Développements

E. Hecky

6 juillet 2023

Table des matières

Chapitre 1

Couplages

1.1 Leçons d'algèbre

- 101 - Groupe opérant sur un ensemble. Exemples et applications.
  - Formule de Burnside et coloriages du cube
  - Pavages du plan
  - Indicateur de Schur-Frobenius
- 102 - Groupe des nombres complexes de module 1. Racines de l'unité. Applications.
  - Irréductibilité des polynômes cyclotomiques
  - Polygones réguliers constructibles
- 103 - Conjugaison dans un groupe. Exemples de sous-groupes distingués et de groupes quotients. Applications.
  - Lie-Kolchin
  - Table de caractères de $\mathfrak{S}_4$
  - Simplicité de $\mathfrak{A}_n$ pour $n \geqslant 5$
  - Bijection drapeaux et Sylow
  - Base de Burnside
  - Simplicité de $\mathrm{SO}_3(\mathbf{R})$
  - Pavages du plan
- 104 - Groupes finis. Exemples et applications.
  - Table de caractères de $\mathfrak{S}_4$
  - Simplicité de $\mathfrak{A}_n$ pour $n \geqslant 5$
  - Loi des cycles d'une permutation aléatoire
- 105 - Groupe des permutations d'un ensemble fini. Applications.
  - Frobenius-Zolotarev
  - Table de caractères de $\mathfrak{S}_4$
  - Simplicité de $\mathfrak{A}_n$ pour $n \geqslant 5$
  - Formule de Burnside et coloriages du cube
  - Loi des cycles d'une permutation aléatoire
- 106 - Groupe linéaire d'un espace vectoriel de dimension finie E, sous-groupes de

GL(E). Applications.
- Frobenius-Zolotarev
- Lie-Kolchin
- Nombre de matrices nilpotentes sur $\mathbf{F}_q$
- Simplicité de $SO_3(\mathbf{R})$
- Pavages du plan

— 108 - Exemples de parties génératrices d'un groupe. Applications.
- Base de Burnside
- Simplicité de $SO_3(\mathbf{R})$

— 120 - Anneaux Z/nZ. Applications.
- Réciprocité quadratique
- Théorème de Sophie Germain

— 121 - Nombres premiers. Applications.
- Réciprocité quadratique
- Polygones réguliers constructibles
- Théorème de Sophie Germain
- Probabilité que deux nombres soient premiers entre eux

— 122 - Anneaux principaux. Exemples et applications.
- Algorithme de Berlekamp
- Forme normale de Smith
- Irréductibilité des polynômes cyclotomiques
- Critère de Dumas + Eisenstein

— 123 - Corps finis. Applications.
- Frobenius-Zolotarev
- Nombre de matrices diagonalisables sur $\mathbf{F}_q$
- Réciprocité quadratique
- Algorithme de Berlekamp
- Décomposition de $X^{p^n} - X$
- Nombre de matrices nilpotentes sur $\mathbf{F}_q$

— 125 - Extensions de corps. Exemples et applications.
- Décomposition de $X^{p^n} - X$
- Polygones réguliers constructibles
- Théorème de Springer

— 126 - Exemples d'équations en arithmétique.
- Réciprocité quadratique
- Théorème de Sophie Germain

— 141 - Polynômes irréductibles à une indéterminée. Corps de rupture. Exemples et applications.
- Algorithme de Berlekamp
- Décomposition de $X^{p^n} - X$
- Irréductibilité des polynômes cyclotomiques
- Théorème de Springer
- Critère de Dumas + Eisenstein

— 142 - PGCD et PPCM, algorithmes de calcul. Applications.

- — Algorithme de Berlekamp
- — Forme normale de Smith
- — Théorème de Sophie Germain

— 144 - Racines d'un polynôme. Fonctions symétriques élémentaires. Exemples et applications.
- — Décomposition de $X^{p^n} - X$
- — Irréductibilité des polynômes cyclotomiques
- — Polygones réguliers constructibles
- — Théorème de Springer

— 148 - Exemples de décompositions de matrices. Applications.
- — Jordan-Dunford-Chevalley algorithmique
- — Nombre de matrices nilpotentes sur $\mathbf{F}_q$

— 149 - Valeurs propres, vecteurs propres. Calculs exacts ou approchés d'éléments propres. Applications.
- — Lie-Kolchin
- — Convergence d'une suite de polygones vers l'isobarycentre

— 151 - Dimension d'un espace vectoriel (on se restreindra au cas de la dimension finie). Rang. Exemples et applications.
- — Algorithme de Berlekamp
- — Base de Burnside
- — Polygones réguliers constructibles

— 152 - Déterminant. Exemples et applications.
- — Frobenius-Zolotarev
- — Conique passant par cinq points
- — Convergence d'une suite de polygones vers l'isobarycentre

— 153 - Polynômes d'endomorphisme en dimension finie. Réduction d'un endomorphisme en dimension finie. Applications.
- — Jordan-Dunford-Chevalley algorithmique
- — Nombre de matrices nilpotentes sur $\mathbf{F}_q$
- — Critère de Klarès

— 154 - Sous-espaces stables par un endomorphisme ou une famille d'endomorphismes d'un espace vectoriel de dimension finie. Applications.
- — Nombre de matrices diagonalisables sur $\mathbf{F}_q$
- — Lie-Kolchin
- — Jordan-Dunford-Chevalley algorithmique
- — Nombre de matrices nilpotentes sur $\mathbf{F}_q$
- — Bijection drapeaux et Sylow
- — Indicateur de Schur-Frobenius

— 155 - Endomorphismes diagonalisables en dimension finie.
- — Nombre de matrices diagonalisables sur $\mathbf{F}_q$
- — Jordan-Dunford-Chevalley algorithmique
- — Critère de Klarès

— 156 - Exponentielle de matrices. Applications.
- — Jordan-Dunford-Chevalley algorithmique

- L'exponentielle sur les matrices symétriques est un homéomorphisme
- Théorème de stabilité de Lyapounov

— 157 - Endomorphismes trigonalisables. Endomorphismes nilpotents.
- Lie-Kolchin
- Jordan-Dunford-Chevalley algorithmique
- Nombre de matrices nilpotentes sur $\mathbf{F}_q$
- Critère de Klarès

— 158 - Matrices symétriques réelles, matrices hermitiennes.
- L'exponentielle sur les matrices symétriques est un homéomorphisme
- Indicateur de Schur-Frobenius

— 159 - Formes linéaires et dualité en dimension finie. Exemples et applications.
- Théorème des extrema liés
- Indicateur de Schur-Frobenius

— 160 - Endomorphismes remarquables d'un espace vectoriel euclidien (de dimension finie).
- Table de caractères de $\mathfrak{S}_4$
- Simplicité de $\mathrm{SO}_3(\mathbf{R})$
- Pavages du plan
- L'exponentielle sur les matrices symétriques est un homéomorphisme

— 161 - Distances dans un espace affine euclidien. Isométries.
- Pavages du plan
- Conserver l'orthogonalité

— 162 - Systèmes d'équations linéaires ; opérations élémentaires, aspects algorithmiques et conséquences théoriques.
- Conique passant par cinq points
- Forme normale de Smith

— 170 - Formes quadratiques sur un espace vectoriel de dimension finie. Orthogonalité, isotropie. Applications.
- Théorème de Springer
- Indicateur de Schur-Frobenius
- Composantes connexes des formes quadratiques

— 171 - Formes quadratiques réelles. Coniques. Exemples et applications.
- Conique passant par cinq points
- Composantes connexes des formes quadratiques

— 181 - Barycentres dans un espace affine réel de dimension finie, convexité. Applications.
- Conique passant par cinq points
- Convergence d'une suite de polygones vers l'isobarycentre

— 190 - Méthodes combinatoires, problèmes de dénombrement.
- Lemme de Sperner + Théorème de Brouwer
- Nombre de matrices diagonalisables sur $\mathbf{F}_q$
- Réciprocité quadratique
- Décomposition de $X^{p^n} - X$
- Nombre de matrices nilpotentes sur $\mathbf{F}_q$

  - Formule de Burnside et coloriages du cube
  - Probabilité que deux nombres soient premiers entre eux
- 191 - Exemples d'utilisation de techniques d'algèbre en géométrie.
  - Conique passant par cinq points
  - Pavages du plan
  - Polygones réguliers constructibles
  - Conserver l'orthogonalité

1.2 Leçons d'analyse

- 201 - Espaces de fonctions. Exemples et applications.
  - Stone-Weierstrass / Bernstein
  - Approximation de Korovkin
  - Trois droites de Hadamard / Riesz-Thorin
  - Espace de Bergman
  - Montel et Osgood
  - Densité des polynômes orthogonaux
- 203 - Utilisation de la notion de compacité.
  - Stone-Weierstrass / Bernstein
  - Approximation de Korovkin
  - Lemme de Sperner + Théorème de Brouwer
  - Montel et Osgood
  - Prokhorov et Lévy
- 204 - Connexité. Exemples et applications.
  - Sunyer i Balaguer
  - Simplicité de $\mathrm{SO}_3(\mathbf{R})$
  - Composantes connexes des formes quadratiques
  - Les isométries locales sont des isométries affines
- 205 - Espaces complets. Exemples et applications.
  - Espace de Bergman
  - Lax-Milgram et $-(pu')' + qu' + u = f$
  - Montel et Osgood
  - Sunyer i Balaguer
- 206 - Exemples d'utilisation de la notion de dimension finie en analyse.
  - Théorème des extrema liés
  - Théorème de stabilité de Lyapounov
- 208 - Espaces vectoriels normés, applications linéaires continues. Exemples.
  - Trois droites de Hadamard / Riesz-Thorin
  - Espace de Bergman
- 209 - Approximation d'une fonction par des fonctions régulières. Exemples et applications.
  - Stone-Weierstrass / Bernstein
  - Approximation de Korovkin

  - Densité des polynômes orthogonaux
  - Théorème taubérien de Hardy-Littlewood
- 213 - Espaces de Hilbert. Bases hilbertiennes. Exemples et applications.
  - Espace de Bergman
  - Lax-Milgram et $-(pu')' + qu' + u = f$
  - Densité des polynômes orthogonaux
- 214 - Théorème d'inversion locale, théorème des fonctions implicites. Exemples et applications en analyse et en géométrie.
  - Théorème des extrema liés
  - Les isométries locales sont des isométries affines
- 215 - Applications différentiables définies sur un ouvert de $\mathbf{R}^n$. Exemples et applications.
  - Théorème des extrema liés
  - Théorème de stabilité de Lyapounov
  - Les isométries locales sont des isométries affines
- 219 - Extremums : existence, caractérisation, recherche. Exemples et applications.
  - Trois droites de Hadamard / Riesz-Thorin
  - Inégalité isopérimétrique
  - Théorème des extrema liés
- 220 - Équations différentielles ordinaires. Exemples de résolution et d'études de solutions en dimension 1 et 2.
  - Lax-Milgram et $-(pu')' + qu' + u = f$
  - Nombre de zéros d'une solution d'équation différentielle
  - Théorème de stabilité de Lyapounov
- 221 - Équations différentielles linéaires. Systèmes d'équations différentielles linéaires. Exemples et applications.
  - Nombre de zéros d'une solution d'équation différentielle
  - Théorème de stabilité de Lyapounov
- 223 - Suites numériques. Convergence, valeurs d'adhérence. Exemples et applications.
  - Sunyer i Balaguer
  - Développement asymptotique d'une suite récurrente
- 224 - Exemples de développements asymptotiques de suites et de fonctions.
  - Développement asymptotique d'une suite récurrente
  - Nombre de zéros d'une solution d'équation différentielle
- 226 - Suites vectorielles et réelles définies par une relation de récurrence $u_{n+1} = f(u_n)$. Exemples. Applications à la résolution approchée d'équations.
  - Développement asymptotique d'une suite récurrente
  - Convergence d'une suite de polygones vers l'isobarycentre
- 228 - Continuité, dérivabilité des fonctions réelles d'une variable réelle. Exemples et applications.
  - Sunyer i Balaguer
  - Bohr-Mollerup et intégrale de log Gamma
  - Prokhorov et Lévy

- 229 - Fonctions monotones. Fonctions convexes. Exemples et applications.
  - Bohr-Mollerup et intégrale de log Gamma
  - Prokhorov et Lévy
- 230 - Séries de nombres réels ou complexes. Comportement des restes ou des sommes partielles des séries numériques. Exemples.
  - Théorème taubérien de Hardy-Littlewood
  - Développement asymptotique d'une suite récurrente
  - Probabilité que deux nombres soient premiers entre eux
- 234 - Fonctions et espaces de fonctions Lebesgue-intégrables.
  - Trois droites de Hadamard / Riesz-Thorin
  - Espace de Bergman
  - Densité des polynômes orthogonaux
- 235 - Problèmes d'interversion en analyse.
  - Théorème taubérien de Hardy-Littlewood
  - Formule des compléments
- 236 - Illustrer par des exemples quelques méthodes de calcul d'intégrales de fonctions d'une ou plusieurs variables.
  - Inégalité isopérimétrique
  - Bohr-Mollerup et intégrale de log Gamma
  - Formule des compléments
- 239 - Fonctions définies par une intégrale dépendant d'un paramètre. Exemples et applications.
  - Bohr-Mollerup et intégrale de log Gamma
  - Densité des polynômes orthogonaux
  - Formule des compléments
- 241 - Suites et séries de fonctions. Exemples et contre-exemples.
  - Stone-Weierstrass / Bernstein
  - Approximation de Korovkin
  - Montel et Osgood
  - Théorème taubérien de Hardy-Littlewood
  - Formule de Poisson
  - Raikov
- 243 - Séries entières, propriétés de la somme. Exemples et applications.
  - Espace de Bergman
  - Théorème taubérien de Hardy-Littlewood
  - Raikov
- 245 - Fonctions d'une variable complexe. Exemples et applications.
  - Trois droites de Hadamard / Riesz-Thorin
  - Espace de Bergman
  - Montel et Osgood
  - Densité des polynômes orthogonaux
  - Formule des compléments
  - Raikov
- 246 - Séries de Fourier. Exemples et applications.

