_1, \dots, x_s, y_1, \dots, y_r)$. On pose $a = (\alpha, \beta) \in \mathbb{R}^s \times \mathbb{R}^r$.

Comme la famille $(Dg_i(a))$ est libre, la matrice

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \dots & \frac{\partial g_1}{\partial x_s}(a) & \frac{\partial g_1}{\partial y_1}(a) & \dots & \frac{\partial g_1}{\partial y_r}(a) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1}(a) & \dots & \frac{\partial g_r}{\partial x_s}(a) & \frac{\partial g_r}{\partial y_1}(a) & \dots & \frac{\partial g_r}{\partial y_r}(a) \end{pmatrix}$$

est de rang r. On peut alors en extraire une sous-matrice de taille $r * r$ inversible, et quitte à changer le nom des variables, on peut supposer que $\det\left(\frac{\partial g_i}{\partial y_j}(a)\right)_{1 \leqslant i,j \leqslant r} \neq 0$. D'après le théorème des fonctions implicites, en notant $g = (g_1, \dots, g_r)$, on peut donc trouver un voisinage ouvert U' de α dans $\mathbb{R}^s$, un voisinage ouvert U'' de β dans $\mathbb{R}^r$ et une fonction $\varphi = (\varphi_1, \dots, \varphi_r) : U' \to U''$ de classe $\mathcal{C}^1$ tels que

$$x \in U',\ y \in U'' \text{ et } g(x, y) = 0 \iff x \in U' \text{ et } y = \varphi(x).$$

Posons $h(x) = f(x, \varphi(x))$. La fonction h admet un extremum local en $x = \alpha$ car $f(\alpha, \varphi(\alpha)) = f(\alpha, \beta) = f(a)$ et $(x, \varphi(x)) \in \Gamma$ pour tout $x \in U'$. Par conséquent, $Dh(\alpha) = 0$, *i.e.* pour tout $i \in [\![1, s]\!]$,

$$\frac{\partial h}{\partial x_i}(\alpha) = \frac{\partial f}{\partial x_i}(a) + \sum_{j=1}^{r} \frac{\partial \varphi_j}{\partial x_i}(\alpha) \frac{\partial f}{\partial y_j}(a) = 0.$$

D'autre part, pour tout $k \in [\![1, r]\!]$, $g_k(x, \varphi(x)) = 0$, donc en dérivant par rapport à x_i, pour tout $i \in [\![1, s]\!]$, on obtient

$$\frac{\partial g_k}{\partial x_i}(a) + \sum_{j=1}^{r} \frac{\partial \varphi_j}{\partial x_i}(\alpha) \frac{\partial g_k}{\partial y_j}(a) = 0.$$

Ces deux dernières égalités impliquent que les s premiers vecteurs colonnes de la matrice

$$M = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \dots & \frac{\partial f}{\partial x_s}(a) & \frac{\partial f}{\partial y_1}(a) & \dots & \frac{\partial f}{\partial y_r}(a) \\ \frac{\partial g_1}{\partial x_1}(a) & \dots & \frac{\partial g_1}{\partial x_s}(a) & \frac{\partial g_1}{\partial y_1}(a) & \dots & \frac{\partial g_1}{\partial y_r}(a) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1}(a) & \dots & \frac{\partial g_r}{\partial x_s}(a) & \frac{\partial g_r}{\partial y_1}(a) & \dots & \frac{\partial g_r}{\partial y_r}(a) \end{pmatrix}$$

s'expriment linéairement en fonction des r derniers. Donc $\mathrm{rg}(M) \leqslant r$. Or le rang des vecteurs lignes est égal au rang des vecteurs colonnes, donc les $r+1$ vecteurs lignes de M forment une famille liée. Il existe donc $\mu_0, \dots, \mu_r$ non tous nuls tels que $\mu_0 Df(a) + \mu_1 Dg_1(a) + \dots + \mu_r Dg_r(a) = 0$. Comme la famille $(Dg_i(a))$ est libre, on a nécessairement $\mu_0 \neq 0$, d'où le résultat en divisant l'égalité par μ_0. □

Une application importante du théorème des extrema liés est le théorème suivant concernant la diagonalisabilité des endomorphismes symétriques :

Théorème. *Soit u un endomorphisme d'un espace euclidien E. Alors E possède une base orthonormée de vecteurs propres pour u.*

Démonstration. Considérons la fonction

$$\begin{array}{rccc} f : & E & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & (u(x), x), \end{array}$$

qui est de classe $\mathcal{C}^\infty$. Comme E est de dimension finie (car euclidien), la sphère unité de E est compacte et donc f atteint son minimum sur cette sphère en un point v.

D'autre part, on peut considérer le problème de minimisation suivant : on cherche à minimiser f sous la contrainte $g(x) = \|x\|^2 - 1 = 0$. Le théorème des extrema liés nous dit alors qu'il existe $\lambda \in \mathbb{R}$ tel que $Df(v) = \lambda Dg(v)$. Or comme u est symétrique, on a $(u(x), h) = (x, u(h))$, et on obtient en développant $f(x+h) = (u(x+h), x+h)$:

$$Df(x).h = 2(u(x), h).$$

Enfin, comme $Dg(x).h = 2(x, h)$, on obtient

$$(u(v), h) = \lambda(v, h),$$

pour tout $h \in E$. On en déduit que $u(v) = \lambda v$, *i.e.* que v est un vecteur propre de u.

On pose maintenant $e_1 = v$ et $E' = \mathrm{Vect}(e_1)^\perp$. Si $x \in E'$, alors

$$(u(x), e_1) = (x, u(e_1)) = \lambda(x, e_1) = 0,$$

i.e. E' est stable par u. On peut donc conclure par récurrence : E possède une base orthonormée de vecteurs propres pour u.

Remarque : si μ est une autre valeur propre de u et qu'on note w un vecteur propre unitaire associé, alors

$$f(w) = \mu\|w\|^2 = \mu \geqslant \inf_{\|x\|=1} f(x) = \lambda,$$

et donc λ est la plus petite valeur propre de u. □

Donnons deux autres applications du théorème des extrema liés :

Application (Inégalité arithmético-géométrique)**.** *Soit $n \in \mathbb{N}$, $n \geqslant 2$. Alors pour tout $(x_1, \dots, x_n) \in (\mathbb{R}_+)^n$,*

$$\left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}} \leqslant \frac{1}{n}\sum_{i=1}^n x_i.$$

Démonstration. Soient $s > 0$, $f : \mathbb{R}^n \to \mathbb{R}$, $(x_1, \dots, x_n) \mapsto \prod_{i=1}^n x_i$, $g : \mathbb{R}^n \to \mathbb{R}$, $(x_1, \dots, x_n) \mapsto \sum_{i=1}^n x_i - s$ et

$$\Gamma = \{(x_1, \dots, x_n) \in \mathbb{R}^n \,/\, g(x) = 0\}.$$

La fonction f est continue sur le compact $K = \Gamma \cap (\mathbb{R}_+)^n$ (compact car fermé borné) : soit $a \in K$ le maximum global de f sur K. On a en réalité $a \in U = \Gamma \cap (\mathbb{R}_+^*)^n$ car $f(x) = 0$ si l'un des x_i est nul.

Donc $f_{|U}$ admet un extremum global en a, et comme U est un ouvert de Γ, $f_{|\Gamma}$ atteint un extremum local en a. Clairement, $Dg(a) \neq 0$ (c'est même vrai pour tout $x \in \mathbb{R}^n$), donc on peut appliquer le théorème des extrema liés : il existe $\lambda \in \mathbb{R}$ tel que $Df(a) = \lambda Dg(a)$, et donc pour tout $i \in [\![1, n]\!]$,

$$\frac{\partial f}{\partial x_i}(a) = \frac{f(a)}{a_i} = \lambda \frac{\partial g}{\partial x_i}(a) = \lambda.$$

Or $f(a) \neq 0$, donc tous les a_i sont égaux. Comme $\sum_{i=1}^n a_i = s$ (car $a \in \Gamma$), on a $a_i = \frac{s}{n}$. Finalement,

$$f(x) = \prod_{i=1}^n x_i \leqslant f(a) = \left(\frac{s}{n}\right)^n,$$

pour tout $x \in K$. Mais pour $x \in K$, on a $s = \sum_{i=1}^n x_i$, donc on obtient

$$\prod_{i=1}^n x_i \leqslant \left(\frac{\sum_{i=1}^n x_i}{n}\right)^n,$$

et cette inégalité est vraie pour tout $x \in (\mathbb{R}_+)^n$ car elle est vraie pour n'importe quel $s > 0$. □

Application. *Soit $n \in \mathbb{N}^*$. On munit $\mathcal{M}_n(\mathbb{R})$ de la norme $\|M\| = \left(\sum_{i,j} m_{ij}^2\right)^{\frac{1}{2}}$. Alors l'ensemble des éléments de $\mathrm{SL}_n(\mathbb{R}) = \{M \in \mathcal{M}_n(\mathbb{R}), \det(M) = 1\}$ de norme minimale est $\mathrm{SO}_n(\mathbb{R}) = \{M \in \mathcal{O}_n(\mathbb{R}), \det(M) = 1\}$.*

Démonstration. On remarque que $\|M\|^2 = \mathrm{tr}({}^tMM)$. Il s'agit donc de minimiser $f(M) = \mathrm{tr}({}^tMM)$ sous la contrainte $g(M) = 0$ avec $g(M) = \det(M) - 1$.

On vérifie les hypothèses du théorème des extrema liés : f et g sont $\mathcal{C}^1$ sur $\mathcal{M}_n(\mathbb{R})$ (car f est une forme quadratique, et g est fonction du déterminant polynomial en les coefficients de la matrice). L'ensemble $\mathrm{SL}_n(\mathbb{R}) = g^{-1}(\{0\})$ est un fermé de $\mathcal{M}_n(\mathbb{R})$, donc le minimum de f sur $\mathrm{SL}_n(\mathbb{R})$ est atteint en un point $A \in \mathrm{SL}_n(\mathbb{R})$. En effet, dans tout espace vectoriel de dimension finie, la distance d de 0 à un fermé F est toujours atteinte puisqu'elle est atteinte sur le compact K des éléments x de F tels que $\|x\| \leqslant d + 1$. La différentielle de g est celle du déterminant, donc $Dg(A).H = \mathrm{tr}({}^t\mathrm{com}(A)H)$. Ainsi $Dg(A) \neq 0$. Le théorème s'applique donc : il existe $\lambda \in \mathbb{R}$ tel que $Df(A) = \lambda Dg(A)$.

En développant $f(A + H)$, on trouve $Df(A).H = 2\mathrm{tr}({}^tAH)$. Donc

$$2\mathrm{tr}({}^tAH) = \lambda \mathrm{tr}({}^t\mathrm{com}(A)H),$$

pour tout $H \in \mathcal{M}_n(\mathbb{R})$. En prenant $H = E_{ij}$, on trouve que $2A = \lambda \mathrm{com}(A)$. Comme A est inversible et $\det(A) = 1$, on a $\mathrm{com}(A) = {}^tA^{-1}$ (avec l'identité ${}^t\mathrm{com}(A)A = \det(A)I_n$). Donc $2^tAA = \lambda I_n$. En prenant le déterminant, il vient $2^n = \lambda^n$, et comme ${}^tAA$ est une matrice symétrique définie positive, on a $\lambda > 0$, donc $\lambda = 2$. D'où : ${}^tAA = I_n$, *i.e.* $A \in \mathrm{SO}_n(\mathbb{R})$.

On a donc démontré que le minimum de f sur $\mathrm{SL}_n(\mathbb{R})$ était nécessairement atteint en un point $A \in \mathrm{SO}_n(\mathbb{R})$, et la valeur de cette distance minimale est égale à $\mathrm{tr}({}^tAA) = \mathrm{tr}(I_n) = n$. Réciproquement, tous les éléments $M \in \mathrm{SO}_n(\mathbb{R})$ vérifient $f(M) = n$, d'où le résultat. □

Théorème (Théorème des fonctions implicites)**.** *Soient U un ouvert de $\mathbb{R}^n \times \mathbb{R}^p$, (a, b) un point de U et $f : (x, y) \mapsto f(x, y)$ une application de classe $\mathcal{C}^1$ de U dans*

$\mathbb{R}^p$. *On suppose que* $f(a,b)=0$ *et que la matrice jacobienne* $D_yf(a,b)$*, formée des dérivées partielles par rapport à* y*, est inversible, i.e.* $\det(D_yf(a,b))\neq 0$.

Alors l'équation $f(x,y)=0$ *peut être résolue localement par rapport aux variables* y *: il existe* V *voisinage ouvert de* a *dans* $\mathbb{R}^n$*,* W *voisinage ouvert de* b *dans* $\mathbb{R}^p$*, avec* $V\times W\subset U$ *et* $D_yf(x,y)$ *inversible pour tout* $(x,y)\in V\times W$*, et une unique application* $\varphi: V\to W$ *telle que*

$$\big[\,x\in V,\, y\in W \text{ et } f(x,y)=0\,\big] \iff \big[\,x\in V \text{ et } y=\varphi(x)\,\big].$$

De plus, φ *est de classe* $\mathcal{C}^1$ *sur* V.

Exemple. *Dans* $\mathbb{R}^2$*, on cherche le ou les points les plus proches de l'origine sur une courbe* C*. Si la courbe* C *est définie paramétriquement par* $x=x(t)$ *et* $y=y(t)$ *où* t *parcourt un intervalle* I*, il s'agit de déterminer le minimum sur* I *de la fonction* $f: t\mapsto x(t)^2+y(t)^2$ *(distance au carré de* $(x(t),y(t))$ *à l'origine). C'est un problème d'extrema libres car* $x(t)$ *et* $y(t)$ *sont indépendants, et on le résout en sachant que les extrema* t *sont des points critiques : nécessairement* $f'(t)=0$.

Si la courbe C *est définie implicitement par l'équation* $g(x,y)=0$*, il s'agit de déterminer le minimum de* $f:(x,y)\mapsto x^2+y^2$ *lorsque* x *et* y *sont liés par cette relation* $g(x,y)=0$*. Le théorème des extrema liés nous dit que* $Df(a,b)=\lambda Dg(a,b)$*. On résout alors le système de 3 équations en les 3 inconnues* a,b,λ *suivant :*

$$\begin{cases} \frac{\partial f}{\partial x}(a,b)=\lambda\frac{\partial g}{\partial x}(a,b)\\ \frac{\partial f}{\partial y}(a,b)=\lambda\frac{\partial g}{\partial y}(a,b)\\ g(a,b)=0. \end{cases}$$

Remarque. Si on suppose de plus $Df(a)\neq 0$, l'hypersurface S de niveau de f passant par a, d'équation $f(x)=f(a)$, est lisse en a (théorème 5.11 dans Rouvière page 200). La condition nécessaire d'extrema liés signifie alors que son espace vectoriel tangent en a, qui est le noyau de $Df(a)$, contient l'espace vectoriel tangent en Γ (*i.e.* tous les vecteurs v tels que $Dg_i(a).v=0$ pour tout i). Donc Γ est tangent en a à S.

En reprenant l'exemple précédent, si on trace des cercles de centre O et de rayons croissants jusqu'à rencontrer la courbe C, on voit sur un dessin que le ou les points de C qui minimisent la distance à l'origine sont tels que C est tangente à un de ces cercles. Les cercles sont en fait des niveaux de $f(x,y)=x^2+y^2$, ceci est donc justifié par ce que l'on vient de dire.

Références :

- Gourdon - *Analyse* - Page 327 (le théorème), page 319 (la première application), page 321 (la deuxième application).
- Benzoni-Gavage - *Calcul différentiel et équations différentielles* - Exercice 3.1 page 87 (l'application à la diagonalisabilité des endormorphismes symétriques).
- Rouvière - *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation* - Page 192 (théorème des fonctions implicites) et page 372 (suppléments sur le théorème des extrema liés).

3.26 Théorème des quatre sommets

Théorème. *Soit $f : \mathbb{R} \to \mathbb{R}^2$ un arc de classe $\mathcal{C}^3$ paramétré par l'abscisse curviligne, τ-périodique, correspondant à une courbe C fermée, simple et convexe (i.e. qui délimite une partie convexe). Alors la courbure γ de C admet au moins quatre points critiques sur chaque période.*

Démonstration. Comme f est de classe $\mathcal{C}^3$, sa tangente unitaire T est de classe $\mathcal{C}^2$, et sa normale unitaire N et sa courbure γ sont de classe $\mathcal{C}^1$. Ces fonctions sont également τ-périodiques et on a

$$\int_0^\tau \gamma'(s)\,\mathrm{d}s = \gamma(\tau) - \gamma(0) = 0.$$

Par intégration par parties, on a de plus

$$\begin{aligned}\int_0^\tau \gamma'(s)f(s)\,\mathrm{d}s &= \underbrace{\big[\gamma(s)f(s)\big]_0^\tau}_{=0} - \int_0^\tau \gamma(s)T(s)\,\mathrm{d}s \\ &= \int_0^\tau N'(s)\,\mathrm{d}s = N(\tau) - N(0) = 0,\end{aligned}$$

où l'on s'est servi de la deuxième formule de Frénet : $N'(s) = -\gamma(s)T(s)$. En écrivant $f(s) = (x(s), y(s))$, on a donc

$$\int_0^\tau \gamma'(s)\,\mathrm{d}s = \int_0^\tau x(s)\gamma'(s)\,\mathrm{d}s = \int_0^\tau y(s)\gamma'(s)\,\mathrm{d}s = 0. \quad (*)$$

Comme γ est continue, elle admet au moins un maximum et un minimum sur chaque période. Ces extrema sont des points critiques, notons-les s_1 et s_2 avec $s_1 < s_2 < s_1 + \tau$.

Si γ est constante, le résultat est immédiat. On suppose donc que γ' n'est pas identiquement nulle, et étant d'intégrale nulle sur une période (par $(*)$), γ' ne peut être de signe constant sur $[s_1, s_1 + \tau]$.

On suppose que s_1 et s_2 sont les seuls points critiques de γ. Comme C délimite un convexe, la droite $(f(s_1)f(s_2))$ sépare C en exactement deux arcs, et par continuité de γ', γ' ne s'annule pas et reste de signe constant sur $[s_1, s_2]$ et sur $[s_2, s_1 + \tau]$. Soit $ax + by + c = 0$ une équation de la droite $(f(s_1)f(s_2))$. Alors

$$g : s \mapsto (ax(s) + by(s) + c)\gamma'(s)$$

doit être de signe constant sur $[s_1, s_1 + \tau]$. Or par $(*)$, $\int_0^\tau g(s)\,\mathrm{d}s = 0$, donc comme g est continue, cela implique $g = 0$, ce qui est absurde car alors toute la courbe C serait incluse dans la droite $(f(s_1)f(s_2))$ et la courbure γ serait constante.

Donc γ admet un troisième point critique s_3 dans $[s_1, s_1+\tau[$ et on peut supposer pour se fixer les idées que $s_1 < s_3 < s_2 < s_1 + \tau$. Supposons à nouveau que ce sont les seuls points critiques de γ. Alors γ' reste de signe constant sur $[s_1, s_3]$, $[s_3, s_2]$ et sur $[s_2, s_1+\tau]$. Il y a donc nécessairement deux arcs adjacents où γ' reste de même signe. S'il s'agit des arcs $[s_1, s_2]$ et $[s_2, s_1+\tau]$, c'est absurde d'après le raisonnement précédent. Les autres cas de figure se traitent de la même façon.

Finalement, γ admet au moins quatre points critiques sur $[s_1, s_1+\tau[$. □

Exemple. *D'après le théorème des quatre sommets, une ellipse admet au moins quatre sommets. Vérifions-le.*

On paramètre l'ellipse par $f : t \mapsto (a\cos(t), b\sin(t))$, $t \in [0, 2\pi]$, $a, b > 0$. *On a*

$$f'(t) = (-a\sin(t), b\cos(t)) \quad et \quad f''(t) = (-a\cos(t), -b\sin(t)),$$

d'où $\det(f'(t), f''(t)) = ab$. *De plus,*

$$\|f'(t)\| = \sqrt{a^2\sin^2(t) + b^2\cos^2(t)},$$

et donc la courbure vaut

$$\gamma(t) = \frac{\det(f'(t), f''(t))}{\|f'(t)\|^3} = \frac{ab}{(a^2\sin^2(t) + b^2\cos^2(t))^{\frac{3}{2}}}.$$

En la dérivant, on trouve

$$\gamma'(t) = \frac{3ab(b^2 - a^2)\cos(t)\sin(t)}{(a^2\sin^2(t) + b^2\cos^2(t))^{\frac{5}{2}}}.$$

Il y a donc en réalité quatre annulations à γ' *: en* $t = 0$, $\frac{\pi}{2}$, π, *et* $\frac{3\pi}{2}$.