  - Inégalité isopérimétrique
  - Formule de Poisson
- 250 - Transformation de Fourier. Applications.
  - Densité des polynômes orthogonaux
  - Formule de Poisson
- 253 - Utilisation de la notion de convexité en analyse.
  - Trois droites de Hadamard / Riesz-Thorin
  - Lemme de Sperner + Théorème de Brouwer
  - Bohr-Mollerup et intégrale de log Gamma
- 261 - Loi d'une variable aléatoire : caractérisations, exemples, applications.
  - Prokhorov et Lévy
  - Raikov
  - Loi des cycles d'une permutation aléatoire
- 262 - Convergences d'une suite de variables aléatoires. Théorèmes limite. Exemples et applications.
  - Prokhorov et Lévy
  - Loi des cycles d'une permutation aléatoire
- 264 - Variables aléatoires discrètes. Exemples et applications.
  - Raikov
  - Loi des cycles d'une permutation aléatoire
- 265 - Exemples d'études et d'applications de fonctions usuelles et spéciales.
  - Bohr-Mollerup et intégrale de log Gamma
  - Formule des compléments
  - Formule de Poisson
- 266 - Illustration de la notion d'indépendance en probabilités.
  - Stone-Weierstrass / Bernstein
  - Raikov
- 267 - Exemples d'utilisation de courbes en dimension 2 ou supérieure.
  - Trois droites de Hadamard / Riesz-Thorin
  - Inégalité isopérimétrique
  - Formule des compléments
  - Nombre de zéros d'une solution d'équation différentielle

Chapitre 2

Développements

2.1 Développements d'algèbre

2.1.1 Algorithme de Berlekamp

Leçons 122, 123, 141, 142, 151

Référence Objectif agrégation

Prérequis. Théorème chinois ; algorithmes de division euclidienne et d'Euclide.

Théorème 2.1.1. *Soient q une puissance d'un nombre premier et $P \in \mathbf{F}_q[X]$ un polynôme de degré n et sans facteur carré. Alors il existe un algorithme donnant la décomposition de P en produit d'irréductibles. Précisément, on trouve un facteur non constant de P en $O(qn^3)$ opérations.*

Démonstration.

Lemme 2.1.2. *Pour $R \in \mathbf{F}_q[X]$, le morphisme :*

$$S_R : \begin{array}{ccc} \mathbf{F}_q[X]/R & \longrightarrow & \mathbf{F}_q[X]/R \\ \overline{Q} & \longmapsto & \overline{Q(X^q)} \end{array}$$

est bien défini et coïncide avec $\overline{Q} \mapsto \overline{Q}^q$.

Démonstration. L'application $S_0 : Q \in \mathbf{F}_q[X] \mapsto Q(X^q) = Q^q$ est un morphisme d'anneaux, on note $S = \pi \circ S : \mathbf{F}_q[X] \to \mathbf{F}_q[X]/R$ avec π la projection modulo R. Comme $S(R) = 0$, le morphisme S passe au quotient et donne S_R. Ensuite :

$$S_R(Q \bmod R) = S_R(\pi(Q)) = \pi(Q(X^q)) = \pi(Q^q) = \pi(Q)^q$$

comme annoncé. □

Voici la description de l'algorithme. On note $\pi : \mathbf{F}_q[X] \to \mathbf{F}_q[X]/P$ la projection et $x = \pi(X)$. On note $\mathcal{B} = (1, x, \ldots, x^{\deg P - 1})$ qui est une base de $\mathbf{F}_q[X]/P$.

— On calcule la matrice de $S_P - \mathrm{id}$ dans $\mathcal{B}$;
— Le nombre de facteurs irréductibles de P est :

$$r = \dim \mathrm{Ker}(S_P - \mathrm{id}) = \deg P - \mathrm{rg}(S_P - \mathrm{id}).$$

Si $r = 1$, on s'arrête car P est irréductible ;
— On calcule un polynôme V non constant modulo P tel que $\pi(V) \in \mathrm{Ker}(S_P - \mathrm{id})$. Avec l'algorithme d'Euclide, on calcule la décomposition :

$$P = \prod_{\alpha \in \mathbf{F}_q} \mathrm{pgcd}(P, V - \alpha),$$

et l'on recommence avec chaque facteur non trivial.

Il faut montrer que cet algorithme termine et est correct. Notons $P = P_1 \cdot \ldots \cdot P_r$ une décomposition de P en produit d'irréductibles, qui sont premiers entre eux parce que P est sans facteur carré. Notons $\kappa_i = \mathbf{F}_q[X]/P_i$. Par théorème chinois, on a alors un isomorphisme :

$$\phi : \mathbf{F}_q[X]/P \to \kappa_1 \times \cdots \times \kappa_r.$$

Notons $\widehat{S_P} = \phi \circ S_P \circ \phi^{-1}$ qui correspond à l'élévation à la puissance q dans $\kappa_1 \times \cdot \times \kappa_r$. Alors :

$$\begin{aligned}(x_1, \ldots, x_r) \in \mathrm{Ker}(\widehat{S_P} - \mathrm{id}) \text{ ssi } & (x_1^q, \ldots, x_r^q) = (x_1, \ldots, x_r) \\ \text{ssi } & \forall i,\ x_i^q = x_i \text{ dans } \kappa_i.\end{aligned}$$

Or, chaque κ_i est une extension de $\mathbf{F}_q$ et l'on y a $\mathbf{F}_q = \{x \in \kappa_i \mid x^q = x\}$. Ainsi, $(x_1, \ldots, x_r) \in \mathrm{Ker}(\widehat{S_P} - \mathrm{id})$ est équivalent à demander que $x_i \in \mathbf{F}_q \subset \kappa_i$ pour tout i. Par conséquent, $\mathrm{Ker}(\widehat{S_P} - \mathrm{id}) = \mathbf{F}_q^r$ et la dimension de ce noyau compte donc bien le nombre de facteurs irréductibles de P.

On suppose alors $r > 1$, ce qui montre que $\mathrm{Ker}(S_P - \mathrm{id})$ n'est pas réduit à une droite, donc qu'il y existe un polynôme V non constant modulo P. Comme $\pi(V) \in \mathrm{Ker}(S_P - \mathrm{id})$, ce qui précède montre que V modulo P_i est constant, égal disons à $\alpha_i \in \mathbf{F}_q \subset \kappa_i$, pour tout i.

Soit $\alpha \in \mathbf{F}_q$. Comme $\mathrm{pgcd}(P, V - \alpha)$ divise P, il est de la forme $\prod_{I_\alpha} P_i$, et comme les P_i sont premiers entre eux, on a $I_\alpha = \{i \mid P_i | V - \alpha\}$. Or, $\alpha_i = \alpha$ si et seulement si P_i divise $V - \alpha$, donc $I_\alpha = \{i \mid \alpha_i = \alpha\}$. Finalement :

$$P = \prod_{\alpha \in \mathbf{F}_q} \prod_{i | \alpha_i = \alpha} P_i = \prod_{\alpha \in \mathbf{F}_q} \mathrm{pgcd}(P, V - \alpha).$$

Comme V modulo P n'est pas constant, au moins deux facteurs de ce produit sont non triviaux donc le degré décroît strictement à chaque exécution de l'algorithme. En particulier, l'algorithme termine. Et d'après tout ce que l'on vient de voir, il est correct.

Calculons la complexité de l'algorithme. La mise sous forme sans facteur carré revient au calcul de $\text{pgcd}(P, P')$ ce qui se fait en $O(n^2)$ opérations.

La matrice est calculée avec n divisions euclidiennes de polynômes de degré $\leqslant nq$ par P, donc ce calcul se fait en complexité $O(n(n(nq-n+1))) = O(qn^3)$.

Le pivot de Gauss pour calculer le noyau de la matrice se fait en $O(n^3)$ opérations.

Dans le pire cas, on doit calculer n pgcd non triviaux de degré toujours $\leqslant n$ donc un calcul en $O(n^3)$.

On trouve donc un facteur non constant de P en $O(qn^3)$ opérations. □

Remarques.

— Il faut savoir exécuter cet algorithme sur un exemple. C'est pas difficile si on sait rapidement faire des divisions euclidiennes de polynômes. Par exemple avec $q = 5$ et $P = X^3 + X^2 + X + 2$, on trouve comme matrice :

$$S_P - \text{id} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 4 & 2 \end{pmatrix}$$

qui est de rang 1. On remarque en fait que la première colonne sera toujours nulle et qu'elle correspond à la droite du noyau formée des polynômes constants modulo P. On trouve alors un autre élément du noyau, comme $(0, 2, 1)$, qui correspond à $V = X^2+2X$. Enfin, il faut calculer les cinq $\text{pgcd}(P, V-\alpha)$. Les deux premiers et le dernier valent 1, le troisième vaut X^2+2X+3 et le quatrième $X+4$. Ces deux polynômes sont les facteurs irréductibles de P. On le sait parce qu'on a trouvé $r = 2$ et deux pgcds non triviaux donc on peut s'arrêter ! On sent vraiment une sorte de magie dans le calcul de V et des pgcds, qui séparent miraculeusement les facteurs irréductibles.

— Ce n'est pas nécessaire de donner la complexité pendant le développement, mais s'il reste du temps c'est de très bon goût. Et c'est bien de la connaître pour répondre facilement à une éventuelle question du jury à ce propos. De même, il faut bien connaître la complexité de la division euclidienne et celle de l'algorithme d'Euclide.

— Quand P a des facteurs carrés, on peut le mettre sous forme sans facteur carré en le divisant par $\text{pgcd}(P, P')$, et exécuter l'algorithme sur ce quotient et sur le pgcd. Attention, P' peut être nul, mais cela n'arrive que lorsque P est de la forme R^p avec p la caractéristique du corps de base. Et dans ce cas, on peut calculer R et exécuter l'algorithme sur ce polynôme.

— L'algorithme est utilisé dans la pratique pour le calcul efficace du logarithme discret, donc c'est utile en cryptographie.

Recasages.

— 122 : On illustre, un peu comme avec la forme normale de Smith, un algorithme de calcul sur des anneaux de polynômes. La démonstration utilise plusieurs fois la principalité de $\mathbf{F}_q[X]$, par exemple pour dire que l'irréductibilité de P_i entraîne

que $\kappa_i = \mathbf{F}_q[X]/P_i$ est un corps. On utilise aussi le fait que c'est un anneau euclidien pour calculer effectivement des restes, et que c'est un anneau factoriel pour les pgcd. Attention à bien motiver la présence de cet algorithme dans la leçon quand même.

— 123 : C'est une très bonne leçon pour ce développement qui est réellement utile en pratique lorsque l'on manipule des corps finis.

— 141 : Parfait. On parle de polynômes irréductibles et de corps de rupture, c'est difficile de faire mieux.

— 142 : On utilise à fond les algorithmes de calcul de division euclidienne et de pgcd. C'est donc une application concrète de l'algorithme d'Euclide. En plus, il y a ce côté magique des pgcds qui font apparaître les facteurs, donc on sent que le calcul des pgcd est central.

— 151 : Voir le nombre de facteurs irréductibles d'un polynôme en calculant la dimension d'un espace vectoriel, ça paraît très fort. Il faut donc mettre l'accent sur ce point. On peut aussi le rapprocher à la factorisation des polynômes cyclotomiques sur les corps finis : Φ_n se décompose sur $\mathbf{F}_q$ en un produit de $\phi(n)/r$ polynômes irréductibles distincts et tous de même degré r égal à l'ordre de q dans $(\mathbf{Z}/n\mathbf{Z})^\times$. Ce théorème utilise le théorème de la base télescopique (de manière équivalente, la description des sous-corps de $\mathbf{F}_{q^r}$), donc le lien avec la leçon n'est pas idiot (cf. Demazure, page 217).

2.1.2 Base de Burnside

Leçons 103, 108, 151

Référence Un max de maths

Prérequis. Sous-groupe dérivé.

Théorème 2.1.3. *Les parties génératrices minimales d'un p-groupe fini G ont toutes le même cardinal.*

Démonstration. On commence par le lemme suivant.

Lemme 2.1.4. *Tout sous-groupe maximal $H < G$ est distingué, et le quotient G/H est cyclique d'ordre p.*

Démonstration. On pose $N = \{g \in G \mid gH = Hg\}$ le normalisateur de H. C'est un sous-groupe de G, il suffit de montrer que $N = G$ pour avoir le résultat. Comme N contient H, il suffit en fait de montrer que $|N| > |H|$ puisque H est maximal.

Le sous-groupe H agit sur l'ensemble G/H par translation à gauche. La formule des classes s'écrit :

$$|G/H| = \left|(G/H)^H\right| + \sum_{|\mathrm{Orb}(gH)|>1} |\mathrm{Orb}(gH)|.$$

Comme p divise le membre de gauche et la somme à droite, on a $\left|(G/H)^H\right| \geqslant p > 1$. Or,

$$(G/H)^H = \{gH \in G/H \mid \forall h \in H,\ hgH = gH\} = \{gH \mid Hg = gH\} = N/H.$$

Donc $|N| = |H|\left|(G/H)^H\right| > |H|$, comme voulu. □

Notons $\mathcal{M}$ l'ensemble des sous-groupes maximaux de G et $\Phi(G) = \bigcap_{H\in\mathcal{M}} H$. D'après le lemme, pour tout $H \in \mathcal{M}$ on a $H \lhd G$ et G/H est alors un groupe d'ordre p, c'est-à-dire C_p, donc il est abélien, d'où $H \supset DG$. Ainsi, $\Phi(G)$ est distingué dans G et contient DG donc $G/\Phi(G)$ est un groupe abélien.

On montre alors le résultat d'abord sur $G/\Phi(G)$ qui est un p-groupe élémentaire : pour $x, y \in G/\Phi(G)$ et $\lambda \in \mathbf{F}_p$, on pose :

$$x \oplus y = xy, \qquad \lambda \otimes x = x^\lambda.$$

Pour vérifier que la multiplication est bien définie, il faut vérifier que $p \otimes x = 0$. Mais si x' est un représentant de x, $H \in \mathcal{M}$ et $\pi : G \to G/H$ est la projection, on a $\pi(x'^p) = p\pi(x') = 0$ car G/H est cyclique d'ordre p. Ainsi $x'^p \in H$ pour tout H donc $x'^p \in \Phi(G)$, d'où $p \otimes x = 0$.

Ces opérations munissent $G/\Phi(G)$ d'une structure de $\mathbf{F}_p$-espace vectoriel de dimension finie, dont les bases sont exactement les parties génératrices minimales de $G/\Phi(G)$. Cela démontre le résultat.

Démonstration dans le cas général. Soit $(g_1, \dots, g_n)$ une partie génératrice de G minimale (qui existe car G est fini). Alors en notant $\pi : G \to G/\Phi(G)$ la projection, la famille $(\pi(g_1), \dots, \pi(g_n))$ est génératrice de $G/\Phi(G)$. Il suffit de montrer qu'elle est libre, puisqu'alors toutes les parties génératrices minimales de G auront pour cardinal la dimension de $G/\Phi(G)$.

Si cette famille était liée, disons $\pi(g_1) = \pi(g_2)^{\alpha_2} \cdot \ldots \cdot \pi(g_n)^{\alpha_n}$, alors on aurait $g_1 y^{-1} \in \Phi(G)$ après avoir posé $y = g_2^{\alpha_2} \cdot \ldots \cdot g_n^{\alpha_n}$. La famille $(g_1 y^{-1}, g_2, \dots, g_n)$ est alors encore une partie génératrice de G (car $g_1 = g_1 y^{-1} y$ et $y \in \langle g_2, \dots, g_n \rangle$). Or, $(g_2, \dots, g_n)$ n'est pas une partie génératrice (par minimalité) donc il existe $H \in \mathcal{M}$ tel que $\langle g_2, \dots, g_n \rangle \subset H$. Et puisque $g_1 y^{-1} \in \Phi(G) \subset H$, la famille $(g_1 y^{-1}, g_2, \dots, g_n)$ ne peut pas générer un sous-groupe plus grand que H... □

Remarques.

- Le lemme utilise la même technique de réduction d'une équation aux classes modulo p que ce qu'on fait quand on montre que le centre d'un p-groupe fini n'est jamais trivial. C'est bien de savoir faire ce parallèle.
- On peut dire à l'oral que $G/\Phi(G)$ est un p-groupe élémentaire, c'est-à-dire que tous les éléments sont de puissance p triviale. On utilise le même argument de $\mathbf{F}_p$-espace vectoriel quand on dit que les groupes dans lesquels tout élément est de carré trivial sont isomorphes à un C_p^n.
- Le sous-groupe $\Phi(G)$ est assez important en théorie des groupes, on l'appelle le *sous-groupe de Frattini* de G. Il est toujours caractéristique dans G, et on peut le voir comme l'ensemble des éléments superflus de G (les éléments qui peuvent être retirés de toute partie génératrice). D'ailleurs la partie où l'on montre que $G/\Phi(G) \cong \mathbf{F}_p^n$ s'appelle le *théorème de Frattini*.
- Il est important de connaître des exemples et des contre-exemples pour accompagner ce théorème. Par exemple, le groupe quaternionique Q_8 est un 2-groupe fini, ses parties génératrices minimales sont de cardinal 2 (par exemple, (i, j)). Le groupe symétrique $\mathfrak{S}_n$ avec $n \geqslant 3$ n'est pas un p-groupe, et il a des parties génératrices minimales de cardinaux différents (par exemple, $((1\ 2), (1\ 2\ \dots\ n))$ et $((i\ i+1))_{i<n}$). Le p-groupe de Prüfer :

  $$\mathbf{Z}_{p^\infty} = \{z \in \mathbf{C} \mid \exists \alpha \in \mathbf{N},\ z^{p^\alpha} = 1\}$$

  est un p-groupe mais il n'a *aucune partie génératrice minimale* (si on prend deux éléments d'une partie génératrice, l'un appartient toujours au sous-groupe engendré par l'autre. En fait, tous les sous-groupes de $\mathbf{Z}_{p^\infty}$ sont cycliques d'ordre une puissance de p, mais ce groupe n'est pas lui-même cyclique ni fini).

Recasages.

- 103 : C'est bien, on n'y parle pas beaucoup de conjugaison parce que les quotients que l'on manipule sont abéliens, mais on fait quand même intervenir de manière cruciale un normalisateur. On obtient un exemple important de sous-groupe distingué et du quotient correspondant (à rajouter au zoo du centre, sous-groupe dérivé, etc).

— 108 : C'est vraiment bien, surtout si l'on fait une partie ou une sous-partie entièrement dédiée à l'étude des parties génératrices minimales. On peut aussi y traiter le cas des groupes abéliens de type fini par exemple (dont le calcul des parties génératrices minimales est très intéressant). On peut aussi donner des bornes sur les cardinaux minimaux des parties génératrices, et tout comparer sur des cas particuliers.
— 151 : On dispose avec ce développement d'un exemple assez rare où la notion de dimension (finie) d'un espace vectoriel apparaît de manière aussi cruciale dans une démonstration. Le développement rentre très naturellement dans cette leçon, surtout si l'on y développe une sous-partie à propos des p-groupes abéliens élémentaires.

2.1.3 Bijection drapeaux et Sylow

Leçons 103, 154

Référence H2G2 tome 1

2.1.4 Conique passant par cinq points

Leçons 152, 162, 171, 181, 191

Référence Géométrie (Eiden)

Prérequis. Classifications des coniques et des formes quadratiques.

Théorème 2.1.5. *Par cinq points distincts du plan affine passe une conique. Elle est unique si et seulement si quatre de ces points ne sont jamais alignés. Elle est non dégénérée si et seulement si trois de ces points ne sont jamais alignés.*

Démonstration. Notons ces points A, B, C, D et E. Si les cinq points sont alignés, il suffit de prendre cette droite et n'importe quelle autre.

Sinon, sans perdre de généralité (A, B, C) est un repère affine du plan : notons-y (X, Y, Z) les coordonnées barycentriques. L'équation d'une conique passant par A, B et C est alors de la forme :

$$pYZ + qXZ + rXY = 0.$$

Notons (x_1, y_1, z_1) et (x_2, y_2, z_2) les coordonnées barycentriques respectives de D et E dans le repère. La conique passe par ces points si et seulement si :

$$\begin{cases} py_1z_1 + qx_1z_1 + rx_1y_1 = 0 \\ py_2z_2 + qx_2z_2 + rx_2y_2 = 0 \end{cases}$$

Ce système de deux équations en les trois inconnues p, q et r est de rang au plus deux, donc il y a toujours des solutions (p, q, r) non nulles. Une conique qui passe par les cinq points existe donc toujours.

Discutons de l'unicité. Plusieurs coniques conviennent si et seulement si le rang de ce système est $\leqslant 1$, ce qui est le cas si et seulement si les mineurs de taille 2 s'annulent. Ces mineurs sont :

$$z_1z_2\begin{vmatrix} 0 & x_1 & x_2 \\ 0 & y_1 & y_2 \\ 1 & z_1 & z_2 \end{vmatrix}, \quad y_1y_2\begin{vmatrix} 0 & x_1 & x_2 \\ 1 & y_1 & y_2 \\ 0 & z_1 & z_2 \end{vmatrix}, \quad x_1x_2\begin{vmatrix} 1 & x_1 & x_2 \\ 0 & y_1 & y_2 \\ 0 & z_1 & z_2 \end{vmatrix}.$$

Dans le cas où les coefficients devant ces déterminants ne sont pas nuls, c'est-à-dire lorsque $x_1y_1z_1x_2y_2z_2 \neq 0$, c'est-à-dire encore lorsque ni D ni E n'appartient à l'une des droites (AB), (AC) et (BC), le système est de rang $\leqslant 1$ si et seulement si $A, B, C \in (DE)$. Ce n'est pas possible, puisque A, B et C ne sont pas alignés. Ainsi, si la conique n'est pas unique, alors par exemple $D \in (AB)$, ce qui donne $z_1 = 0$, $x_1y_1 \neq 0$ et $y_1z_2 = 0$. En particulier $E \in (AB)$ et quatre points sont alignés.

Réciproquement, si quatre points sont alignés, alors il y a une infinité de coniques qui passent par les cinq points : il suffit de prendre la droite qui aligne les quatre points, et une autre droite qui passe par le cinquième.

Discutons de la dégénérescence. La conique est non dégénérée si et seulement si la forme quadratique de matrice :

$$\begin{pmatrix} 0 & r & q \\ r & 0 & p \\ q & p & 0 \end{pmatrix}$$

n'est pas dégénérée. Son discriminant est $2pqr$: s'il est nul, disons $p = 0$, alors l'équation de la conique devient :

$$0 = qXZ + rXY = X(qZ + rY)$$

qui est la réunion de deux droites. Il est impossible de placer cinq points sur deux droites sans que trois ne soient alignés.

Réciproquement, si trois points sont alignés, il existe clairement une réunion de deux droites qui passe par les cinq points, c'est donc l'unique conique en question et elle est bien dégénérée. □

Corollaire 2.1.6. *Deux coniques du plan affine soit sont confondues, soit s'intersectent au plus quatre fois, soit partagent une droite en commun.*

Remarques.

- Mieux vaut pour ce développement être très au clair au sujet de la classification des coniques, et de la classification des formes quadratiques (pour les éventuelles questions du jury).
- Le développement n'est pas dur et peut facilement s'apprendre par cœur, mais il faut faire attention de ne pas s'embrouiller dans l'argument reliant le rang du système à l'alignement des points.
- Il faut absolument faire des dessins au tableau. Quand trois ou quatre points sont alignés, on dessine les droites correspondantes et on rappelle que c'est une conique. De même, pour dire que $D \in (AB)$ si et seulement si $z_1 = 0$ par exemple, on peut faire le dessin du repère affine et rappeler le principe des coordonnées barycentriques. C'est un développement de géométrie, donc s'il n'y a pas de dessin, le jury ne sera pas content.

Recasages.

- 152 : On fait un bel usage du déterminant et des mineurs pour détecter quand le système est de rang trop petit. Les trois déterminants que l'on écrit sont très élégants et il faut comprendre pourquoi ils sont de cette forme. De manière générale, on peut dans le plan expliciter les liens entre déterminants et alignements de points.
- 162 : Comme avant, c'est le rang d'un système linéaire qui détermine le nombre de coniques qui passent par les cinq points. Il ne faut pas oublier que résoudre un système linéaire revient à calculer une intersection d'hyperplans, ici on calcule une intersection de deux droites projectives dans un plan projectif. Le cas où la conique est unique est le cas général, et le cas où il existe une infinité de coniques

est le cas où ces deux droites projectives sont confondues. On voit alors le lien entre le fait que choisir cinq points dont quatre (ou trois) sont alignés est un choix de mesure nulle, tout comme choisir deux droites dans le plan projectif et espérer qu'elles soient confondues.

— 171 : C'est parfait, en plus le développement renforce le lien déjà fort entre les coniques et les formes quadratiques. En démontrant ce théorème, on montre que l'on a compris de quoi on parle.

— 181 : Les coordonnées barycentriques rendent la résolution du problème infiniment plus simple et élégante que si l'on avait choisi des coordonnées linéaires. Donc on tient ici une jolie application du concept de barycentre.

— 191 : Beaucoup de développements à propos de coniques rentrent dans cette leçon, car les coniques sont des objets géométriques que l'on manipule grâce à notre connaissance des formes quadratiques. On transforme alors le problème originellement géométrique en un problème à propos de systèmes linéaires, que les techniques d'algèbre linéaire savent bien traiter.

2.1.5 Conserver l'orthogonalité

Leçons 161, 191

Référence Oraux Algèbre 3

Prérequis. Le seul morphisme de corps $\mathbf{R} \to \mathbf{R}$ est l'identité.

Théorème 2.1.7. *Soit $f : \mathbf{R}^2 \to \mathbf{R}^2$ une application conservant l'orthogonalité : si ABC est un triangle rectangle en B, alors $f(A)f(B)f(C)$ est un triangle rectangle en $f(B)$. Alors f est une similitude affine.*

Démonstration. On démontre tout par étapes.

f est injective. Si A et B sont distincts, on prend C tel que ABC soit rectangle. Alors $f(A)f(B)f(C)$ est rectangle donc $f(A) \neq f(B)$.

f conserve l'alignement. Soient A, B et C trois points alignés distincts. On choisit D tel que (AC) et (CD) soient perpendiculaires. Les triangles ACD et BCD sont rectangles en C, donc $f(A)f(C)f(D)$ et $f(B)f(C)f(D)$ sont rectangles en $f(C)$. Ainsi les droites $(f(C)f(A))$ et $(f(C)f(B))$ sont toutes les deux perpendiculaires à $(f(C)f(D))$ et passent par $f(C)$ dont $f(A)$, $f(B)$ et $f(C)$ sont alignés.

En fait l'ordre est conservé. Supposons par exemple que $C \in [AB]$. Alors on peut choisir D comme avant en demandant en plus que ABD soit rectangle en D (avec un cercle). Alors $f(A)f(B)f(D)$ est rectangle en $f(D)$ et le point $f(C)$ appartient à $[f(A)f(B)]$ puisque c'est le projeté orthogonal de $f(D)$ sur $(f(A)f(B))$.

f conserve le parallélisme. Soient (AB) et (CD) des droites parallèles distinctes. La perpendiculaire à (AB) passant par A coupe (CD) en un point E que l'on peut peut supposer différent de C (quitte à échanger C et D). Les triangles $f(A)f(B)f(E)$ et $f(A)f(E)f(C)$ sont rectangles respectivement en $f(A)$ et en $f(E)$. Ainsi $(f(A)f(B))$ et $(f(C)f(E)) = (f(C)f(D))$ sont toutes les deux perpendiculaires à la même droite $(f(A)f(E))$.