On peut également déterminer la développée de l'ellipse, *i.e.* le lieu de ses centres de courbure. Le vecteur tangent $T(t)$ vaut

$$T(t) = \frac{f'(t)}{\|f'(t)\|} = \frac{1}{\sqrt{a^2\sin^2(t) + b^2\cos^2(t)}}(-a\sin(t), b\cos(t)),$$

et le vecteur normal correspondant est donc

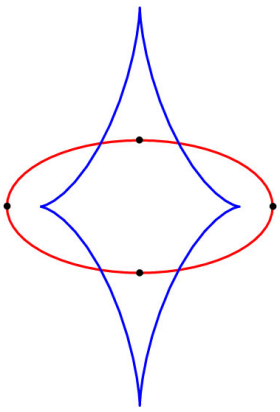
$$N(t) = \frac{1}{\sqrt{a^2\sin^2(t) + b^2\cos^2(t)}}(-b\cos(t), -a\sin(t)),$$

où pour éviter les calculs, on peut prendre un vecteur $N(t)$ perpendiculaire à $T(t)$ tel que $(T(t), N(t))$ soit direct, puis le normaliser. Les coordonnées des centres de courbure $C(t)$ sont donc :

$$C(t) \;=\; f(t) + \frac{1}{\gamma(t)}N(t)$$

$$\begin{aligned} &= (a\cos(t), b\sin(t)) + \frac{a^2\sin^2(t) + b^2\cos^2(t)}{ab}(-b\cos(t), -a\sin(t)) \\ &= \left(\frac{a^2-b^2}{a}\cos^3(t), \frac{b^2-a^2}{b}\sin^3(t)\right). \end{aligned}$$

En bleu la développée de l'ellipse :



Remarque. Soit C une courbe fermée simple. Le théorème de Jordan, très intuitif, dit que le complémentaire de la courbe C dans le plan $\mathbb{R}^2$ est composé de deux composantes connexes dont l'une seulement est bornée. On note A l'adhérence dans $\mathbb{R}^2$ de la composante bornée. On dit que la courbe est *convexe* si pour tous points $P_1, P_2 \in C$, le segment $[P_1P_2]$ est inclu dans A.

Il est aussi intuitif qu'une courbe fermée simple est convexe si et seulement si sa courbure reste de signe constant. Pour s'en convaincre, on peut aussi utiliser la troisième formule de Frénet : $\gamma(s) = \theta'(s)$. Si la courbure γ reste positive, alors le paramètre angulaire θ (angle entre la tangente et l'axe des abscisse) est croissant. On pourrait aussi décomposer la courbe en deux courbes : on note A le point le plus à gauche, B le point le plus à droite, $f_1 : x \mapsto y_1$ la courbe "du bas", et $f_2 : x \mapsto y_2$ la courbe "du haut". Alors f_1 est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ convexe, qui reste donc au-dessus de ses tangentes, et f_2 concave, qui reste en-dessous de ses tangentes.

Définition. Une courbe continue $f : [a, b] \to \mathbb{R}^2$ est dite *fermée* si $f(a) = f(b)$. Elle est dite *simple* si $f_{|[a,b[}$ est injective, *i.e.* si la courbe ne se recoupe pas elle-même.

Définition. Soit $f : I \to \mathbb{R}^2$ de classe $\mathcal{C}^1$. On dit que le paramétrage f est *régulier* si $f'(t) \neq 0$ pour tout $t \in I$.

Définition. Soient (I, f) et (J, g) deux arcs paramétrés. On dit qu'ils sont $\mathcal{C}^k$-équivalents s'il existe un $\mathcal{C}^k$-difféomorphisme θ de I sur J tel que $f = g \circ \theta$. On dit aussi que (J, g) est un *paramétrage admissible* de l'arc (I, f).

Définition. Soit (I, f) un arc paramétré de classe $\mathcal{C}^1$. Un reparamétrage (J, g) de l'arc est dit *normal* s'il vérifie $\|g'(u)\| = 1$ pour tout $u \in J$.

Définition. La longueur d'une courbe paramétrée $f : [a, b] \to \mathbb{R}^2$ est donnée par

$$l(f) = \sup \sum_{k=0}^{n-1} \|f(t_{i+1}) - f(t_i)\|$$

où le sup est pris sur toutes les subdivisions $t_0 = a < t_1 < \cdots < t_n = b$ de l'intervalle $[a, b]$, n étant quelconque.

Théorème. *Soit $f : [a, b] \to \mathbb{R}^2$ une courbe paramétrée de classe $\mathcal{C}^1$. Alors*

$$l(f) = \int_a^b \|f'(t)\| \, \mathrm{d}t.$$

Démonstration. Soit $t_0 = a < t_1 < \cdots < t_n = b$ une subdivision de l'intervalle $[a, b]$. Pour $i \in [\![1, n-1]\!]$, on a

$$\|f(t_{i+1}) - f(t_i)\| = \left\| \int_{t_i}^{t_{i+1}} f'(t) \, \mathrm{d}t \right\| \leqslant \int_{t_i}^{t_{i+1}} \|f'(t)\| \, \mathrm{d}t.$$

Cela implique, en sommant :

$$\sum_{k=0}^{n-1} \|f(t_{i+1}) - f(t_i)\| \leqslant \int_a^b \|f'(t)\| \, \mathrm{d}t.$$

En passant au sup sur toutes les subdivisions, on obtient

$$l(f) \leqslant \int_a^b \|f'(t)\| \, \mathrm{d}t.$$

Pour montrer l'égalité, on introduit la fonction $\varphi : [a, b] \to \mathbb{R}$ qui donne la longueur de la courbe entre a et t. Soient $t \in [a, b]$ et $h \in \mathbb{R}$ tel que $t + h \in [a, b]$. La longueur de la courbe entre les paramètres t et $t + h$ est plus grande que la longueur du segment reliant $f(t)$ à $f(t + h)$ (le plus court chemin est la ligne droite), donc

$$\|f(t + h) - f(t)\| \leqslant \varphi(t + h) - \varphi(t).$$

On en déduit l'encadrement suivant :

$$\frac{\|f(t + h) - f(t)\|}{h} \leqslant \frac{\varphi(t + h) - \varphi(t)}{h} \leqslant \frac{1}{h} \int_t^{t+h} \|f'(u)\| \, \mathrm{d}u,$$

la deuxième inégalité étant donnée par la majoration de $l(f)$ établie en premier. Les membres de droite et de gauche tendent vers la même limite $\|f'(t)\|$ quand $h \to 0$, donc par encadrement, φ est dérivable en t et $\varphi'(t) = \|f'(t)\|$. Finalement

$$\varphi(t) = \int_a^t \|f'(t)\| \, \mathrm{d}t$$

où l'on intègre à partir de a pour avoir $\varphi(a) = 0$, puis

$$l(f) = \varphi(b) = \int_a^b \|f'(t)\| \, \mathrm{d}t.$$

□

Définition. On appelle *abscisse curviligne* de (I, f) toute application s de I dans $\mathbb{R}$ telle que si $t_1, t_2 \in I$ avec $t_1 < t_2$, alors $s(t_2) - s(t_1)$ est la longueur de l'arc $([t_1, t_2], f)$.

Proposition. *Les abscisses curvilignes de l'arc orienté (I, f) sont exactement les primitives de la fonction $t \mapsto \|f'(t)\|$.*

Démonstration. Cela découle de la définition et du théorème précédent. □

Théorème. *Si (I, f) est un arc paramétré de classe $\mathcal{C}^1$ régulier, alors toute abscisse curviligne est un paramétrage admissible de classe $\mathcal{C}^1$.*

Démonstration. Soit s une abscisse curviligne. Pour tout $t \in I$, $s'(t) = \|f'(t)\| > 0$ car l'arc est régulier. Donc s est strictement croissante sur I. Comme s est continue (car $\mathcal{C}^1$), $s : I \to J = s(I)$ est bijective. De plus, s' ne s'annule pas sur I, donc s^{-1} est dérivable sur J et

$$(s^{-1})' = \frac{1}{s' \circ s^{-1}}$$

qui est continue sur J. Par conséquent, s est un $\mathcal{C}^1$-difféomorphisme, c'est donc un paramétrage admissible de (I, f) de classe $\mathcal{C}^1$. □

Proposition. *Si (I, f) est un arc paramétré de classe $\mathcal{C}^1$ régulier, alors toute abscisse curviligne est un paramétrage normal de (I, f).*

Démonstration. On note $J = s(I)$ et $F = f \circ s^{-1}$. L'arc (J, F) est un arc paramétré équivalent à (I, f). En dérivant $f = F \circ s$, on obtient $f'(t) = s'(t)F'(s(t))$, soit encore

$$F'(s(t)) = \frac{f'(t)}{s'(t)} = \frac{f'(t)}{\|f'(t)\|},$$

car toute abscisse curviligne est une primitive de $t \mapsto \|f'(t)\|$. On en déduit $\|F'(s)\| = 1$ et donc s définit un paramétrage normal (J, F) de (I, f). □

Définition. Soit $f : I \to \mathbb{R}^2$ une courbe paramétrée de classe $\mathcal{C}^2$. On dit que la courbe f est *birégulière* si $f'(t)$ et $f''(t)$ sont linéairement indépendants pour tout $t \in I$.

Définition. Soit $f : I \to \mathbb{R}^2$ une courbe régulière. Le *vecteur tangent unitaire* à la courbe à l'instant t est le vecteur

$$T(t) = \frac{f'(t)}{\|f'(t)\|}.$$

Définition. Soit $f : I \to \mathbb{R}^2$ une courbe paramétrée birégulière. La *première normale* ou *normale principale* en $t \in I$ est le vecteur

$$N(t) = \frac{T'(t)}{\|T'(t)\|}.$$

Définition. Soit $f : I \to \mathbb{R}^2$ une courbe régulière de classe $\mathcal{C}^2$ paramétrée par l'abscisse curviligne. Le nombre $\gamma(s) = \|f''(s)\| = \|T'(s)\|$ est appelé *courbure* de la courbe f en $s \in I$. C'est la norme du vecteur accélération d'un mobile parcourant la courbe à vitesse constante égale à 1.

Théorème (Formules de Frénet)**.** *Soit $f : I \to \mathbb{R}^2$ une courbe régulière de classe $\mathcal{C}^2$ paramétrée par l'abscisse curviligne. On note T le vecteur tangent unitaire, N le vecteur normal unitaire, γ la courbure et θ le paramètre angulaire (que l'on définira dans la démonstration). On a alors les formules de Frénet :*

$$T'(s) = \gamma(s)N(s),\ N'(s) = -\gamma(s)T(s),\ \textit{et}\ \ \theta'(s) = \gamma(s).$$

Démonstration. Comme la courbe est paramétrée par l'abscisse curviligne, $\gamma(s) = \|f''(s)\| = \|T'(s)\|$. Donc $N(s) = \frac{T'(s)}{\|T'(s)\|} = \frac{T'(s)}{\gamma(s)}$, ce qui donne la première formule de Frénet.

Ensuite, T est une fonction de classe $\mathcal{C}^1$ de I vers le cercle unité de $\mathbb{R}^2$, donc d'après le théorème de relèvement, il existe une fonction $\theta : I \to \mathbb{R}$ de classe $\mathcal{C}^1$ telle que $T(s) = \big(\cos(\theta(s)), \sin(\theta(s))\big)$, pour tout $s \in I$. On appelle θ le paramètre angulaire. On dérive :

$$T'(s) = \theta'(s)\big(-\sin(\theta(s)), \cos(\theta(s))\big).$$

Comme $T'(s) = \gamma(s)N(s)$, en prenant la norme dans ces deux égalités, on obtient $\theta'(s) = \gamma(s)$. On en déduit ensuite que $N(s) = \big(-\sin(\theta(s)), \cos(\theta(s))\big)$. On dérive maintenant l'expression de N et on trouve :

$$N'(s) = -\theta'(s)T(s),$$

i.e. $N'(s) = -\gamma(s)T(s)$. □

Proposition. *Soit (I, f) un arc paramétré de classe $\mathcal{C}^2$ régulier. La courbure de l'arc est donnée par la formule :*

$$\gamma(t) = \frac{\det(f'(t), f''(t))}{\|f'(t)\|^3}.$$

Démonstration. On prend s une abscisse curviligne de l'arc, et on note $F = f \circ s^{-1}$ et $J = s(I)$. Alors (J, F) est un reparamétrage de (I, f) par l'abscisse curviligne. On a donc les formules de Frénet : en particulier, $\frac{\mathrm{d}T}{\mathrm{d}s}(s) = \gamma(s)N(s)$.