Ainsi, f conserve les parallélogrammes.

f est affine. Posons :

$$\varphi : \begin{array}{ccc} \mathbf{R}^2 & \longrightarrow & \mathbf{R}^2 \\ a & \longmapsto & f(a) - f(0) \end{array}$$

dont on doit montrer qu'elle est linéaire.

φ est un morphisme de groupes. Soient $a, b \in \mathbf{R}^2$ et $c = a + b$. Si a et b ne sont pas colinéaires alors $0acb$ est un parallélogramme, donc $f(0)f(a)f(c)f(b)$ aussi. Ainsi $f(c) - f(a) = f(b) - f(0)$, c'est-à-dire $\varphi(c) = \varphi(a) + \varphi(b)$.

Si a et b sont colinéaires, cela revient à montrer le point suivant :

Pour $a \in \mathbf{R}^2$ et $t \in \mathbf{R}$, on a $\varphi(ta) = t\varphi(a)$. C'est évident si $a = 0$. Sinon, ta est sur la droite $(0a)$ donc $f(ta) \in (f(0)f(a))$, et il existe t' tel que :

$$f(ta) - f(0) = t'(f(a) - f(0)).$$

Comme $a \neq 0$, $f(a) \neq f(0)$ donc ce t' est unique. On pose $\sigma : t \mapsto t'$ et l'on montre que c'est l'identité.

σ est un morphisme de corps de $\mathbf{R}$. On sait déjà que $\sigma(0) = 0$ et que $\sigma(1) = 1$. Soient $t, t' \in \mathbf{R}$ non nuls.

On choisit b tel que a et b ne soient pas colinéaires. Les points 0, ta, $b + ta$ et b forment un parallélogramme, et de même pour les points $t'a$, $(t + t')a$, $b + ta$ et b. Les images de ces deux parallélogrammes sont des parallélogrammes, donc :

$$f((t + t')a) - f(t'a) = f(ta) - f(0)$$

ce qui se réécrit $\varphi((t + t')a) = \varphi(ta) + \varphi(t'a)$ d'où $\sigma(t + t') = \sigma(t) + \sigma(t')$.

Les droites (ab) et $(tatb)$ sont parallèles donc leurs images aussi, et de même $(t'ab)$ et $(tt'atb)$ sont parallèles donc leurs images aussi. Tout cela se réécrit en $\sigma(tt') = \sigma(t)\sigma(t')$.

On conclut que σ est un morphisme de corps de $\mathbf{R}$ donc c'est l'identité, donc φ est bien linéaire donc f est affine.

f est une similitude. Comme f est injective φ l'est aussi donc c'est un automorphisme. Soit (e_1, e_2) une base orthonormée de $\mathbf{R}^2$. Alors $(\varphi(e_1), \varphi(e_2))$ est aussi une base orthogonale. Comme $e_1 + e_2$ et $e_1 - e_2$ sont orthogonaux, $\varphi(e_1) + \varphi(e_2)$ et $\varphi(e_1) - \varphi(e_2)$ le sont aussi. Donc $\varphi(e_1)$ et $\varphi(e_2)$ ont la même norme $k > 0$. On voit alors que φ est la composée d'une homothétie de rapport k et d'une transformation orthogonale. □

Remarques.

- C'est un peu trop long si on veut écrire beaucoup, mais il faut dire tout ça à l'oral et faire des dessins en n'écrivant que les noms des étapes. Il faut profiter de l'originalité de ce développement : contrairement à beaucoup d'autres, il ressemble un peu à une chasse au trésor où l'on utilise l'hypothèse très faible pour trouver à chaque fois des propriétés de plus en plus fortes.
- Il faut évidemment faire énormément de dessins. Souvent quelques traits sont suffisants pour que l'argument soit convaincant. On peut passer vite sur certains points, sinon on n'en finit jamais.
- La raison pour laquelle l'identité est le seul morphisme d'anneaux $\mathbf{R} \to \mathbf{R}$ est la suivante. Il est clair qu'un morphisme σ doit fixer $\mathbf{Q}$ qui est le sous-corps premier de $\mathbf{R}$ (on peut le faire à la main). Ensuite, il suffit de montrer que σ est forcément continu. Pour cela, on remarque qu'il préserve l'ordre : si $x \geqslant 0$ alors disons $x = y^2$, puis $\sigma(x) = \sigma(y)^2 \geqslant 0$. En particulier, si $|a - b| \leqslant 1/n$, on obtient $|\sigma(a) - \sigma(b)| \leqslant |\sigma(1)|/n = 1/n$. Et le seul prolongement continu de l'identité sur $\mathbf{Q}$ à tout $\mathbf{R}$ est l'identité.
- Il faut bien insister sur le côté magique du résultat. L'hypothèse sur f est vraiment, vraiment très faible !

Recasages.

- 161 : On peut carrément faire une partie dans le plan qui traite des similitudes, si l'on veut. On fait à un moment usage de la métrique sur la plan affine, quand on dit qu'il existe un D qui rende ABD triangle. Pour le construire, on construit le cercle de diamètre $[AB]$ et on choisit D à l'intersection entre ce cercle et la perpendiculaire à (AB) passant par C. En quelque sorte, on peut donc mettre ce développement après des propriétés des triangles isocèles si l'on veut. Bon, ce n'est quand même pas le cœur du développement.
- 191 : On ne s'attend absolument pas, en lisant l'énoncé, à devoir reconnaître des morphismes de corps ! Bon à part ça, c'est un peu léger, alors il faut mettre ce fait en valeur.

2.1.6 Critère de Dumas + Eisenstein

Leçons 122, 141

Référence Petit compagnon des nombres

Prérequis. Inclure au moins la définition et la proposition suivantes dans le plan. Soient p un nombre premier et τ un réel quelconque.

Définition 2.1.8. La *valuation p-adique penchée de pente* τ est la fonction :

$$\begin{aligned} \mathbf{Q}[t] &\longrightarrow \mathbf{R} \cup \{\infty\} \\ v_{p,\tau} : \sum_i a_i t^i &\longmapsto \min\{v_p(a_i) - \tau i : a_i \neq 0\} \end{aligned}$$

en ayant convenu que $v_{p,\tau}(0) = \infty$. Pour $f = \sum_i a_i t^i \in \mathbf{Q}[t]$, on pose aussi $m_{p,\tau}(f)$ (resp. $M_{p,\tau}(f)$) le plus petit (resp. le plus grand) i tel que $v_{p,\tau}(f) = v_p(a_i) - \tau i$. Une *pente du polygone de Newton p-adique de* f est une pente τ telle que la *largeur* $\Delta_{p,\tau}(f) = M_{p,\tau}(f) - m_{p,\tau}(f)$ soit strictement positive.

Proposition 2.1.9. *On a :*

1. $v_{p,\tau}(f) = \infty$ *si et seulement si* $f = 0$;
2. $v_{p,\tau}(fg) = v_{p,\tau}(f) + v_{p,\tau}(g)$;
3. $v_{p,\tau}(f+g) \geqslant \min(v_{p,\tau}(f), v_{p,\tau}(g))$;
4. *si* f *et* g *sont non nuls, alors* $M_{p,\tau}(fg) = M_{p,\tau}(f) + M_{p,\tau}(g)$ *et* $m_{p,\tau}(fg) = m_{p,\tau}(f) + m_{p,\tau}(g)$;
5. $\Delta_{p,\tau}(fg) = \Delta_{p,\tau}(f) + \Delta_{p,\tau}(g)$.

Démonstration. 1. Immédiat.

2. Si $f = \sum_i a_i t^i$ et $g = \sum_i b_i t^i$ alors $fg = \sum_i c_i t^i$ avec $c_i = \sum_{j=0}^{i} a_j b_{i-j}$. On a alors $v_p(c_i) \geqslant \min\{v_p(a_j) + v_p(b_{i-j})\}$ d'où :

$$v_p(c_i) - \tau i \geqslant \min\{(v_p(a_j) - \tau j) + (v_p(b_{i-j}) - \tau(i-j))\}$$

donc $v_{p,\tau}(fg) \geqslant v_{p,\tau}(f) + v_{p,\tau}(g)$.

Pour l'inégalité réciproque : on pose $r = m_{p,\tau}(f)$ et $s = m_{p,\tau}(g)$. Alors $c_{r+s} = \sum_{j=0}^{r+s} a_j b_{r+s-j}$, le terme en $j = r$ étant de valuation $v_p(a_r b_s) = \tau(r+s) + v_{p,\tau}(f) + v_{p,\tau}(g)$ et les autres termes étant de valuation plus grande. Ainsi :

$$v_p(c_{r+s}) = \tau(r+s) + v_{p,\tau}(f) + v_{p,\tau}(g)$$

d'où :

$$v_{p,\tau}(fg) \leqslant v_{p,\tau}(f) + v_{p,\tau}(g).$$

3. Immédiat.

4. Avec les notations qui précèdent, si $i < r+s$ alors chaque terme de $\sum_{j=0}^{i} a_j b_{i-j}$ est de valuation $> \tau i + v_{p,\tau}(f) + v_{p,\tau}(g)$ donc $m_{p,\tau}(fg) = r+s$. La démonstration est la même pour $M_{p,\tau}$.
5. Immédiat avec le point précédent.

□

Définition 2.1.10. On définit le *dénominateur réduit* d'une pente $\tau \in \mathbf{Q}$ comme le plus petit entier $q > 0$ tel que $\tau q \in \mathbf{Z}$.

Lemme 2.1.11. *Soit $\tau \in \mathbf{Q}$ de dénominateur réduit q. Alors l'image de la fonction $v_{p,\tau}$ (en enlevant 0 qui est envoyé sur ∞) est $\frac{1}{q}\mathbf{Z}$.*

Démonstration. Il est clair que $v_{p,\tau}(\mathbf{Q}[t] \setminus \{0\})$ est inclus dans $\frac{1}{q}\mathbf{Z}$. Réciproquement, tout élément de $\frac{1}{q}\mathbf{Z}$ s'écrit $a + \tau b$ avec $a \in \mathbf{Z}$ et $b \in \mathbf{N}$, et alors $p^a t^b$ convient. □

Théorème 2.1.12 (critère de Dumas)**.** *Soit $h \in \mathbf{Q}[t]$ tel que $h(0) \neq 0$. S'il existe un nombre premier p tel que h admette une pente τ de largeur $\Delta_{p,\tau}(h) = \deg h$ et de dénominateur réduit $\deg h$ alors h est irréductible.*

Démonstration. Comme la largeur est égale à $\deg h$ on a nécessairement $m_{p,\tau}(h) = 0$ et $M_{p,\tau}(h) = \deg h$. Si h se factorise en $h = fg$ alors d'après la proposition on a aussi $m_{p,\tau}(f) = m_{p,\tau}(g) = 0$, $M_{p,\tau}(f) = \deg f$ et $M_{p,\tau}(g) = \deg g$. Ainsi d'après le lemme, $\Delta_{p,\tau}(f) = \deg f$ et $\Delta_{p,\tau}(g) = \deg g$ sont multiples du dénominateur réduit qui est $\deg h$. Comme $\deg f + \deg g = \deg h$, l'un des deux degrés de f ou de g doit être nul. □

Théorème 2.1.13 (critère d'Eisenstein)**.** *Soit $h = \sum_{i=0}^{d} a_i t^i \in \mathbf{Q}[t]$. S'il existe un nombre premier p tel que $v_p(a_0) = 1$, $v_p(a_d) = 0$ et $v_p(a_i) \geqslant 1$ pour les autres i, alors h est irréductible.*

Démonstration. C'est une application directe du critère de Dumas, avec la pente $\tau = -\frac{1}{\deg h}$. □

Remarques.

— Bien sûr, tout faire est trop long. Justement, c'est bien : on peut choisir ce que l'on veut bien démontrer. C'est surtout les points 4 et 5 qui nous intéressent dans la proposition, donc bien sûr c'est inutile de démontrer le 2 dans le développement. Le lemme est assez immédiat et c'est le sens facile qui nous intéresse. Le cœur du développement est donc évidemment dans les critères de Dumas et Eisenstein, qui une fois les polygones de Newton introduits, sont très visuels.

— C'est de bon goût, je pense, de faire une partie ou une sous-partie qui parle des polygones de Newton. Ça donne des jolis dessins et une manière très visuelle de voir l'irréductibilité des polynômes. Le livre *Petit compagnon des nombres* en parle très bien, et c'est intéressant d'au moins lire la page Wikipédia à leur sujet.

— Le jury risque de poser une question sur la démonstration "usuelle" du critère d'Eisenstein, qui consiste à réduire la factorisation $h = fg$ modulo p puis à discuter des valuations des coefficients. La discussion risque d'être intéressante dans la leçon 122, mais pas tant dans la leçon 141. C'est quand même bien de connaître au moins l'idée, et l'intérêt de passer par les polygones de Newton (en l'occurrence, le critère de Dumas est beaucoup plus général).

— C'est important de connaître les interprétations visuelles. On trace dans le plan les points $(i, v_p(a_i))$. Le polygone de Newton est formé de l'enveloppe convexe de ces points et de tout ce qu'il y a au-dessus. La valuation penchée $v_{p,\tau}(f)$ est alors la plus haute ordonnée à l'origine d'une droite qui reste en-dessous de tous ces points (cette droite est un hyperplan de soutien du polygone de Newton). Les nombres $m_{p,\tau}(f)$ et $M_{p,\tau}(f)$ sont respectivement la plus petite et la plus grande abscisse où la droite touche le polygone. Le critère de Dumas dit que s'il existe un polygone de Newton de f avec un segment reliant le point d'abscisse 0 et celui d'abscisse $\deg f$ ne rencontrant aucun autre point, alors f est irréductible. Le critère d'Eisenstein est le cas particulier avec les points $(0, 1)$ et $(\deg f, 0)$.