On dérive $t \mapsto f(t) = F \circ s(t)$:

$$f'(t) = s'(t)\frac{\mathrm{d}F}{\mathrm{d}s}(s(t)) = s'(t)T(s(t)).$$

On dérive une deuxième fois :

$$f''(t) = s''(t)T(s(t)) + s'(t)^2 \underbrace{\frac{\mathrm{d}T}{\mathrm{d}s}(s(t))}_{=\gamma(s(t))N(s(t))}.$$

Comme (T, N) est une base directe, on a $\det(T, N) = 1$, donc :

$$\det(f'(t), f''(t)) = \det\big(s'(t)T,\, s'(t)^2\gamma(s(t))N\big) = s'(t)^3\gamma(s(t)).$$

Enfin, s étant une abscisse curviligne, c'est une primitive de $\|f'(t)\|$, donc $s'(t) = \|f'(t)\|$. On remarque alors que la courbure $\gamma(s(t))$ ne dépend pas de l'abscisse curviligne s choisie, et on la note $\gamma(t)$. On a donc démontré ce qu'on voulait. □

Références :

– Francinou, Gianella, Nicolas - *Oraux X-ENS, Analyse 4* - Page 335.

Chapitre 4

Développements bonus

4.1 Différentielle d'une limite et application exponentielle

Théorème. *Soient E et F deux espaces normés, U un ouvert de E et une suite d'applications $f_k : U \to F$ différentiables sur U. On suppose que, pour $k \to \infty$:*

(i) La suite (f_k) converge simplement sur U vers une application f.

(ii) La suite des différentielles $Df_k(x)$ converge dans $\mathcal{L}(E,F)$ uniformément pour $x \in U$ vers une application $g(x)$.

Alors l'application $\lim_{k\to\infty} f_k$ est différentiable sur U et

$$D(\lim_k f_k) = \lim_k Df_k.$$

De plus si les f_k sont de classe $\mathcal{C}^1$ sur U, il en est de même de leur limite.

Démonstration. Soit $\varepsilon > 0$. L'hypothèse (ii) donne l'existence d'un entier N tel que pour tout $k \geqslant N$ et tout $x \in U$,

$$\|Df_k(x) - g(x)\| = \sup_{\|h\|=1} \|Df_k(x).h - g(x).h\| \leqslant \varepsilon.$$

D'où, pour tout $j, k \geqslant N$,

$$\|Df_j(x) - Df_k(x)\| \leqslant 2\varepsilon.$$

Soit maintenant $a \in U$. Pour $r > 0$ assez petit, la boule de centre a et de rayon r est un convexe contenu dans U. On applique l'inégalité de la moyenne à $f_j - f_k$ entre a et $a + h$, pour $\|h\| \leqslant r$:

$$\big\|\big(f_j(a+h) - f_j(a)\big) - \big(f_k(a+h) - f_k(a)\big)\big\| \leqslant 2\varepsilon\|h\|.$$

Fixons $k \geqslant N$. Comme f_k est différentiable en a, il existe $\alpha < r$ tel que

$$\|f_k(a+h) - f_k(a) - Df_k(a).h\| = o\left(\|h\|\right) \leqslant \varepsilon\|h\|$$

pour $\|h\| \leqslant \alpha$. Par ailleurs, par définition de la norme d'application, et d'après la toute première inégalité, on a :

$$\|Df_k(a).h - g(a).h\| \leqslant \|Df_k(x) - g(x)\|\|h\| \leqslant \varepsilon\|h\|.$$

Enfin, par l'inégalité triangulaire et les trois inégalités précédentes, on obtient :

$$\|f_j(a+h) - f_j(a) - g(a).h\| \leqslant 4\varepsilon\|h\|,$$

pour tout $j \geqslant N$ et tout $\|h\| \leqslant \alpha$. On fait maintenant tendre j vers l'infini :

$$\|f(a+h) - f(a) - g(a).h\| \leqslant 4\varepsilon\|h\|,$$

ce qui établit la différentiabilité de f en a, avec $Df(a) = g(a)$.

Si les f_k sont de classe $\mathcal{C}^1$, les Df_k sont continues sur U, donc $Df = \lim_{k\to\infty} Df_k$ aussi par convergence uniforme, ce qui prouve que f est $\mathcal{C}^1$. □

Application. *L'application exponentielle est de classe $\mathcal{C}^1$ sur $\mathcal{M}_n(\mathbb{C})$.*

Démonstration. Soit la suite d'applications $(f_k)_{k\in\mathbb{N}}$ définie par :

$$\begin{array}{rccc} f_k : & \mathcal{M}_n(\mathbb{C}) & \longrightarrow & \mathcal{M}_n(\mathbb{C}) \\ & M & \longmapsto & \sum_{p=0}^{k} \frac{M^p}{p!}. \end{array}$$

Pour tout $p \geqslant 1$, l'application $M \mapsto M^p$ est de classe $\mathcal{C}^1$ sur $\mathcal{M}_n(\mathbb{C})$ car polynomiale en les coefficients de la matrice M. On développe ensuite :

$$\begin{aligned} (M+H)^p &= (M+H)(M+H)\dots(M+H) \\ &= M^p + (M^{p-1}H + M^{p-2}HM + \cdots + HM^{p-1}) + o\left(\|H\|\right), \end{aligned}$$

où l'on a bien sûr pris une norme d'algèbre sur $\mathcal{M}_n(\mathbb{C})$ (par exemple la norme d'application linéaire). Ainsi, pour $p \geqslant 1$, l'application $u_p : M \mapsto \frac{M^p}{p!}$ a pour différentielle :

$$Du_p(M).H = \frac{1}{p!}(M^{p-1}H + M^{p-2}HM + \cdots + HM^{p-1}).$$

Alors $\|Du_p(M).H\| \leqslant \frac{1}{p!}p\|M\|^{p-1}\|H\|$, ce qui nous donne :

$$\|Du_p(M)\| \leqslant \frac{\|M\|^{p-1}}{(p-1)!}.$$

Par conséquent, la série $\sum_{p\geqslant 1} Du_p(M)$ est absolument convergente dans l'espace complet $\mathcal{L}(\mathcal{M}_n(\mathbb{C}), \mathcal{M}_n(\mathbb{C}))$, et la convergence est uniforme sur toute boule $\|M\| \leqslant R$. Comme les

$$f_k(M) = \sum_{p=0}^{k} u_p(M)$$

tendent vers $\exp(M)$ lorsque $k \to \infty$, en appliquant le théorème, on obtient que l'application exponentielle est de classe $\mathcal{C}^1$ sur toute boule, donc sur $\mathcal{M}_n(\mathbb{C})$, et

$$\begin{aligned} D\exp(M).H &= \sum_{p=1}^{\infty} Du_p(M).H \\ &= \sum_{p=1}^{\infty} \frac{1}{p!}(M^{p-1}H + M^{p-2}HM + \cdots + HM^{p-1}). \end{aligned}$$

□

Remarque. On peut en fait montrer que l'application exponentielle est de classe $\mathcal{C}^\infty$.

Remarque. L'expression de la différentielle obtenue ici n'est pas très élégante. On peut montrer par une autre méthode que

$$D\exp(M).H = \exp(M)\left(\sum_{k=1}^{\infty}\frac{1}{k!}(-\varphi_M)^{k-1}(H)\right)$$

où $\varphi_M(H) = MH - HM$. On peut déduire cette expression de la notre de la façon suivante : on considère la série double

$$\begin{aligned}\left(e^M\right)^{-1}\sum_{p=1}^{\infty}Du_p(M).H &= \left(\sum_{q=0}^{\infty}\frac{(-M)^q}{q!}\right)\left(\sum_{p=1}^{\infty}Du_p(M).H\right)\\ &= \sum_{k=1}^{\infty}\left(\sum_{p+q=k}\frac{(-M)^q}{q!}Du_p(M).H\right).\end{aligned}$$

On montre ensuite par récurrence sur k que

$$\sum_{p+q=k}\frac{(-M)^q}{q!}Du_p(M).H = \frac{1}{k!}(-\varphi_M)^{k-1}(H).$$

Le calcul est élémentaire mais un peu long.

Théorème (Inégalité de la moyenne)**.** *Soient E et F deux espaces normés sur $\mathbb{R}$ ou $\mathbb{C}$, U un ouvert de E et $f : U \to F$ une application différentiable en tout point de U. Soient $[a, b]$ un segment de droite tout entier contenu dans U et k une constante positive. On suppose que pour tout $x \in [a, b]$, on a*

$$\|Df(x)\|_{\mathcal{L}(E,F)} = \sup_{\|h\|_E=1}\|Df(x).h\|_F \leqslant k.$$

Alors on a l'inégalité de la moyenne :

$$\|f(b) - f(a)\|_F \leqslant k\|b - a\|_E.$$

Références :

– Rouvière - *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation* - Page 117.

4.2 Points extrémaux de la boule unité de $\mathcal{M}_n(\mathbb{R})$

On note B la boule unité fermée de $\mathcal{M}_n(\mathbb{R})$ pour la norme

$$\|M\| = \sup_{\|X\|=1} \|MX\|,$$

où $\|MX\|$ est le norme euclienne de MX sur $\mathbb{R}^n$.

Théorème. *L'ensemble des points extrémaux de B est le groupe orthogonal $\mathcal{O}_n(\mathbb{R})$.*

Démonstration. Soit $P \in \mathcal{O}_n(\mathbb{R})$. Montrons que P est un point extrémal de B. On remarque d'abord que $P \in B$:

$$\|P\| = \sup_{\|X\|=1} \|PX\| = 1$$

car les matrices orthogonales conservent la norme. On suppose maintenant que $P = \frac{1}{2}(U+V)$ avec $U, V \in B$. Soit X un vecteur colonne unitaire. Alors :

$$\|X\| = 1 = \|PX\| = \frac{1}{2}\|UX + VX\| \leqslant \frac{1}{2}(\|UX\| + \|VX\|) \leqslant \frac{1}{2}(\|U\| + \|V\|) \leqslant 1.$$

On en déduit que toutes les inégalités écrites sont en fait des égalités. Cela implique que $\|U\| = \|V\| = 1$, $\|UX\| = \|VX\| = 1$ et de plus les vecteurs UX et VX sont positivement liés. Pour ce dernier point, il s'agit du cas d'égalité dans l'inégalité triangulaire pour une norme euclidienne. Par conséquent, on a $UX = VX$, et cela pour tout vecteur unitaire X, donc par linéarité, $U = V$. La matrice P est donc bien un point extrémal de B.

Inversement, soit $A \in B$ avec $A \notin \mathcal{O}_n(\mathbb{R})$. On va montrer que A n'est pas extrémal.

On écrit la décomposition polaire de A : il existe $O \in \mathcal{O}_n(\mathbb{R})$ et $S \in \mathcal{S}_n^+$ tels que $A = OS$. La matrice S est orthogonalement semblable à une matrice diagonale $D = \mathrm{Diag}(d_1, \ldots, d_n)$ avec les $d_i \geqslant 0$, et qu'on peut supposer rangés par ordre croissant : $d_1 \leqslant \cdots \leqslant d_n$. On écrit donc $S = {}^tPDP$ avec $P \in \mathcal{O}_n(\mathbb{R})$. Comme $S = O^{-1}A$ et que $O^{-1} \in \mathcal{O}_n(\mathbb{R})$ préserve la norme, on en déduit immédiatement que

$$\|S\| = \sup_{\|X\|=1} \|O^{-1}AX\| = \sup_{\|X\|=1} \|AX\| = \|A\| \leqslant 1.$$

Or :

$$\|S\| = \sup_{\|X\|=1} \|{}^tPDPX\| = \sup_{\|X\|=1} \|DPX\| = \sup_{\|Y\|=1} \|DY\|,$$

la dernière égalité étant justifiée par le changement de variable $X = {}^tPY$ possible car ${}^tP$ est inversible et conserve la norme. Donc $\|S\| \leqslant 1$ implique que $d_i \leqslant 1$ pour tout i (sinon $\|De_i\| = d_i > 1$, ce qui contredit $\|D\| \leqslant 1$).

Si $d_1 = 1$, alors $D = I_n$, et donc $S = I_n$ et $A = O \in \mathcal{O}_n(\mathbb{R})$. Donc nécessairement, $d_1 < 1$. On écrit alors $d_1 = \frac{\alpha+\beta}{2}$ avec $-1 \leqslant \alpha < \beta \leqslant 1$, car d_1 peut être vu comme le milieu de deux points de $[-1, 1]$. On pose alors

$$D' = \mathrm{Diag}(\alpha, d_2, \ldots, d_n) \quad \text{et} \quad D'' = \mathrm{Diag}(\beta, d_2, \ldots, d_n).$$

On a ainsi $D = \frac{1}{2}(D' + D'')$ avec $D' \neq D''$, et

$$A = \frac{1}{2}(O^t P D' P + O^t P D'' P).$$

On vérifie que les matrices $O^t P D' P$ et $O^t P D'' P$ sont dans B :

$$\|O^t P D' P\| = \sup_{\|X\|=1} \|O^t P D' P X\| = \sup_{\|X\|=1} \|D' P X\| = \sup_{\|Y\|=1} \|D' Y\| \leqslant 1$$

car les coefficients diagonaux de D' sont compris entre -1 et 1.

On a donc écrit A comme milieu de deux points distincts de B, ce qui signifie que A n'est pas un point extrémal de B. □

Admettons le théorème suivant :

Théorème (Krein-Milman)**.** *Tout convexe compact d'un espace affine de dimension finie est enveloppe convexe de l'ensemble de ses points extrémaux.*

On déduit immédiatement des deux théorèmes précédents :

Théorème. *L'enveloppe convexe de $\mathcal{O}_n(\mathbb{R})$ est B.*

Définition. Un point x d'un convexe X est dit *extrémal* si $X \backslash \{x\}$ est encore convexe. Cela équivaut à dire que x ne peut pas s'écrire comme milieu de deux points distincts de X.

Exemple. *Les points extrémaux d'un carré plein de $\mathbb{R}^2$ sont ses quatre sommets.*

Proposition (Cas d'égalité dans l'inégalité triangulaire)**.** *Soient E un espace pré-hilbertien et $u, v \in E$. Si $\|u + v\| = \|u\| + \|v\|$, alors il existe $\lambda, \mu \in \mathbb{R}_+$ tels que $\lambda u = \mu v$ (i.e. u et v sont positivement liés).*

Démonstration. Pour montrer l'inégalité triangulaire, on écrit :

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2 + 2\mathrm{Re}(u|v),$$

puis $\mathrm{Re}(u|v) \leqslant |(u|v)| \leqslant \|u\|\|v\|$ par l'inégalité de Cauchy-Schwarz. Donc l'égalité $\|u + v\| = \|u\| + \|v\|$ implique que toutes les inégalités précédentes sont en fait des égalités. L'égalité dans Cauchy-Schwarz implique l'existence de $\lambda \in \mathbb{C}$ tel que $v = \lambda u$ (si on suppose par exemple $u \neq 0$). D'où $(u|v) = \overline{\lambda}\|u\|^2$ et comme $\mathrm{Re}(u|v) = |(u|v)|$, on a que $(u|v)$ est un réel positif. Donc λ est un réel positif. □

Références :

– Francinou, Gianella, Nicolas - *Oraux X-ENS, Algèbre 3* - Page 130.

4.3 Sous-espaces de dimension finie de $\mathcal{C}(\mathbb{R},\mathbb{C})$ stables par translations

Théorème. *On note $E=\mathcal{C}(\mathbb{R},\mathbb{C})$ et F un sous-espace vectoriel de E de dimension finie. Pour $a\in\mathbb{R}$, on définit l'endomorphisme τ_a de E par $\tau_a f(x)=f(x-a)$.*

Alors F est de dimension n et est stable par tous les τ_a, $a\in\mathbb{R}$, si et seulement si F est l'espace des solutions d'une équations différentielle linéaire homogène d'ordre n à coefficients constants.

Démonstration. Si F est l'espace des solutions d'une équations différentielle linéaire homogène d'ordre n à coefficients constants, alors F est un sous-espace vectoriel de E. De plus, F est de dimension n d'après le théorème de Cauchy-Lipschitz linéaire (qui nous donne un isomorphisme entre les solutions de l'équation et les conditions initiales). De plus, il est clair qu'un tel sous-espace est stable par translations.

Réciproquement, supposons que F est de dimension n et est stable par translations. Soit $(f_1,\dots,f_n)$ une base de F. Pour $a\in\mathbb{R}$ et $i\in[\![1,n]\!]$, on a $\tau_{-a}f_i\in F$, donc il existe des scalaires $\lambda_{i1}(a),\dots,\lambda_{in}(a)$ tels que

$$f_i(x+a)=\sum_{k=1}^{n}\lambda_{ik}(a)f_k(x)\quad(*),$$

pour tout $x\in\mathbb{R}$. Pour $x\in\mathbb{R}$, notons $F_k(x)=\int_0^x f_k(t)\,\mathrm{d}t$. En intégrant $(*)$, on obtient :

$$\underbrace{\int_0^x f_i(t+a)\,\mathrm{d}t}_{=F_i(x+a)-F_i(a)}=\sum_{k=1}^{n}\lambda_{ik}(a)F_k(x).$$

Les f_i étant linéairement indépendantes, les F_i le sont aussi (sinon on obtiendrait par dérivation une relation entre les f_i). Nous démontrerons dans le lemme ci-après que cela implique l'existence de réels $x_1,\dots,x_n$ tels que la matrice $A=(F_i(x_j))_{1\leqslant i,j\leqslant n}$ soit inversible. En évaluant $(*)$ en x_j, on obtient :

$$F_i(x_j+a)-F_i(a)=\sum_{k=1}^{n}\lambda_{ik}(a)F_k(x_j),$$

pour tout i,j, soit encore $B(a)=\Lambda(a)A$ en notant $B(a)=(F_i(x_j+a)-F_i(a))_{1\leqslant i,j\leqslant n}$ et $\Lambda(a)=(\lambda_{ij}(a))_{1\leqslant i,j\leqslant n}$. On en déduit que $\Lambda(a)=B(a)A^{-1}$ pour tout $a\in\mathbb{R}$. Les f_i étant continues, les F_i sont de classe $\mathcal{C}^1$ et donc aussi l'application $a\mapsto B(a)$. Ainsi $a\mapsto\Lambda(a)$ est de classe $\mathcal{C}^1$, et donc les λ_{ij} sont de classe $\mathcal{C}^1$. En évaluant en $x=0$ dans $(*)$, on trouve

$$f_i(a)=\sum_{k=1}^{n}f_k(0)\lambda_{ik}(a),$$

pour tout $a \in \mathbb{R}$, ce qui montre que les f_i sont de classe $\mathcal{C}^1$. Par une récurrence immédiate, les f_i sont en fait de classe $\mathcal{C}^\infty$ et on obtient $F \subset \mathcal{C}^\infty(\mathbb{R},\mathbb{C})$.

On dérive maintenant $(*)$ par rapport à a et on prend $a = 0$, ce qui nous donne :

$$f_i'(x) = \sum_{k=1}^{n} \lambda_{ik}'(0) f_k(x),$$

pour tout $x \in \mathbb{R}$. D'où $f_i' \in F$ et F est donc stable par l'endomorphisme de dérivation D. Comme F est de dimension finie n et que $D_{|F}$ est un endomorphisme de F, $D_{|F}$ admet un polynôme minimal qu'on note μ, de degré $d \leqslant n$. Comme $\mu(D) = 0$ sur F, on a $F \subset \mathrm{Ker}(\mu(D)) = \{f \in E \,/\, \mu(D)(f) = 0\}$ qui est un sous-espace vectoriel de dimension $d = \deg(\mu)$ d'après le théorème de Cauchy-Lipschitz linéaire. Donc $n \leqslant d$, et finalement $d = n$, ce qui implique $F = \mathrm{Ker}(\mu(D))$ et termine la démonstration. □

Lemme. *Soient $h_1, \dots, h_n$ des fonctions de $\mathbb{R}$ dans $\mathbb{C}$ linéairement indépendantes. Alors il existe des réels $x_1, \dots, x_n$ tels que la matrice $(h_i(x_j))_{1\leqslant i,j\leqslant n}$ soit inversible.*

Démonstration. On note $K = \mathrm{Vect}(h_1, \dots, h_n)$. Pour $x \in \mathbb{R}$, on note δ_x la forme linéaire $f \mapsto f(x)$. Soient $\Gamma = \{\delta_x,\, x \in \mathbb{R}\}$ et $G = \mathrm{Vect}(\Gamma) \subset K^*$. On note ensuite

$$G^0 = \{f \in K \,/\, g(f) = 0, \text{ pour tout } g \in G\}.$$

Alors on vérifie facilement que $G^0 = \Gamma^0 = \{0\}$. D'autre part, on a $\dim(G) + \dim(G^0) = \dim(K^*)$. En effet, soit $(g_1^*, \dots, g_q^*)$ une base de G. On définit l'application linéaire v par :

$$\begin{array}{rccc} v: & K & \longrightarrow & K^* \\ & x & \longmapsto & g_1^*(x)g_1^* + \cdots + g_q^*(x)g_q^*. \end{array}$$

Alors $x \in \mathrm{Ker}(v)$ si et seulement si $g_i^*(x) = 0$ pour tout i, *i.e.* si et seulement si $\varphi(x) = 0$ pour tout $\varphi \in G$. Donc $\mathrm{Ker}(v) = G^0$. Si on note $(g_1, \dots, g_q)$ la base antéduale de $(g_1^*, \dots, g_q^*)$, alors pour tout i, $v(g_i) = g_i^*$. On en déduit que $\mathrm{Im}(v) = G$. Finalement, par le théorème du rang, $\dim(G) + \dim(G^0) = \dim(K) = \dim(K^*)$.