Recasages.

— 122 : Pas la meilleure leçon mais le critère d'Eisenstein est valable dans tous les anneaux factoriels. Attention cette démonstration fonctionne dans les anneaux de la forme $K[t]$ avec K un corps valué, genre $\mathbf{Q}$ avec les valuations p-adiques ou $k(x)$ avec l'ordre d'annulation en un point. Donc pas tous les anneaux factoriels.

— 141 : Là c'est vraiment parfait, même qu'une partie sur les polygones de Newton est bienvenue. Attention, ça peut demander un peu de recul sur la notion : il existe par exemple des polytopes de Newton importants en géométrie tropicale.

2.1.7 Critère de Klarès

Leçons 153, 155, 157

Référence Mansuy - Mneimné 3e Ed

Prérequis. Décompositions de Dunford et de Jordan ; deux endomorphismes diagonalisables qui commutent sont codiagonalisables.

Définition 2.1.14. Soient E un $\mathbf{C}$-espace vectoriel de dimension finie et $u \in \mathcal{L}(E)$. On définit :

$$\mathrm{ad}_u : \begin{array}{ccc} \mathcal{L}(E) & \longrightarrow & \mathcal{L}(E) \\ v & \longmapsto & uv - vu. \end{array}$$

Théorème 2.1.15. *Soient E un $\mathbf{C}$-espace vectoriel de dimension finie et $u \in \mathcal{L}(E)$. Alors u est diagonalisable si et seulement si* $\mathrm{Ker}(\mathrm{ad}_u) = \mathrm{Ker}(\mathrm{ad}_u^2)$.

Démonstration. Commençons par écrire la décomposition de Dunford $u = d + n$, où d est diagonalisable, n nilpotent et $dn = nd$.

Montrons qu'il existe $v \in \mathcal{L}(E)$ tel que $n = uv - vu$. On commence par traiter le cas où u est le bloc de Jordan standard J_k de taille k. On pose $M_k = \mathrm{diag}(1, 2, \ldots, k)$ et l'on vérifie immédiatement que $J_k M_k - M_k J_k = J_k$, qui est bien la partie nilpotente de J_k.

Dans le cas général, notons F_λ le sous-espace caractéristique de E pour u associé à une valeur propre λ. La restriction de u à F_λ est alors de la forme $\lambda \mathrm{id}_{F_\lambda} + n_{F_\lambda}$ avec n_{F_λ} nilpotent. Par décomposition de Jordan, dans une certaine base de F_λ, la matrice N_λ de n_{F_λ} est diagonale par blocs $\mathrm{diag}(J_{k_1}, \ldots, J_{k_s})$ avec $k_1 + \ldots + k_s = \dim F_\lambda$. On voit immédiatement que $V_\lambda = \mathrm{diag}(M_{k_1}, \ldots, M_{k_s})$ vérifie $N_\lambda = N_\lambda V_\lambda - V_\lambda N_\lambda$. Comme l'identité commute avec V_λ et $u|_{F_\lambda} = \lambda \mathrm{id}_{F_\lambda} + n_{F_\lambda}$, l'endomorphisme v_λ de F_λ associé à V_λ dans la base de jordanisation convient : $n_{F_\lambda} = u|_{F_\lambda} v_\lambda - v_\lambda u|_{F_\lambda}$.

Il suffit alors de recoller tous les v_λ en un $v \in \mathcal{L}(E)$, ce qui est possible puisque $E = \bigoplus_\lambda F_\lambda$.

Supposons que $\mathrm{Ker}(\mathrm{ad}_u) = \mathrm{Ker}(\mathrm{ad}_u^2)$. Cette condition équivaut bien sûr à $\mathrm{Ker}(\mathrm{ad}_u) \cap \mathrm{Im}(\mathrm{ad}_u) = 0$. Comme n commute avec u (il commute avec d et lui-même donc avec $n + d = u$, ou plus court : c'est un polynôme en u), on a $n \in \mathrm{Ker}(\mathrm{ad}_u)$. La première étape montre que $n \in \mathrm{Im}(\mathrm{ad}_u)$ donc $n = 0$ et u est diagonalisable.

Réciproquement, supposons u diagonalisable. Pour $f \in \mathcal{L}(E)$, posons :

$$\psi_f : \begin{array}{ccc} \mathcal{L}(E) & \longrightarrow & \mathcal{L}(E) \\ v & \longmapsto & fv, \end{array} \qquad \phi_f : \begin{array}{ccc} \mathcal{L}(E) & \longrightarrow & \mathcal{L}(E) \\ v & \longmapsto & vf. \end{array}$$

Alors $\psi_f^k = \psi_{f^k}$ et $\phi_f^k = \phi_{f^k}$. Ainsi, les polynômes annulateurs de u annulent aussi ψ_u et ϕ_u : ces deux applications linéaires sont donc diagonalisables. Mais ils commutent, donc

ils sont diagonalisables dans une même base, qui diagonalise alors aussi $\mathrm{ad}_u = \psi_u - \phi_u$. Ainsi ad_u est diagonalisable, d'où $\mathrm{Ker}(\mathrm{ad}_u) = \mathrm{Ker}(\mathrm{ad}_u^2)$. □

Remarques.

- Il faut absolument savoir expliquer la décomposition de Dunford qui est centrale ici. Soit on peut le justifier de la manière "usuelle" (décomposer l'espace en sous-espaces caractéristiques, etc) soit on peut expliquer la manière algorithmique qui consiste à appliquer une méthode de Newton en partant de u pour arriver à la partie diagonalisable.
- De même, il faut bien connaître la décomposition de Jordan. On peut soit connaître les idées de la démonstration "usuelle", soit savoir que l'on peut la déduire du théorème de structure des modules (de torsion et) de type fini sur un anneau principal (c'est un corollaire de l'existence des formes normales de Smith, dont on déduit aussi le théorème de structure des groupes abéliens de type fini, et le théorème de réduction de Frobenius).
- On peut passer assez vite sur le passage de bloc de Jordan à sous-espace caractéristique à cas général. Une fois qu'on a fait le cas de Jordan et qu'on a expliqué pourquoi ça marchait avec une valeur propre car l'identité commute avec tout, le recollement ne demande pas d'effort. Bien sûr, il faut savoir justifier pourquoi $E = \bigoplus_\lambda F_\lambda$: c'est parce que le corps de base est $\mathbf{C}$ (ou n'importe quel corps algébriquement clos).
- L'équivalence entre $\mathrm{Ker}(T) = \mathrm{Ker}(T^2)$ et $\mathrm{Ker}(T) \cap \mathrm{Im}(T) = 0$ doit bien sûr être maîtrisée parfaitement, surtout qu'elle est vraiment facile (et on n'a besoin que du sens le plus facile).
- Il faut savoir expliquer pourquoi deux endomorphismes diagonalisables qui commutent sont codiagonalisables. Les sous-espaces propres de l'un sont stables par l'autre.
- On parle de ad_u, il faut donc savoir pourquoi on l'appelle comme ça. C'est la *représentation adjointe* de l'algèbre de Lie des matrices carrées à coefficients complexes.
- On peut penser qu'on tient ici un super algorithme pour déterminer si une matrice est diagonalisable. Attention, la diagonalisabilité d'une matrice de taille n s'exprime ici avec des noyaux d'endomorphismes de $\mathcal{L}(\mathbf{K}^n)$, c'est-à-dire des pivots de Gauss sur des matrices de taille n^2. Comme le pivot de Gauss sur une matrice de taille m se fait en $O(m^3)$ opérations, on finit avec un critère de diagonalisabilité qui se calcule explicitement en $O(n^6)$ opérations. Il y a plus rapide : on peut calculer le polynôme minimal μ de M en calculant la forme normale de Smith de $XI - M$, et regarder si μ et μ' sont premiers entre eux. Avec un pivot de Gauss puis un algorithme d'Euclide, on obtient une complexité en $O(n^3)$.

Recasages.

- 153 : C'est vraiment bien ici. On utilise les critères de diagonalisabilité avec les polynômes, la codiagonalisabilité et la décomposition de Dunford. C'est difficile

de trouver meilleur développement.

— 155 : On a un critère de diagonalisabilité qui s'exprime avec le calcul de deux noyaux, c'est quand même assez joli. Voir la dernière remarque quand même : ce n'est pas vraiment viable comme algorithme. On peut quand même s'amuser à regarder ce que ça fait sur des matrices 2×2 ou (avec beaucoup de courage pour les noyaux de matrices 9×9) sur des matrices 3×3.

— 157 : C'est une jolie application de la décomposition de Dunford et de la décomposition de Jordan. Donc ça rentre à la fois dans le côté trigonalisable et dans le côté nilpotent ! (C'est important de ne pas trop séparer ces deux notions dans la leçon.)

2.1.8 Décomposition de $X^{p^n} - X$

Leçons 123, 125, 141, 144, 190

Référence Oraux Algèbre 1

2.1.9 Forme normale de Smith

Leçons 122, 142, 162

Référence Objectif agrégation

Prérequis. Aucun.

Théorème 2.1.16. *Soient A un anneau euclidien et M une matrice à coefficients dans A. Alors il existe une famille $(d_1, \dots, d_s)$ d'éléments non nuls de A avec $d_s | \dots | d_1$, telle que M soit équivalente à :*

$$\begin{pmatrix} d_s & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & d_1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Cette famille est unique à association près.

Démonstration. On décrit un algorithme pour passer de M à sa forme normale de Smith. Le résultat reste vrai si A est seulement principal, mais alors on n'a plus d'algorithme.

1. Si $U = 0$ on s'arrête.

2. Permuter des lignes et colonnes pour placer l'élément de stathme minimal tout en haut à gauche.

3. On traite la première colonne.

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

$$m_{i,1} = q m_{1,1} + r_1$$

et soustraire q fois la première ligne à la i-ème.

3b. Si $r_i \neq 0$, échanger les lignes 1 et i, et retourner en 3a.

3c. Si $r_i = 0$ et i n'était pas la dernière ligne, faire $i \leftarrow i + 1$ et retourner en 3a.

3d. Si $r_i = 0$ et i était la dernière ligne, passer à l'étape 4.

4. Ici, tous les coefficients sous le premier sont nuls dans la première colonne. On traite la première ligne.

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

$$m_{1,j} = qm_{1,1} + s_j$$

et soustraire q fois la première colonne à la j-ème.

4b. Si $s_j \neq 0$, échanger les colonnes 1 et j, pleurer un bon coup, et retourner en 3.

4c. Si $s_j = 0$ et j n'était pas la dernière colonne, faire $j \leftarrow j + 1$ et retourner en 4a.

4d. Si $s_j = 0$ et j était la dernière colonne, passer à l'étape 5.

5. Ici, tous les coefficients de la première ligne et tous les coefficients de la première colonne sont nuls, sauf celui en haut à gauche.

5a. Si $m_{1,1}$ ne divise pas tous les coefficients de la sous-matrice en bas à droite, disons qu'un tel coefficient se trouve sur la colonne $j_{\text{putain de bordel de merde}}$. Ajouter la $j_{\text{putain de bordel de merde}}$-ème colonne à la première et retourner à l'étape 3.

5b. Recommencer à l'étape 1 avec la sous-matrice en bas à droite.

Maintenant, il faut montrer que l'algorithme termine, qu'il est correct, et l'unicité des facteurs invariants.

L'algorithme termine. Le stathme de $m_{1,1}$ décroît strictement à chaque retour en arrière dans l'algorithme.

L'algorithme est correct. Chaque étape de l'algorithme est une opération élémentaire sur les lignes ou les colonnes, donc la matrice que l'on obtient à la fin est bien équivalente à M. Ensuite, elle est bien de la forme attendue puisque l'on ne passe, en étape 5b, à la sous-matrice en bas à droite que lorsque le coefficient en haut à gauche est tout seul, et l'on a le droit d'atteindre 5b que si ce coefficient divise tous les coefficients de la sous-matrice en bas à droite. Donc les d_i se divisent les uns les autres.

Unicité des facteurs invariants à association près. Notons $I_k(M)$ l'idéal de A engendré par les mineurs de taille k de M. Un calcul direct montre que $I_k(PM) \subset I_k(M)$, et de même $I_k(MQ) \subset I_k(M)$. Ainsi si M' est équivalente à M alors on a d'une part $I_k(M') \subset I_k(M)$ et d'autre part $I_k(M) \subset I_k(M')$. Finalement, $I_k(M)$ reste constant au cours de l'algorithme. Une fois terminé, l'algorithme donne :

$$I_k(M) = \langle d_s d_{s-1} \cdot \ldots \cdot d_{s-k+1} \rangle.$$

Ainsi si l'on dispose d'autres facteurs invariants $(d'_s, \dots, d'_1)$, alors :

$$\begin{aligned} \langle d_s \rangle &= \langle d'_s \rangle \\ \langle d_s d_{s-1} \rangle &= \langle d'_s d'_{s-1} \rangle \\ \langle d_s d_{s-1} d_{s-2} \rangle &= \langle d'_s d'_{s-1} d'_{s-2} \rangle \\ \dots &= \dots \end{aligned}$$

La première égalité montre que d_s et d'_s sont associés, puis la seconde montre que d_{s-1} et d'_{s-1} sont associés, et ainsi de suite. □

Remarques.

- Il faut absolument présenter ce développement comme un algorithme, et ne pas présenter d'étape qui soit abstraite. Il faut qu'on comprenne que l'on est capable de calculer cette forme normale à la main si l'on veut.
- Dans la même lignée, il faut absolument savoir exécuter cet algorithme sur un exemple simple (disons, une matrice 3×3). Sinon, on présente un truc qu'on ne connaît pas bien.
- C'est mieux, quand on met ce développement dans une leçon, de l'accompagner de quelques unes au moins de ses innombrables et monstrueuses conséquences. Classification des groupes abéliens de type fini et de manière générale, des modules de type fini sur un anneau principal, réduction de Frobenius, théorème de Jordan, théorème de Cayley-Hamilton, ... Faire une sous-partie ou une partie à propos des modules de type fini sur les anneaux principaux (ou euclidiens) montre que l'on a vraiment compris l'essence de ce théorème.

Recasages.

- 122 : Parfait, pour faire une partie sur les modules de type fini sur les anneaux principaux notamment. Il peut convenir tout seul, mais c'est vraiment moins bien.
- 142 : Au final, on fait tellement de divisions euclidiennes qu'on peut se demander si l'on n'a pas calculé plein de pgcds. La réponse est si bien sûr, puisque le coefficient en haut à gauche est le pgcd des coefficients de M. Et de manière générale le produit des k premiers coefficients est le pgcd des mineurs de taille k de M, c'est d'ailleurs ce que l'on démontre quand on montre l'unicité des facteurs invariants. D'ailleurs, ça permet de trouver les facteurs invariants d'une matrice 2×2 sans avoir à réfléchir ni à exécuter l'algorithme.
- 162 : La forme normale de Smith peut être utile pour résoudre des systèmes $\mathbf{Z}$-linéaires. C'est exactement l'analogue du pivot de Gauss sur des anneaux euclidiens au lieu des corps, et donc ça résout le même genre de problèmes. En plus il y a *opérations élémentaires* dans le titre de la leçon.

2.1.10 Formule de Burnside et coloriages du cube

Leçons 101, 105, 190

Référence H2G2 tome 1

2.1.11 Frobenius-Zolotarev

Leçons 105, 106, 123, 152

Référence Objectif agrégation

2.1.12 Indicateur de Schur-Frobenius

Leçons 101, 154, 158, 159, 170

Référence

Prérequis. Théorie des représentations et des caractères, lemme de Schur, diagonalisation des matrices hermitiennes.

Théorème 2.1.17. *Soit $\rho : G \to \mathrm{GL}(V)$ une représentation linéaire irréductible d'un groupe fini G sur $\mathbf{C}$, de caractère χ. Pour que χ soit à valeurs réelles (resp. pour que ρ se réalise sur $\mathbf{R}$), il faut et il suffit qu'il existe une forme bilinéaire (resp. forme bilinéaire symétrique) non dégénérée sur V, invariante par G. De plus, ρ se réalise sur $\mathbf{R}$ si et seulement si l'*indicateur de Frobenius-Schur *vaut* 1 :

$$\frac{1}{|G|}\sum_{g\in G}\chi(g^2) = 1.$$

Démonstration. On rappelle qu'avec la représentation ρ vient automatiquement la représentation duale V^*, et que son caractère χ^* est donné par :

$$\chi^*(g) = \chi(g)^* = \chi(g^{-1}).$$

Ainsi, χ est à valeurs réelles si et seulement si $\chi = \chi^*$, si et seulement si V et V^* sont des représentations isomorphes. Mais $\mathrm{Hom}(V, V^*) \cong \mathrm{Hom}(V \otimes V, \mathbf{C})$, et l'existence d'un isomorphisme est donc équivalente à l'existence d'une forme bilinéaire G-invariante non nulle sur V, laquelle est alors nécessairement non dégénérée.

Supposons maintenant que ρ est réalisable sur $\mathbf{R}$. Cela veut dire que l'on peut trouver une base $(e_1, \ldots, e_n)$ de V dans laquelle la matrice de chaque $\rho(g)$ est à coefficients réels. On pose alors $V_0 = \bigoplus_i \mathbf{R}e_i$ de sorte que $V = V_0 \oplus \mathrm{i}V_0$. On choisit par exemple β le produit scalaire qui rend la base (e_i) orthonormée, et l'on pose comme d'habitude :

$$\beta' = \frac{1}{|G|}\sum_{g\in G} g\cdot\beta$$

qui définit maintenant une forme bilinéaire symétrique G-invariante sur V_0. On l'étend à tout $V = V \otimes_{\mathbf{R}} \mathbf{C}$ par extension des scalaires, définissant une forme bilinéaire symétrique non dégénérée et G-invariante sur V.

Réciproquement, supposons que B est une forme bilinéaire symétrique non dégénérée et G-invariante sur V. En moyennant comme ci-dessus, on peut fabriquer un produit hermitien invariant par G que l'on note $\langle -, - \rangle$. Par théorème de Riesz, il existe une unique application $\phi : V \to V$ anti-linéaire et bijective telle que pour $x, y \in V$:

$$B(x, y) = \langle \phi(x), y\rangle^*.$$

Pour $x, y \in V$ on a alors :

$$\langle \phi^2(x), y\rangle = B(\phi(x), y)^* = B(y, \phi(x))^* = \langle \phi(y), \phi(x)\rangle$$

et de manière symétrique, $\langle \phi(x), \phi(y) \rangle = \langle \phi^2(y), x \rangle$ ce qui montre que $\langle \phi^2(x), y \rangle = \langle x, \phi^2(y) \rangle$. Ainsi ϕ^2 est un automorphisme hermitien de V et comme $\langle \phi^2(x), x \rangle = \langle \phi(x), \phi(x) \rangle$, il est positif.

On peut alors lui trouver une racine carrée v qui est un polynôme à coefficients réels en ϕ (diagonaliser ϕ^2 et prendre le polynôme d'interpolation de Lagrange pour les valeurs propres en envoyant λ_i sur $\sqrt{\lambda_i}$). On pose alors $\sigma = \phi v^{-1}$. Comme v est un polynôme en ϕ ils commutent, donc σ^2 est l'identité. On décompose alors $V = V_+ \oplus V_-$ selon les valeurs propres $+1$ et -1 de σ.

Comme σ est anti-linéaire, la multiplication par i envoie V_+ sur V_-, donc $V = V_+ \oplus \mathrm{i}V_+$. Comme $B(-,-)$ et $\langle -, - \rangle$ sont G-invariants, les applications ϕ, v et σ commutent avec les $\rho(g)$, donc V_+ et V_- sont G-stables. Cela donne une réalisation de ρ sur $\mathbf{R}$.

Montrons la dernière équivalence. L'homomorphisme défini par $\theta(x \otimes y) = y \otimes x$ est un automorphisme de la représentation $V \otimes V$, de carré l'identité. Ainsi il décompose cette représentation en :

$$V \otimes V = \mathrm{Sym}^2(V) \oplus \mathrm{Alt}^2(V),$$

les facteurs étant respectivement le *carré symétrique* et le *carré alterné* de V. Le caractère χ_σ de $\mathrm{Sym}^2(V)$ (resp. le caractère χ_α de $\mathrm{Alt}^2(V)$) est donné par :

$$\chi_\sigma(g) = \frac{1}{2}(\chi(g)^2 + \chi(g^2))$$
$$\chi_\alpha(g) = \frac{1}{2}(\chi(g)^2 - \chi(g^2))$$

et l'on vérifie bien que $\chi_\sigma + \chi_\alpha = \chi$. Le nombre a_σ (resp. a_α) de fois que le carré symétrique (resp. alterné) de ρ contient la représentation triviale est alors :

$$a_\sigma = \langle 1, \chi_\sigma \rangle \qquad a_\alpha = \langle 1, \chi_\alpha \rangle.$$

Mais le dual de $\mathrm{Sym}^2(V)$ est (comme déjà utilisé plus haut) l'espace des formes bilinéaires symétriques sur V, donc il existe une forme bilinéaire symétrique G-invariante si et seulement si $\mathrm{Sym}^2(V)^*$ n'est pas triviale, si et seulement si $a_\sigma = 1$ et $a_\alpha = 0$, ce qui donne le résultat puisque l'indicateur de Frobenius-Schur vaut $a_\sigma - a_\alpha$. □

Remarques.

- Tout faire risque d'être un peu lourd. Le mieux est de caser la définition du carré symétrique et du carré alterné plus tôt dans le plan, et d'utiliser le calcul de leurs caractères sans réfléchir. Pour que ça rentre en quinze minutes, il faut probablement passer sur des détails (comme la technique de moyenner pour obtenir des objets G-invariants).
- Il faut être au clair sur le lemme de Schur : si V et V' sont deux représentations irréductibles alors $\mathrm{Hom}(V, V')$ est soit isomorphe à $\mathbf{C}$ soit nul, selon que V et V' sont des représentations isomorphes ou non. On utilise ce lemme à plusieurs reprises.

— Le calcul des caractères χ_σ et χ_α se fait comme suit. Soit $g \in G$. On peut alors choisir une base (e_i) de V qui diagonalise $\rho(g)$, avec les valeurs propres (λ_i). On a alors $\chi(g) = \sum_i \lambda_i$ et $\chi(g^2) = \sum_i \lambda_i^2$. D'autre part :

$$(\rho(g) \otimes \rho(g))(e_i \otimes e_j + e_j \otimes e_i) = \lambda_i \lambda_j (e_i \otimes e_j + e_j \otimes e_i)$$

d'où :

$$\chi_\sigma(g) = \sum_{i \leqslant j} \lambda_i \lambda_j = \frac{1}{2} \left((\sum_i \lambda_i)^2 + \sum_i \lambda_i^2 \right)$$

et de même pour χ_α en remplaçant les $+$ par des $-$ aux bons endroits.

Recasages.

— 101 : Si l'on fait une partie sur les représentations, sinon c'est complètement hors sujet. Mais c'est un développement de très haut niveau, alors si l'on veut parler de représentations dans cette leçon, ça place la barre très haut directement.

— 154 : La question des sous-espaces stables en théorie des représentations est vraiment centrale, et ici la réalisation d'une représentation complexe sur $\mathbf{R}$ en est un cas. On regarde quand est-ce qu'il existe un sous-R-espace vectoriel de V qui est stable par les $\rho(g)$, donc ça rentre bien dans la leçon. Attention à bien gérer la théorie des représentations dans cette leçon, elle ne doit pas y être omniprésente je pense.

— 158 : On utilise la diagonalisation des matrices hermitiennes, et l'existence d'une racine carrée quand la matrice est positive (qui est facile, et on n'est pas obligé de l'expliquer à l'oral).

— 159 : On utilise à fond la dualité, qui est vraiment pratique dans la théorie des représentations, parce que ça fait apparaître des Hom et donc des conditions très fortes. Et puis le développement parle énormément de formes bilinéaires, ce qui revient au même.

— 170 : Ça commence à être limite mais c'est justifiable. Il faut un peu insister sur le fait que la forme bilinéaire symétrique est alors la forme polaire d'une forme quadratique, et que le fait qu'elle soit G-invariante revient à dire que la représentation se réalise dans le groupe orthogonal de cette forme quadratique.

2.1.13 Irréductibilité des polynômes cyclotomiques

Leçons 102, 122, 141, 144

Référence Perrin

Prérequis. Les polynômes cyclotomiques sont à coefficients entiers ; factorialité de $\mathbf{Z}[X]$.

Théorème 2.1.18. *Le polynôme cyclotomique Φ_n est irréductible sur $\mathbf{Q}$ (donc aussi sur $\mathbf{Z}$).*

Démonstration. Soient ζ une racine primitive n-ème de l'unité de p un nombre premier ne divisant pas n. Comme p et n sont premiers entre eux, ζ^p est aussi une racine primitive n-ème de l'unité.

Par factorialité de $\mathbf{Z}[X]$, on peut écrire $\Phi_n = f_1^{\alpha_1} \cdot \ldots \cdot f_r^{\alpha_r}$ avec les f_i irréductibles distincts. Φ_n est unitaire donc chaque f_i l'est aussi, et ζ est une racine de l'un d'eux, disons f_i qui est alors égal à μ_ζ. De même, ζ^p est racine de l'un d'eux, disons f_j qui est alors égal à μ_{ζ^p}. En particulier, les polynômes minimaux μ_ζ et μ_{ζ^p} sont à coefficients entiers. On va montrer que ces deux polynômes sont égaux.

S'ils sont différents, alors leur produit divise Φ_n. En fait ζ est aussi racine de $\mu_{\zeta^p}(X^p)$ donc μ_ζ le divise, dans $\mathbf{Q}[X]$ donc aussi dans $\mathbf{Z}[X]$. On écrit alors :

$$\mu_{\zeta^p}(X^p) = \mu_\zeta h, \text{ avec } h \in \mathbf{Z}[X].$$

On réduit cette égalité modulo p, ce qui avec le morphisme de Frobenius donne :

$$\overline{\mu_{\zeta^p}}^p = \overline{\mu_\zeta}\,\overline{h}.$$

Si φ est un facteur irréductible de $\overline{\mu_\zeta}$ dans $\mathbf{F}_p[X]$, il divise alors aussi $\overline{\mu_{\zeta^p}}$, et donc d'après la remarque au début du paragraphe, φ^2 divise $\overline{\Phi_n}$. Dans un corps de rupture de φ, le polynôme $\overline{\Phi_n}$ a donc une racine double... ce qui est impossible car $X^n - 1$ est à racines simples dans les corps de caractéristique p (sa dérivée s'annule uniquement en 0 qui n'est pas racine). C'est absurde.

Donc $\mu_\zeta = \mu_{\zeta^p}$. Toutes les racines primitives n-èmes de l'unité sont des puissances de ζ d'exposant premier avec n, donc elles ont toutes le même polynôme minimal, qui doit alors être multiple de Φ_n, donc leur polynôme minimal est Φ_n. En particulier, Φ_n est irréductible dans $\mathbf{Q}[X]$. Comme il est unitaire, il est donc aussi irréductible dans $\mathbf{Z}[X]$. □

Remarques.