Donc $\dim(G) = \dim(K^*)$, d'où $G = K^*$. Il existe donc des réels $x_1, \dots, x_n$ tels que $(\delta_{x_1}, \dots, \delta_{x_n})$ soit une base de K^*. On décompose δ_{x_j} sur la base $(h_i^*)_i$:

$$\delta_{x_i} = \lambda_1 h_1^* + \cdots + \lambda_n h_n^*,$$

et en évaluant en h_i, on trouve $\lambda_i = \delta_{x_j}(h_i) = h_i(x_j)$. Finalement, $(h_i(x_j))_{1\leqslant i,j\leqslant n}$ est la matrice de passage de la base duale (h_i^*) à la base (δ_{x_i}), elle est donc inversible. □

Références :

- Aucune.

4.4 Théorème de Brauer

Théorème. *Soit K un corps de caractéristique quelconque. Pour toute permutation $\sigma \in S_n$, on note $P(\sigma) \in \mathrm{GL}_n(K)$ la matrice associée à la permutation de la base canonique de K^n par σ. Alors deux permutations σ et τ sont conjuguées dans S_n si et seulement si $P(\sigma)$ et $P(\tau)$ sont conjuguées dans $\mathrm{GL}_n(K)$ (i.e. semblables sur K).*

Démonstration. Le sens direct est immédiat : si $\tau = \varphi\sigma\varphi^{-1}$, alors

$$P(\tau) = P(\varphi\sigma\varphi^{-1}) = P(\varphi)P(\sigma)P(\varphi)^{-1},$$

car on peut vérifier à la main que $P(\sigma_1\sigma_2) = P(\sigma_1)P(\sigma_2)$ pour tout $\sigma_1, \sigma_2 \in S_n$.

Supposons maintenant qu'il existe $M \in \mathrm{GL}_n(K)$ telle que

$$P(\tau) = MP(\sigma)M^{-1}.$$

Pour montrer que σ et τ sont conjuguées, on va montrer que leurs décompositions en produits de cycles à supports disjoints comportent le même nombre de cycles de chaque longueur. On note alors $c_k(\sigma)$ le nombre de cycles de longueur k dans la décomposition de σ, et $c_k(\tau)$ celui de τ.

On remarque que

$$P(\tau^m) = P(\tau)\dots P(\tau) = MP(\sigma)\dots P(\sigma)M^{-1} = MP(\sigma^m)M^{-1}.$$

On va alors s'intéresser aux cycles de τ^m et σ^m.

Lemme. *Si φ est un cycle de longueur k, alors φ^m est le produit de d cycles à supports disjoints de longueurs $\frac{k}{d}$ où $d = \mathrm{pgcd}(k, m)$.*

Démonstration. Si m divise k, c'est clair : on écrit $\varphi = (1, 2, \dots, k)$ et on constate que

$$\varphi^m = (1, m+1, 2m+1, \dots, k-m+1)(2, m+2, \dots, k-m+2)\dots(m, 2m, \dots, k).$$

Sinon, soit $d = \mathrm{pgcd}(k, m)$. Alors les permutations φ^m et φ^d engendrent le même sous-groupe H de S_n. En effet, $\varphi^m = (\varphi^d)^{\frac{m}{d}}$ et le théorème de Bézout donne $d = mu + kv$ qui implique $\varphi^d = (\varphi^m)^u$. On considère, pour $x \in [\![1, n]\!]$, les orbites

$$O_x = \{\psi(x),\ \psi \in H\} = \{(\varphi^m)^i(x),\ i \in \mathbb{N}\} = \{(\varphi^d)^i(x),\ i \in \mathbb{N}\},$$

orbites de l'action de H sur $[\![1, n]\!]$. Comme la décomposition en cycles de $\sigma \in S_n$ est l'ensemble des orbites de l'action de $<\sigma>$ sur $[\![1, n]\!]$, on en déduit que φ^m et φ^d se décomposent en cycles de la même façon. Comme d divise k, d'après le premier cas, φ^d est le produit de d cycles à supports disjoints de longueur $\frac{k}{d}$, d'où le résultat pour φ^m. □

On applique ce lemme : chaque cycle de longueur k composant σ se décompose en $\mathrm{pgcd}(k,m)$ cycles dans σ^m. Ainsi, le nombre de cycles de σ^m vaut

$$\sum_{k=1}^{n} \mathrm{pgcd}(k,m)c_k(\sigma).$$

De même pour τ^m.

On pose $V = K^n$. En notant $(e_1,\dots,e_n)$ la base canonique de K^n, on définit une action de S_n sur K^n par $(\varphi, e_i) \mapsto e_{\varphi(i)}$. Un élément $v = (v_1,\dots,v_n) \in V$ est invariant sous une permutation $\varphi \in S_n$ lorsque ses coordonnées vérifient $v_i = v_{\varphi(i)}$ pour tout i, *i.e.* lorsqu'elles sont constantes sur les orbites de φ. En notant

$$V^{\varphi} = \{v \in V \,/\, \varphi(v) = v\},$$

qui est clairement un sous-espace vectoriel de V, on a donc que $\dim_K(V^{\varphi})$ est égal au nombre d'orbites de φ (y compris celles réduites à un point). Donc

$$\dim_K(V^{\sigma^m}) = \sum_{k=1}^{n} \mathrm{pgcd}(k,m)c_k(\sigma),$$

et de même pour τ^m. Enfin, comme $P(\tau^m) = MP(\sigma^m)M^{-1}$, on a clairement l'isomorphisme

$$\begin{array}{rccc} f: & V^{\sigma^m} & \longrightarrow & V^{\tau^m} \\ & v & \longmapsto & Mv, \end{array}$$

(injectivité immédiate car M inversible, on vérifie la surjectivité à la main), donc $\dim_K(V^{\sigma^m}) = \dim_K(V^{\tau^m})$. D'où :

$$\sum_{k=1}^{n} \mathrm{pgcd}(k,m)c_k(\sigma) = \sum_{k=1}^{n} \mathrm{pgcd}(k,m)c_k(\tau),$$

et cette égalité est vraie pour tout entier $m \geqslant 1$. On note $C(\sigma)$ le vecteur colonne constitué par les $c_k(\sigma)$ pour $1 \leqslant k \leqslant n$. De même pour τ. On note ensuite $A = (a_{ij})$ la matrice telle que $a_{ij} = \mathrm{pgcd}(i,j)$. Alors

$$AC(\sigma) = AC(\tau).$$

Il ne reste plus qu'à prouver que la matrice A est inversible pour avoir les égalités $c_k(\sigma) = c_k(\tau)$ pour tout k. On va montrer que le déterminant de A est non nul. Il s'agit du déterminant de Smith. Pour le calculer, on se rappelle que pour tout entier $d \geqslant 1$, on a $d = \sum_{k|d} \varphi(k)$ où φ est l'indicatrice d'Euler. Ainsi,

$$\mathrm{pgcd}(i,j) = \sum_{k|\mathrm{pgcd}(i,j)} \varphi(k) = \sum_{\substack{k=1 \\ k|i,\, k|j}}^{n} \varphi(k) = \sum_{k=1}^{n} r_{ik}s_{kj},$$

où $r_{ik} = \varphi(k)$ si $k|i$ et 0 sinon, $s_{kj} = 1$ si $k|j$ et 0 sinon. En notant $R = (r_{ij})$ et $S = (s_{ij})$, on obtient que $A = RS$. Mais R est S sont des matrices triangulaires (l'une inférieure, l'autre supérieure) de déterminants évidents : $\det(R) = \varphi(1)\varphi(2)\dots\varphi(n)$ et $\det(S) = 1$. D'où

$$\det(A) = \det(R)\det(S) = \varphi(1)\varphi(2)\dots\varphi(n) \neq 0.$$

□

Enoncé du théorème adapté pour la leçon *Exemples d'actions de groupes sur les espaces de matrices* :

Théorème. *Soit K un corps de caractéristique quelconque. Pour toute permutation $\sigma \in S_n$, on note $P(\sigma) \in \mathrm{GL}_n(K)$ la matrice associée à la permutation de la base canonique de K^n par σ. On considère l'action de $\mathrm{GL}_n(K)$ sur lui-même par conjugaison :*

$$(P, Q) \mapsto QPQ^{-1}.$$

Alors deux permutations σ et τ sont conjuguées dans S_n si et seulement si $P(\sigma)$ et $P(\tau)$ sont dans la même orbite.

Proposition. *Deux permutations sont conjuguées si et seulement si leurs décompositions en produits de cycles à supports disjoints comportent le même nombre de cycles de chaque longueur.*

Démonstration. Sens direct : on suppose $\sigma = \varphi\tau\varphi^{-1}$. On décompose τ en produit de cycles à supports disjoints : $\tau = c_1 \dots c_k$. Alors $\sigma = (\varphi c_1 \varphi^{-1}) \dots (\varphi c_k \varphi^{-1})$ et les $\varphi c_i \varphi^{-1}$ sont des cycles de même longueur que c_i et à supports disjoints. En effet :

$$\varphi c_i \varphi^{-1} = \varphi(\alpha_1 \dots \alpha_p)\varphi^{-1} = (\varphi(\alpha_1) \dots \varphi(\alpha_p)).$$

Sens indirect : si $\sigma = (a_1 \dots a_k)$ et $\tau = (b_1 \dots b_k)$ sont deux cycles de longueur k, alors toute permutation φ telle que $\varphi(a_i) = b_i$ vérifie $\tau = \varphi\sigma\varphi^{-1}$. On en déduit sans problème le résultat pour σ et τ lorsque leurs décompositions en produits de cycles à supports disjoints comportent le même nombre de cycles de chaque longueur. □

Références :

– Ferrand et Raoult - *Sur des matrices de permutations conjuguées* (document PDF).