— Il faut se souvenir de pourquoi les polynômes cyclotomiques sont à coefficients entiers. On le fait par récurrence avec la formule $X^n - 1 = \prod_{d|n} \Phi_d$, avec une division euclidienne licite dans $\mathbf{Z}[X]$ car le diviseur est unitaire.

— Il faut savoir expliquer pourquoi l'irréductibilité d'un polynôme unitaire dans $\mathbf{Q}[X]$ entraîne son irréductibilité dans $\mathbf{Z}[X]$. C'est à cause du lemme de Gauss sur le contenu. Ce même lemme de Gauss est celui qui permet de montrer que $\mathbf{Z}[X]$ est un anneau factoriel, ce que l'on utilise de manière cruciale au début du développement. De manière générale, si l'on expose ce développement il faut s'attendre à des questions sur les polynômes irréductibles dans $A[X]$ en fonction des irréductibles dans $\mathrm{Frac}(A)[X]$, lorsque A est un anneau factoriel. Ce sont les constantes irréductibles dans A, et les polynômes primitifs qui sont irréductibles dans $\mathrm{Frac}(A)[X]$.
— Une application de ce théorème est le théorème de Gauss-Wantzel qui énonce qu'un polygone régulier à n côtés est constructible à la règle et au compas si et seulement si n est une puissance de deux multipliée par un produit de nombres premiers de Fermat distincts.
— C'est assez important pour les questions et le recul de savoir que ce n'est pas du tout comme ça que ça se passe sur les corps finis. Si le développement est trop court, on pourra même expliquer ce qui se passe. Par exemple, $\Phi_8 = 1 + X^4$ est réductible sur tous les corps finis. De manière générale, Φ_n se décompose en $\varphi(n)/r$ facteurs irréductibles dans $\mathbf{F}_q[X]$, chacun de degré r égal à l'ordre de q dans $(\mathbf{Z}/n\mathbf{Z})^\times$. La démonstration est relativement courte et peut se rajouter ici (cf. Demazure, page 217).

Recasages.

— 102 : Les polynômes cyclotomiques sont par définition fortement reliés aux racines de l'unité. On peut de cette manière enchaîner ce développement avec le théorème de Gauss-Wantzel.
— 122 : On utilise des techniques vraies dans tous les anneaux factoriels, mais les théorèmes qui concernent le contenu des polynômes sont trop importants pour ne pas être mis dans cette leçon. Ainsi, on pourrait en profiter pour y mettre ce développement. Mais ce n'est bien sûr pas la meilleure leçon.
— 141 : Rien de mieux. On montre l'irréductibilité d'une classe importante de polynômes.
— 144 : Les relations entre les racines de Φ_n permettent de montrer qu'il est irréductible. Ce n'est pas forcément la meilleure leçon pour le développement, mais on peut l'imaginer quand même, avec d'autres résultats à propos des racines de l'unité vues comme des racines de polynômes cyclotomiques.

2.1.14 Jordan-Dunford-Chevalley algorithmique

Leçons 148, 153, 154, 155, 156, 157

Référence NH2G2 tome 1

Prérequis. Aucun.

Théorème 2.1.19. *Soient k un corps algébriquement clos de caractéristique nulle et A une matrice carrée de taille n à coefficients dans k. On note χ son polynôme caractéristique et $P = \chi/\operatorname{pgcd}(\chi, \chi')$ la version sans facteur carré de χ. Alors la suite (A_r) définie par :*

$$A_0 = A \text{ et } A_{r+1} = A_r - P(A_r)P'(A_r)^{-1}$$

est bien définie et elle stationne dès $r = \lceil \log_2(n) \rceil$. La limite $A_\infty = D$ est la partie diagonalisable de la décomposition de Dunford $A = D + N$.

Démonstration. On montre par récurrence sur r les trois points suivants simultanément :

1. $P(A_r)$ est nilpotente d'indice $\nu_r \leqslant 1 + (n-1)2^{-r}$;
2. $P'(A_r)$ est inversible ;
3. A_{r+1} est un polynôme en A.

Initialisation. Les racines de P sont les racines de χ, donc χ divise P^n, et par Cayley-Hamilton, $P(A)^n = 0$. Donc 1. est vrai pour $r = 0$. Ensuite, les valeurs propres de $P'(A)$ sont les $P'(\lambda)$ avec λ valeur propre de A, donc ne sont pas nulles (P est sans facteur carré). Donc 2. est vrai pour $r = 0$. Enfin 3. est vrai parce que $P'(A)^{-1}$ est un polynôme en A (l'inverse d'une matrice est toujours un polynôme en la matrice).

Hérédité. Supposons 1., 2. et 3. pour un certain r. Il existe $Q \in k[X, Y]$ tel que :

$$P(X + Y) = P(X) + YP'(X) + Y^2Q(X, Y).$$

On évalue alors la formule de Taylor ci-dessus en $X = A_r$ et $Y = -P(A_r)P'(A_r)^{-1}$:

$$\begin{aligned} P(A_{r+1}) &= P(A_r) - P(A_r)P'(A_r)^{-1}P'(A_r) + P(A_r)^2P'(A_r)^{-2}Q(A_r, -P(A_r)P'(A_r)^{-1}) \\ &= P(A_r)^2P'(A_r)^{-2}Q(A_r, -P(A_r)P'(A_r)^{-1}). \end{aligned}$$

Comme ces matrices commutent, $P(A_{r+1})$ est bien nilpotente d'indice :

$$\nu_{r+1} \leqslant \frac{\nu_r + 1}{2}$$

d'où 1. au rang $r + 1$.

Comme P est à racines simples il est premier avec P' et Bézout donne $UP + VP' = 1$. En évaluant en A_{r+1} on obtient $V(A_{r+1})P'(A_{r+1}) = I_n - U(A_{r+1})P(A_{r+1})$. Comme toutes ces matrices commutent et comme $U(A_{r+1})P(A_{r+1})$ est nilpotente par 1., la

matrice $V(A_{r+1})P'(A_{r+1})$ n'a que 1 comme valeur propre donc $P'(A_{r+1})$ est inversible, d'où 2. au rang $r+1$.

Reste à montrer 3. au rang $r+1$, qui est immédiat avec 3. au rang r car l'inverse d'une matrice est un polynôme en cette matrice.

Conclusion. Maintenant que 1., 2. et 3. sont démontrés par récurrence, on déduit de 2. que la suite est bien définie, et de 1. que pour $r \geqslant \log_2 n$, la matrice $P(A_r)$ est nilpotente d'ordre 1. Elle est donc nulle, et donc $A_{r+1} = A_r$. Comme P est scindé à racines simples, A_r est diagonalisable. Enfin, on peut écrire :

$$A - A_r = P(A_0)P'(A_0)^{-1} + \cdots + P(A_{r-1})P'(A_{r-1})^{-1},$$

somme dans laquelle chaque terme est nilpotent et où tout commute (tout est un polynôme en A). Donc $A - A_r$ est nilpotente et c'est aussi un polynôme en A, donc A_r et $A - A_r$ commutent. On a bien obtenu la décomposition de Jordan-Chevalley-Dunford de A. □

Remarques.

— C'est vraiment une bonne idée de motiver chaque argument par son analogue analytique. On cherche à récupérer D dans la décomposition $A = D + N$. Comme N est nilpotente, on doit y penser comme à un élément très petit, et donc imaginer que A est déjà très proche de la solution D. Ensuite, D est censée être solution de $P(D) = 0$ donc on peut imaginer adapter une méthode de Newton comme on fait d'habitude en analyse numérique : c'est exactement ce qui définit la suite (A_r). Pour démontrer que $P(A_r)$ est nilpotente d'indice de plus en plus petit par récurrence, on utilise une formule de Taylor : c'est normal, c'est ce que l'on ferait en analyse pour montrer que la méthode de Newton s'approche bien (et rapidement) de la solution. Donc toutes les étapes ici sont complètement naturelles si on sait faire les parallèles analytiques.

— On se place ici sur un corps algébriquement clos et de caractéristique nulle. On demande à k d'être algébriquement clos pour ne pas avoir à se soucier du scindage des polynômes, mais bien sûr la méthode marche encore sur un corps algébriquement clos, tant que le polynôme caractéristique est scindé. Et même s'il n'est pas scindé, ça marche, mais on obtient un truc qui n'est pas forcément diagonalisable (mais qui l'est dans une clôture algébrique). On demande à k d'être de caractéristique nulle pour ne pas avoir de problème du type $P' = 0$ alors que P serait non nul. On peut alors remplacer la condition de caractéristique nulle par la perfection du corps.

— Il faut savoir démontrer que si M est une matrice inversible alors M^{-1} est un polynôme en M. Par exemple, $X \mapsto MX$ est un automorphisme de $k[M]$ car il est injectif en dimension finie. Ainsi, l'identité qui est dans $k[M]$ doit être l'image par cet automorphisme d'une certaine matrice $X \in k[M]$...

— Il est bon de connaître la complexité de cet algorithme. Avec les pivots de Gauss, l'algorithme d'Euclide et la borne sur l'indice où la suite stationne, on doit avoir un $O(n^4 \log n)$ opérations.

Recasages.

— 148 : La décomposition de Jordan-Chevalley-Dunford est centrale dans la leçon, alors sa démonstration, surtout si elle est algorithmique et fait des liens avec l'analyse, est vraiment bienvenue.

— 153 : On utilise dans tous les sens des polynômes de matrices, et on utilise aussi plusieurs fois le fait que l'inverse d'une matrice est un polynôme en cette matrice. Et c'est pas seulement dans la démonstration, le fait que les parties diagonalisable et nilpotentes de la décomposition soient des polynômes en la matrice a des conséquences importantes (par exemple, on l'utilise pour démontrer le critère de Klarès).

— 154 : Pas la meilleure leçon, parce que peu d'espaces stables apparaissent vraiment. Mais on peut recaser si on en a vraiment besoin, en insistant lourdement sur les arguments théoriques sur les valeurs propres par exemple.

— 155 : Pas besoin de dire pourquoi ce développement rentre dans cette leçon, on fabrique une matrice diagonalisable importante à partir de n'importe quelle matrice à coefficients dans un corps algébriquement clos.

— 156 : On peut y mettre ce développement car il permet de calculer en pratique l'exponentielle d'une matrice dont le polynôme caractéristique est scindé. On calcule rapidement la décomposition, et l'exponentielle devient aussi facile à calculer que du beurre à étaler.

— 157 : Parfait pour le côté nilpotent, surtout qu'on en parle beaucoup pendant le développement. Le lien avec la trigonalisabilité est assez vague, on peut l'oublier.

2.1.15 L'exponentielle sur les matrices symétriques est un homéomorphisme

Leçons 156, 158, 160

Référence H2G2 tome 1

Prérequis. Théorème spectral.

Théorème 2.1.20. *L'exponentielle matricielle* $\exp : \mathcal{S}_n(\mathbf{R}) \to \mathcal{S}_n^{++}(\mathbf{R})$ *est un homéomorphisme.*

Démonstration. Soit $S \in \mathcal{S}_n(\mathbf{R})$, que l'on diagonalise en $S = P\mathrm{diag}(\lambda_i)P^{-1}$ avec P orthogonale. On a alors $\exp(S) = P\mathrm{diag}(\exp(\lambda_i))P^{-1} \in \mathcal{S}_n^{++}(\mathbf{R})$, donc l'application dans l'énoncé est bien définie. Et bien sûr, elle est continue.

Surjectivité. Soit $B \in \mathcal{S}_n^{++}(\mathbf{R})$. On la diagonalise en $B = P\mathrm{diag}(\mu_i)P^{-1}$ avec $\mu_i > 0$, et alors $A = P\mathrm{diag}(\ln \mu_i)P^{-1}$ vérifie $\exp A = B$.

Injectivité. Soient $A, A' \in \mathcal{S}_n(\mathbf{R})$ telles que $\exp A = \exp A'$. On note Q le polynôme interpolateur de Lagrange qui envoie les valeurs propres $\exp \lambda_i$ de $\exp A$ sur les λ_i. Alors A' commute avec $Q(\exp A') = Q(\exp A) = A$. On peut alors diagonaliser A et A' dans une même base pour vérifier que $A = A'$.

Continuité de la réciproque. Soit $(B_p = \exp A_p)$ une suite de $\mathcal{S}_n^{++}(\mathbf{R})$ qui converge vers $B = \exp A$. On montre donc que $A_p \to A$.

La suite (B_p) est bornée pour $\| - \|_2$. De même pour la suite (B_p^{-1}) qui converge vers B^{-1} par continuité du passage à l'inverse. Pour $M \in \mathcal{S}_n^{++}(\mathbf{R})$, on a en fait :

$$\|M\|_2 = \sqrt{\rho(M^\top M)} = \sqrt{\rho(M^2)} = \rho(M).$$

En appliquant ceci aux B_p et à leurs inverses B_p^{-1}, on remarque que les valeurs propres des B_p sont toutes contenues dans un compact $K = [C', C] \subset]0, +\infty[$. Leurs logarithmes qui sont les valeurs propres des A_p sont alors contenues dans le compact $[\ln C', \ln C]$ de $\mathbf{R}$. Ainsi, (A_p) est bornée pour $\| - \|_2$.

Or, si $A_{\varphi(p)} \to A'$ alors en passant à l'exponentielle puis à la limite on a $B = \exp A'$ donc $A' = A$, ce qui montre que A est la seule valeur d'adhérence de la suite (A_p). Ainsi cette suite converge, vers A. □

Remarques.

— Le développement est ennuyeux à mourir. J'avais juste besoin d'un développement pour remplir mes leçons 156 et 158.

— Il faut savoir démontrer qu'une suite bornée ayant une unique valeur d'adhérence converge. Si elle ne converge pas vers la valeur d'adhérence, alors on peut construire une sous-suite qui reste loin de celle-ci, et qui reste bornée. Mais alors Bolzano et Weierstrass ne sont pas contents.
— L'utilité de ce théorème est assez limitée.
— C'est mieux de savoir démontrer que $\|M\|_2 = \sqrt{\rho(M^\top M)}$ mais ça n'intéresse personne.