Chapitre 5

Quelques résultats intéressants

5.1 Application du théorème de Chevalley-Warning

Théorème (Erdös-Ginzburg-Ziv). *Soient $n \geqslant 1$ et $a_1, \dots, a_{2n-1}$ des entiers. Alors il existe des indices $i_1, \dots, i_n$ tels que*

$$a_{i_1} + \cdots + a_{i_n} \equiv 0\,[n].$$

Démonstration. Supposons n premier, qu'on note $n = p$. On considère :

$$P_1 = \sum_{i=1}^{2p-1} a_i X_i^{p-1} \quad \text{et} \quad P_2 = \sum_{i=1}^{2p-1} X_i^{p-1}.$$

Alors $\deg(P_1) + \deg(P_2) = 2p - 2 < 2p - 1$ où $2p-1$ est le nombre d'indéterminées. On peut donc appliquer le théorème de Chevalley-Warning : le nombre N de zéros communs à P_1 et P_2 est divisible par p. Or $(0, \dots, 0)$ est un zéro commun, donc $N \geqslant p \geqslant 2$. Donc il existe $(x_1, \dots, x_{2p-1}) \in \mathbb{F}_p^{2p-1}$ non nul tel que $P_i(x_1, \dots, x_{2p-1}) = 0$ pour $i \in \{1, 2\}$. Mais si $x \neq 0$, $x^{p-1} = 1$ dans $\mathbb{F}_p$. Donc

$$P_2(x_1, \dots, x_{2p-1}) = \mathrm{Card}\left(\{i \,/\, x_i \neq 0\}\right) 1_{\mathbb{F}_p} = 0.$$

Mais $1 \leqslant \mathrm{Card}\left(\{i \,/\, x_i \neq 0\}\right) \leqslant 2p-1$, donc nécessairement, $\mathrm{Card}\left(\{i \,/\, x_i \neq 0\}\right) = p$. D'où l'existence de $i_1, \dots, i_p$ tels que $x_k = 1$ si k est l'un des i_j et 0 sinon, puis :

$$P_1(x_1, \dots, x_{2p-1}) = \sum_{i=1}^{2p-1} a_i x_i^{p-1} = \sum_{j=1}^{p} a_{i_j} = 0.$$

Pour montrer que le théorème est vrai pour n entier $\geqslant 1$ quelconque, on va montrer que s'il est vrai pour deux entiers m et n, il est vrai pour mn. Le résultat s'en suivra immédiatement en décomposant n en produit de facteurs premiers.

Supposons donc que le théorème est vrai pour deux entiers $m, n \geqslant 1$, et prenons des entiers $a_1, \dots, a_{2mn-1}$. On applique le théorème avec les $2n-1$ premiers entiers $a_1, \dots, a_{2n-1}$: il existe un sous-ensemble I_1 de $[\![1, 2n-1]\!]$ tel que $\mathrm{Card}(I_1) = n$ et

$$\sum_{i \in I_1} a_i \equiv 0\,[n].$$

On applique à nouveau le théorème avec les $2n-1$ premiers entiers a_i pour $i \in [\![1, 2mn-1]\!] \backslash I_1$: il existe I_2 tel que $\mathrm{Card}(I_2) = n$ et

$$\sum_{i \in I_2} a_i \equiv 0\,[n].$$

On recommence ainsi jusqu'à avoir effectué $2m-1$ étapes au total : en effet, au bout de $2m-2$ étapes, il reste $2mn - 1 - (2m-2)n = 2n - 1$ entiers, et au bout

de $2m-2$ étapes, il n'en reste plus que $n-1$, ce qui n'est plus suffisant. En fin de compte, pour tout $j \in [\![1, 2m-1]\!]$, il existe $c_j \in \mathbb{Z}$ tel que

$$\sum_{i \in I_j} a_i = nc_j.$$

Comme on a $2m-1$ entiers c_j, on peut appliquer le théorème pour extraire un sous-ensemble J de $[\![1, 2m-1]\!]$ tel que $\mathrm{Card}(J) = m$ et

$$\sum_{j \in J} c_j \equiv 0\,[m].$$

Finalement,

$$\underbrace{\sum_{j \in J} \sum_{i \in I_j} a_i}_{\text{somme de } mn \text{ entiers } a_i} = n \sum_{j \in J} c_j = nmk \equiv 0\,[mn],$$

ce qui termine la démonstration. □

Références :

- Nathanson - *Additive number theory* - Page 50.

5.2 Cyclicité du groupe multiplicatif d'un corps fini

Définition. Soit G un groupe fini. On appelle *exposant* de G le plus petit entier strictement positif n tel que $x^n = e$ pour tout $x \in G$. L'exposant est aussi le PPCM des ordres des éléments de G.

Proposition. *Soit G un groupe abélien fini. On note p l'exposant de G. Alors il existe un élément de G d'ordre p.*

Démonstration. On va montrer que l'ensemble des ordres des éléments de G est stable par PPCM. Ainsi le PPCM de tous les ordres des éléments de G, qui est égal en fait à l'exposant de G, est l'ordre d'un élément de G.

Soient $x, y \in G$, m l'ordre de x et n l'ordre de y.

Montrons que si m et n sont premiers entre eux, xy est d'ordre mn. On a d'une part $(xy)^{mn} = e$, et d'autre part, si $l > 0$ vérifie $(xy)^l = e$, alors $(xy)^{lm} = (x^m)^l(y^{lm}) = y^{lm} = e$. Donc lm est un multiple de n, et comme m et n sont premiers entre eux, l est un multiple de n. De la même façon, l est un multiple de m, et donc l est un multiple de mn. Cela prouve que l'ordre de xy vaut mn.

Soit maintenant $l = \text{PPCM}(m, n)$. On va écrire $l = m'n'$ avec m' et n' définis ainsi : si la décomposition en produit de puissances de nombres premiers de m contient $p_i^{m_i}$ et celle de n contient $p_i^{n_i}$, on mettra dans la décomposition de m', $p_i^{m_i}$ si $m_i \geqslant n_i$ et p_i^0 sinon, et dans celle de n', p_i^0 si $m_i \geqslant n_i$ et $p_i^{n_i}$ sinon. Alors m' et n' sont premiers entre eux, m' divise m, n' divise n, $x^{\frac{m}{m'}}$ est d'ordre m' et $y^{\frac{n}{n'}}$ est d'ordre n'. D'après ce qui précède, $x^{\frac{m}{m'}}y^{\frac{n}{n'}}$ est d'ordre $m'n' = l$, *i.e.* il existe un élément de G d'ordre $l = \text{PPCM}(m, n)$.

L'ensemble des ordres des éléments de G est donc bien stable par PPCM. □

Proposition. *Le groupe multiplicatif d'un corps fini est cyclique.*

Démonstration. Soit $\mathbb{F}_q$ un corps fini à q éléments. Alors $\text{Card}(\mathbb{F}_q^*) = q - 1$. Soit r l'exposant de $\mathbb{F}_q^*$. D'après le théorème de Lagrange, $x^{q-1} = 1$ pour tout $x \in \mathbb{F}_q^*$, donc $r \leqslant q - 1$.

Par définition de l'exposant, le polynôme $X^r - 1$ s'annule sur $\mathbb{F}_q^*$ donc possède $q - 1$ racines distinctes, donc $r \geqslant q - 1$. Finalement, $r = q - 1$.

Comme r est l'exposant de $\mathbb{F}_q^*$, d'après la proposition précédente, il existe $x \in \mathbb{F}_q^*$ d'ordre r, *i.e.* d'ordre $q - 1$, et donc $\mathbb{F}_q^* = <x>$. □

5.3 Groupes d'ordre p^2

Proposition. *Soit G un groupe qui agit sur un ensemble E. On note* $\mathrm{Stab}(x) = \{g \in G \,/\, g.x = x\}$ *le stabilisateur d'un élément x de E, et $O_x = \{g.x, g \in G\}$ l'orbite de x. Alors l'application*

$$\begin{array}{rccc} \varphi : & G/\mathrm{Stab}(x) & \longrightarrow & O_x \\ & \overline{g} & \longmapsto & g.x \end{array}$$

est bien définie et est une bijection. En particulier, si G est fini, $\mathrm{Card}(O_x)$ *divise* $\mathrm{Card}(G)$.

Remarque. L'application φ est une bijection ensembliste, on ne parle pas d'isomorphisme car $G/\mathrm{Stab}(x)$ n'est pas nécessairement un groupe. En effet, $\mathrm{Stab}(x)$ n'est en général pas distingué dans G.

Démonstration. Soient $g, g' \in G$ tels que $\overline{g} = \overline{g'}$. Alors il existe $h \in \mathrm{Stab}(x)$ tel que $g' = gh$. Ainsi,

$$g'.x = (gh).x = g.(h.x) = g.x,$$

ce qui prouve que φ est bien définie.

La surjectivité de φ est immédiate par définition d'une orbite.

Montrons l'injectivité : soient $g, g' \in G$ tels que $g.x = g'.x$. Alors

$$(g'^{-1}g).x = g'^{-1}.(g.x) = g'^{-1}.(g'.x) = (g'^{-1}g').x = e.x = x.$$

Donc $g'^{-1}g \in \mathrm{Stab}(x)$, et donc $\overline{g'} = \overline{g}$. □

Proposition. *Soit G un p-groupe (p premier) et Z son centre :*

$$Z = \{x \in G \,/\, gx = xg, \forall g \in G\}.$$

Alors Z n'est pas réduit au neutre.

Démonstration. On fait agir G sur lui-même par conjugaison. Soit $x \in G$. L'orbite de x est :

$$O_x = \{gxg^{-1}, g \in G\}.$$

On a alors clairement :

$$x \in Z \iff O_x = \{x\}.$$

D'après la proposition précédente, le cardinal d'une orbite divise le cardinal de G, il vaut donc soit 1, soit p, soit p^2. L'ensemble des orbites réalisant une partition de G, on a :

$$\mathrm{Card}(G) = \mathrm{Card}(Z) + \sum_{i=1}^{k} \mathrm{Card}(O_i),$$

où les O_i désignent les orbites de cardinaux p ou p^2. Comme p divise $\mathrm{Card}(G)$ et $\mathrm{Card}(O_i)$ pour tout i, on obtient que p divise $\mathrm{Card}(Z)$. En particulier, $\mathrm{Card}(Z) \geqslant p$ et donc Z n'est pas réduit au neutre. □

Proposition. *Tout groupe d'ordre p^2 est commutatif.*

Démonstration. Comme Z est un sous-groupe de G, $\mathrm{Card}(Z)$ divise $\mathrm{Card}(G) = p^2$, donc vaut soit p, soit p^2 (il ne peut pas valoir 1 car Z est non réduit au neutre d'après la proposition précédente). Supposons que $\mathrm{Card}(Z) = p$ et prenons $x \in G \backslash Z$. On fait agir G sur lui-même par conjugaison et on considère le stabilisateur de x :

$$\mathrm{Stab}(x) = \{g \in G \,/\, gxg^{-1} = x\}.$$

Alors de façon évidente, $Z \subset \mathrm{Stab}(x)$ et $x \in \mathrm{Stab}(x)$. Donc $\mathrm{Card}(\mathrm{Stab}(x)) > p$. Mais $\mathrm{Stab}(x)$ est un sous-groupe de G, donc son cardinal divise p^2, donc $\mathrm{Card}(\mathrm{Stab}(x)) = p^2$, *i.e.*

$$\mathrm{Stab}(x) = G.$$

Ceci signifie que pour tout $g \in G$, $gxg^{-1} = x$, soit encore que $x \in Z$. C'est absurde et cela prouve que $\mathrm{Card}(Z) = p^2$, *i.e.* $Z = G$ et G est commutatif. □

5.4 Réunion de sous-espaces stricts

Proposition. *Soit E un K-espace vectoriel de dimension finie. Alors si K est infini, E ne peut pas s'écrire comme réunion de sous-espaces stricts.*

Démonstration. Supposons que $E = V_1 \cup \cdots \cup V_N$ où les V_i sont des sous-espaces stricts de E. On a nécessairement $N \geqslant 2$ car sinon $E = V_1$. Quitte à retirer un certain nombre de sous-espaces, on peut supposer qu'aucun n'est contenu dans la réunion des autres. Il existe donc $x \in V_N$ tel que $x \notin V_1 \cup \cdots \cup V_{N-1}$. D'autre part, on n'a pas $V_1 \cup \cdots \cup V_{N-1} \subset V_N$ car sinon $E = V_N$. Donc il existe $y \in V_1 \cup \cdots \cup V_{N-1}$ tel que $y \notin V_N$.

On considère maintenant le vecteur $y + \lambda x$ pour $\lambda \in K$. Il n'appartient pas à V_N car sinon $y \in V_N$. Donc il existe $i_\lambda \in [\![1, N-1]\!]$ tel que $y + \lambda x \in V_{i_\lambda}$. On obtient ainsi une application

$$\begin{array}{rccc} \varphi : & K & \longrightarrow & [\![1, N-1]\!] \\ & \lambda & \longmapsto & i_\lambda. \end{array}$$

Cette application est injective : en effet, si λ et μ sont tels que $i_\lambda = i_\mu$, alors $y + \lambda x$ et $y + \mu x$ sont dans V_{i_λ}, donc leur différence $(\lambda - \mu)x$ aussi. Comme $x \notin V_{i_\lambda}$, on obtient $\lambda = \mu$.

Par conséquent, K ne peut être infini et on a $\mathrm{Card}(K) \leqslant N - 1$. □

Corollaire. *Soit E un K-espace vectoriel de dimension finie avec K corps infini. Soient $F_1, \ldots, F_p$ des sous-espaces vectoriels de E de même dimension r. Alors il existe un supplémentaire G commun à tous les F_i.*

Démonstration. Il existe des sous-espaces G tels que $F_i \cap G = \{0\}$ pour tout i : le sous-espace $\{0\}$ en est un. On en considère un de dimension maximale, que l'on note G. Si $r + \dim(G) < \dim(E)$, les sous-espaces $F_i \oplus G$ sont des sous-espaces stricts de E. D'après la proposition, il existe un vecteur x n'appartenant à aucun de ces sous-espaces. Mais alors le sous-espace $G' = G \oplus Kx$ est en somme directe avec tous les F_i, ce qui contredit la maximalité de G. On a donc $r + \dim(G) = \dim(E)$ et G est un supplémentaire commun à tous les F_i. □

5.5 Sous-groupes à un paramètre de $\mathrm{GL}_n(K)$

Proposition. *Soit $\varphi : \mathbb{R} \to \mathrm{GL}_n(K)$ une application continue telle que pour tout $s, t \in \mathbb{R}$, $\varphi(t+s) = \varphi(t)\varphi(s)$. Alors il existe $A \in \mathcal{M}_n(K)$ telle que pour tout $t \in \mathbb{R}$, $\varphi(t) = e^{tA}$.*

Démonstration. L'hypothèse $\varphi(t+s) = \varphi(t)\varphi(s)$ entraîne $\varphi(0) = \varphi(0)^2$, et comme $\varphi(0)$ est inversible, on obtient $\varphi(0) = I_n$.

Comme φ est continue, on peut l'intégrer et on a :

$$\frac{1}{2t}\int_{-t}^{t} \varphi(s)\,\mathrm{d}s \underset{t\to 0}{\longrightarrow} \varphi(0) = I_n.$$

Mais $\mathrm{GL}_n(K)$ est ouvert, donc il existe $\varepsilon > 0$ tel que $\int_{-\varepsilon}^{\varepsilon} \varphi(s)\,\mathrm{d}s \in \mathrm{GL}_n(K)$. On intègre maintenant l'équation $\varphi(t+s) = \varphi(t)\varphi(s)$ à t fixé pour $s \in [-\varepsilon, \varepsilon]$:

$$\int_{-\varepsilon}^{\varepsilon} \varphi(t+s)\,\mathrm{d}s = \varphi(t)\int_{-\varepsilon}^{\varepsilon} \varphi(s)\,\mathrm{d}s,$$

soit encore

$$\varphi(t) = \left(\int_{t-\varepsilon}^{t+\varepsilon} \varphi(s)\,\mathrm{d}s\right)\left(\int_{-\varepsilon}^{\varepsilon} \varphi(s)\,\mathrm{d}s\right)^{-1}.$$

Cette écriture montre que φ est dérivable et on a

$$\varphi'(t+s) = \varphi'(t)\varphi(s)$$

pour tout $t, s \in \mathbb{R}$. En particulier, on a $\varphi'(s) = \varphi'(0)\varphi(s)$ pour tout $s \in \mathbb{R}$. En posant $A = \varphi'(0)$ et en résolvant l'équation différentielle, on obtient

$$\varphi(s) = \varphi(0)e^{sA}$$

pour tout $s \in \mathbb{R}$, d'où le résultat car $\varphi(0) = I_n$. □

5.6 Tout hyperplan de $\mathcal{M}_n(K)$ rencontre $\mathrm{GL}_n(K)$

Proposition. *Soit $n \geqslant 2$. Alors tout hyperplan de $\mathcal{M}_n(K)$ rencontre $\mathrm{GL}_n(K)$.*

Démonstration. On sait qu'il existe une forme linéaire φ sur $\mathcal{M}_n(K)$ telle que $H = \mathrm{Ker}(\varphi)$.

On note E_{ij} les matrices élémentaires ayant un 1 en position (i,j) et des zéros ailleurs.

S'il existe i,j avec $i \neq j$ tels que $\varphi(E_{ij}) \neq 0$, alors $A = I_n - \frac{\varphi(I_n)}{\varphi(E_{ij})}E_{ij}$ est inversible et $\varphi(A) = 0$. Donc H rencontre $\mathrm{GL}_n(K)$.

Sinon, la matrice

$$A = E_{2,1} + E_{3,2} + \cdots + E_{n-1,n} + E_{1,n} = \begin{pmatrix} 0 & \cdots & \cdots & 1 \\ 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}$$

est clairement inversible avec $\varphi(A) = 0$ et on obtient encore que H rencontre $\mathrm{GL}_n(K)$. □

5.7 Un polynôme irréductible

Proposition. *Le polynôme $P = X^p - X - 1$ est irréductible sur $\mathbb{F}_p$.*

Démonstration. Soit K un corps de décomposition de P sur $\mathbb{F}_p$. Soit $\alpha \in K$ une racine de P. Alors pour tout $i \in [\![0, p-1]\!]$, $\alpha + i$ est aussi racine de P. En effet :

$$P(\alpha + i) = (\alpha + i)^p - (\alpha + i) - 1 = (\alpha^p + i^p) - (\alpha + i) - 1 = i^p - i,$$

où l'on a utilisé Frobenius pour dire que $(\alpha + i)^p = \alpha^p + i^p$. Enfin, comme $i \in \mathbb{F}_p$, on a $i^p - i = 0$, d'où

$$P(\alpha + i) = 0.$$

Supposons maintenant P réductible sur $\mathbb{F}_p$: on écrit $P = QR$ avec $Q, R \in \mathbb{F}_p[X]$, $\deg(Q), \deg(R) < p$. Dans $K[X]$, on a

$$Q = \prod_{k=1}^{d} (X - \alpha - i_k),$$

où $d = \deg(Q)$ et les $i_k \in [\![0, p-1]\!]$. On regarde le coefficient de X^{d-1} dans Q, il vaut :

$$-\big((\alpha + i_1) + (\alpha + i_2) + \cdots + (\alpha + i_k)\big) = -(d\alpha + i_1 + \cdots + i_k).$$

Comme $Q \in \mathbb{F}_p[X]$, ce coefficient appartient à $\mathbb{F}_p$. Donc $d\alpha \in \mathbb{F}_p$. Mais comme $d < p$, on a $d \in \mathbb{F}_p^*$, donc $\alpha \in \mathbb{F}_p$. Et le fait que $\alpha \in \mathbb{F}_p$ implique $\alpha^p = \alpha$, ce qui contredit $\alpha^p - \alpha - 1 = 0$. □

Rappelons le théorème suivant donnant un critère d'irréductibilité pour un polynôme (on regarde son irréductibilité modulo un idéal premier) :

Théorème. *On note $K = \mathrm{Frac}(A)$. Soient I un idéal premier de A et $B = A/I$: B est un anneau intègre de corps de fractions L. Soit $P = a_n X^n + \cdots + a_0 \in A[X]$ tel que $\overline{a_n} \neq 0$ dans B. Alors si $\overline{P}$ est irréductible sur B ou L, P est irréductible sur K (donc aussi dans $A[X]$ si $c(P) = 1$).*

Proposition. *Le polynôme $P = X^p - X - 1$ est irréductible sur $\mathbb{Z}$.*

Démonstration. On a prouvé que P est irréductible sur $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. Le théorème implique alors qu'il est irréductible sur $\mathbb{Q}$, et puisque son contenu vaut 1 (il est unitaire), il est irréductible sur $\mathbb{Z}$. □

5.8 Une fonction donnant tous les nombres premiers

Rappelons le théorème de Wilson qui va nous servir à démontrer la proposition qui suit :

Théorème (Wilson). *Un entier $p > 1$ est un nombre premier si et seulement si*

$$(p-1)! + 1 \equiv 0 \ [p].$$

Proposition. *Soit la fonction*

$$\begin{array}{rccl} f: & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & n & \longmapsto & 2 + 2n! \ [n+1], \end{array}$$

où la congruence est à comprendre ici comme l'unique entier congru à $2+2n!$ modulo $n+1$ et compris entre 1 et $n+1$. Alors f est à valeurs dans l'ensemble des nombres premiers et elle les atteint tous.

Démonstration. Distinguons deux cas :

(i) Si $n+1$ est premier, alors par le théorème de Wilson, $n! \equiv -1 \ [n+1]$, et donc $2 + 2n! \equiv 0 \ [n+1]$. Ainsi $f(n) = n + 1$.

(ii) Si $n + 1$ n'est pas premier, on écrit $n + 1 = ab$ avec $a, b > 1$. Si $a \neq b$, alors $n! = (ab-1)!$ est divisible par ab car il contient a et b. Donc $n+1$ divise $n!$ et $2 + 2n! \equiv 2 \ [n+1]$.

Si on ne peut pas écrire $n+1 = ab$ avec $a, b > 1$ et $a \neq b$, c'est que $n+1$ est le carré d'un nombre premier : $n+1 = p^2$. Si $p > 2$, alors p et $2p$ apparaissent dans $n!$, donc $n! \equiv 0 \ [n+1]$. Si $p = 2$, alors $n+1 = 4$ et on a $2+2n! = 14 \equiv 2 \ [n+1]$.

Finalement, dans tous les cas, si $n + 1$ n'est pas premier, $f(n) = 2$.

□