Recasages.
— 156 : Évidemment.
— 158 : Pareil.
— 160 : Si on a vraiment besoin d'un développement, parce qu'il y a quand même beaucoup mieux.

2.1.16 Lie-Kolchin

Leçons 103, 106, 149, 154, 157

Référence Algèbre corporelle

Prérequis. Sous-groupe dérivé ; groupe résoluble ; cotrigonalisabilité.

Théorème 2.1.21. *Soit G un sous-groupe connexe et résoluble de $\mathrm{GL}_n(\mathbf{C})$. Alors toutes les matrices de G sont trigonalisables dans une même base.*

Démonstration. Si G est abélien, le résultat est déjà démontré parce que des matrices trigonalisables qui commutent sont cotrigonalisables. On suppose donc G non abélien, et $n \geqslant 2$: on va montrer qu'il existe un sous-espace de $\mathbf{C}^n$ non trivial stable par G.

Soit alors $m \geqslant 2$ tel que $D^m G = 1$, et posons $H = D^{m-1}G$ qui n'est pas trivial. Comme $DH = 1$, les matrices de H sont cotrigonalisables donc il existe un vecteur $x \in \mathbf{C}^n$ non nul et propre pour toutes les matrices de H.

Pour $g \in G$ et $h \in H$, comme H est distingué dans G on a $g^{-1}hg \in H$, soit λ sa valeur propre pour x. On a :

$$h(g(x)) = gg^{-1}hg(x) = g(\lambda x) = \lambda g(x)$$

donc $g(x)$ est aussi propre pour tout H (il n'est pas nul car g est inversible et x n'est pas nul).

Pour tout vecteur $y \in \mathbf{C}^n$ propre pour tout H, on pose :

$$\Lambda_y \colon \begin{array}{ccc} H & \longrightarrow & \mathbf{C} \\ h & \longmapsto & \text{la valeur propre de } h \text{ en } y. \end{array}$$

Ce qui précède montre que $\Lambda_{g(x)}(h) = \Lambda_x(g^{-1}hg)$, donc :

$$\Lambda_{-(x)}(h) \colon \begin{array}{ccc} G & \longrightarrow & \mathbf{C} \\ g & \longmapsto & \Lambda_{g(x)}(h) \end{array}$$

est une application continue pour tout $h \in H$. Son image est donc connexe et incluse dans le spectre de h donc est un singleton, et l'on en déduit que $V = \mathrm{Vect}(G \cdot x)$ est un espace (stable par G et) propre pour tout élément de H.

Il ne reste plus qu'à montrer que V n'est pas trivial. Clairement $V \neq 0$ parce que $x \in V$, supposons alors $V = \mathbf{C}^n$. Cela entraîne que H n'est composé que d'homothéties, sauf que H est aussi un sous-groupe dérivé donc engendré par des commutateurs, donc tous les éléments de H sont de déterminant 1, donc ces homothéties sont de rapport une racine de l'unité. Le sous-groupe dérivé d'un groupe connexe est toujours connexe, donc $H = 1$ ce qui est absurde.

On a donc trouvé $0 \neq V \neq \mathbf{C}^n$ un sous-espace stable par G. On en choisit un supplémentaire W, et dans une base adaptée, les éléments $g \in G$ s'écrivent :

$$g = \begin{pmatrix} g_1 & * \\ 0 & g_2 \end{pmatrix}.$$

Les applications :

$$\begin{array}{ll} G \longrightarrow \mathrm{GL}(V) & G \longrightarrow \mathrm{GL}(\mathbf{C}^n/V) \\ g \longmapsto g_1, & g \longmapsto g_2 \end{array}$$

sont des morphismes de groupes continus, donc d'image connexe et résoluble. Une récurrence montre alors que les g_1 et les g_2 se trigonalisent tous dans une même base bien choisie, et cela termine la démonstration. □

Remarques.

- Le développement cache beaucoup de choses qu'il faut absolument savoir démontrer. Si on sait les démontrer, il fait un bel effet :
  - L'image d'un groupe résoluble par un morphisme de groupes est encore résoluble :
    Soit $\phi : G \to H$ avec G résoluble. Pour tout n, on montre que $\phi(D^nG) \subset D^nH$ avec égalité si et seulement si ϕ est surjectif. On le fait par récurrence, en remarquant que $\phi([D^nG, D^nG]) = [\phi(D^nG), \phi(D^nG)]$ donc $\phi(D^{n+1}G) = D\phi(D^nG) = DD^n \operatorname{Im}(\phi) = D^{n+1} \operatorname{Im}(\phi)$.
  - L'image d'un connexe par une application continue est connexe :
    C'est le théorème des valeurs intermédiaires. Un espace X est connexe si et seulement si toute application continue $X \to \{0,1\}$ est constante (c'est juste une reformulation de la définition de la connexité). Ainsi si $f : X \to Y$ est continue, alors pour toute application continue $g : f(X) \to \{0,1\}$ on obtient une application continue $gf : X \to \{0,1\}$ qui est donc constante, et comme f est surjective sur $f(X)$, on en déduit que g est constante.
  - Le sous-groupe dérivé DG d'un groupe G est toujours distingué dans G :
    Il est même caractéristique. Si ϕ est un automorphisme de G, alors pour tous $g, h \in G$ on a $\phi([g,h]) = [\phi(g), \phi(h)]$. Ainsi DG est stable par tout automorphisme, en particulier les automorphismes intérieurs.
  - Le sous-groupe dérivé DG d'un groupe connexe G est encore connexe :
    Soit S l'ensemble des commutateurs dans G. On dispose de l'application commutateur $G \times G \to S$ qui est continue et surjective, ce qui montre que S est connexe. Ensuite, pour $m \in \mathbf{N}$, on note S_m l'ensemble des produits de m commutateurs. On dispose de l'application de produit $S^m \to S_m$ qui est aussi continue et surjective, donc S_m est connexe. Enfin, DG est la réunion de tous les S_m, qui sont tous connexes et s'intersectent tous en 1.
  - L'application Λ_y est continue :
    Soit $U \subset \mathbf{C}$ un ouvert. Alors $\Lambda_y^{-1}(U) = \{h \in H \mid \exists u \in U,\ h(y) = uy\} = \{h \in H \mid h(y) \in Uy\}$. Comme Uy est un ouvert de $\mathbf{C}^n$ et l'application $h \mapsto h(y)$ est continue, $\Lambda_y^{-1}(U)$ est bien ouvert.
  - Des matrices trigonalisables qui commutent sont trigonalisables dans une même base :
    Comme elles commutent, les espaces propres de l'une sont stables par les autres. On peut par exemple faire une récurrence en disant qu'il existe un vecteur propre pour tout le monde, puis quotienter.

— Il est aussi important de maîtriser ce qu'on entend par $g \mapsto g_1$ et $g \mapsto g_2$ à la fin. La première application ne pose pas de problème parce que V est stable par g donc g_1 est bien défini comme un endomorphisme de V, mais g_2 n'est pas un endomorphisme de W (car W n'est pas stable a priori). Par contre, comme g stabilise V, il induit un endomorphisme sur le quotient $\mathbf{C}^n/V$ (qui est isomorphe à W). C'est cet endomorphisme qui est représenté par g_2. Ce raisonnement ne pose aucun problème pour montrer que c'est un morphisme de groupe continu.
— On utilise le fait que $\mathbf{C}$ est algébriquement clos, pour dire que toutes nos matrices sont trigonalisables. Par exemple, $\mathrm{SO}(2, \mathbf{R})$ est connexe et résoluble (il est abélien, difféomorphe au cercle) mais bien sûr presque aucune matrice de rotation n'est trigonalisable sur $\mathbf{R}$.
— La démonstration est vraie si l'on considère la topologie de Zariski sur $\mathrm{GL}_n(\mathbf{C})$, ce qui est plus fort car il y a plus de connexes. Et la démonstration ne change pas.
— Même après avoir demandé sur StackExchange, je n'ai pas trouvé d'exemple de sous-groupe résoluble et connexe de $\mathrm{GL}_n(\mathbf{C})$ qui ne soit pas déjà trivialement trigonalisable. De mémoire de mon premier oral blanc, le deuxième groupe spécial orthogonal sur $\mathbf{C}$ (qui est en général remplacé par le groupe spécial unitaire) est un exemple. Par contre le théorème a pour conséquence immédiate que les représentations irréductibles de dimension finie des groupes algébriques linéaires connexes et résolubles sont toutes de dimension un. C'est aussi vrai pour les algèbres de Lie résolubles ; c'est le théorème de Lie.

Recasages.

— 103 : Pas la meilleure leçon pour ce développement, mais on y utilise le sous-groupe dérivé et la continuité de la conjugaison dans les groupes topologiques. C'est un peu lointain comme lien, mais ça peut se justifier, si on entoure le développement d'une sous-partie topologique.
— 106 : Le développement est très bien dans cette leçon. On peut par exemple le voir dans une partie à propos de la topologie dans les groupes linéaires, ou bien si l'on donne des exemples de groupes de matrices résolubles.
— 149 : J'avais surtout besoin d'un développement dans cette leçon que je déteste. Le développement s'y justifie quand même, puisque l'on explique que des matrices qui commutent sont cotrigonalisables. On passe son temps dans ce développement à chercher un sous-espace stable à coups d'applications Λ qui calculent des valeurs propres, donc pourquoi pas. C'est surtout la démonstration qui est dans le sujet, plus que l'énoncé : donc il faut bien expliquer que la démonstration utilisera ce qui est vu dans le plan.
— 154 : C'est parfait aussi, on utilise bien le raisonnement usuel par récurrence sur la dimension, qui nous permet de passer de l'existence d'un sous-espace stable à la cotrigonalisabilité.
— 157 : Pareil, c'est parfait. Quoi de mieux dans la leçon sur les endomorphismes trigonalisables que d'en trigonaliser une infinité d'un seul coup ?

2.1.17 Nombre de matrices diagonalisables sur $\mathbf{F}_q$

Leçons 123, 154, 155, 190

Référence Oraux Algèbre 1

2.1.18 Nombre de matrices nilpotentes sur $\mathbf{F}_q$

Leçons 106, 123, 148, 153, 154, 157, 190

Référence H2G2 tome 2

Prérequis. Décomposition de Fitting.

Théorème 2.1.22. *Soient q une puissance d'un nombre premier, $n \geqslant 1$ et E un $\mathbf{F}_q$-espace vectoriel de dimension n. Alors il y a $\nu_n = q^{n(n-1)}$ endomorphismes de E qui sont nilpotents.*

Démonstration. Définissons pour tout entier n :
- le q-entier $n_q = q^{n-1}(q^n - 1)$;
- la q-factorielle $n!_q = n_q \cdot (n-1)_q \cdot \ldots \cdot 2_q \cdot 1_q$;
- pour tout $k \in \{0, \ldots, n\}$, le q-coefficient binomial $\binom{n}{k}_q = \frac{n!_q}{k!_q (n-k)!_q}$;
- pour toute suite de nombres complexes $a \in \mathbf{C}^{\mathbf{N}}$, la q-série génératrice exponentielle $\Phi_a(x) = \sum_{i \geqslant 0} a_i \frac{x^i}{i!_q}$.

Soient maintenant $n \geqslant 1$ et E un $\mathbf{F}_q$-espace vectoriel de dimension n.

Combien y a-t-il de bases de E ? Il y en a :

$$\begin{aligned} |\mathrm{GL}_n(\mathbf{F}_q)| &= (q^n - 1)(q^n - q) \cdot \ldots \cdot (q^n - q^{n-1}) \\ &= (q^n - 1) \cdot q(q^{n-1} - 1) \cdot \ldots \cdot q^{n-1}(q - 1) \\ &= 1_q \cdot 2_q \cdot \ldots \cdot n_q = n!_q. \end{aligned}$$

Combien y a-t-il de décompositions $E = F \oplus G$ avec $\dim(F) = d$? Choisir une telle décomposition revient à choisir une base $(e_1, \ldots, e_n)$ de E et à choisir $F = \langle e_1, \ldots, e_d \rangle$ et $G = \langle e_{d+1}, \ldots, e_n \rangle$ (il y a donc $n!_q$ choix). Cependant il ne faut pas compter plusieurs fois la même décomposition lorsque l'on choisit une autre base de F ou de G, donc il faut diviser le nombre total par $d!_q$ et par $(n-d)!_q$. On obtient alors $\binom{n}{d}_q$ choix.

Posons maintenant $\alpha_n = |\mathrm{Aut}(\mathbf{F}_q^n)|$, $\varepsilon_n = |\mathrm{End}(\mathbf{F}_q^n)|$ et $\nu_n = |\mathrm{Nil}(\mathbf{F}_q^n)|$, définissant trois suites α, ε et ν. On va utiliser leurs q-séries génératrices exponentielles pour obtenir une expression de ν_n.

Calculons Φ_α. D'après la première étape, on a :

$$\Phi_\alpha(x) = \sum_{i \geqslant 0} |\mathrm{GL}_n(\mathbf{F}_q)| \frac{x^i}{i!_q} = \sum_{i \geqslant 0} i!_q \frac{x^i}{i!_q} = \sum_{i \geqslant 0} x^i = \frac{1}{1-x}.$$

Exprimons ε en fonction de α et ν. D'après la décomposition de Fitting, pour chaque $u \in \mathrm{End}(\mathbf{F}_q^n)$ il existe un entier $m \geqslant 0$ tel que l'on ait la décomposition :

$$E = \mathrm{Ker}(u^m) \oplus \mathrm{Im}(u^m),$$

les deux espaces étant stables par u, la restriction de u sur le noyau étant nilpotente et celle sur l'image étant inversible. Réciproquement, étant donnée une décomposition de la forme $E = F \oplus G$, un endomorphisme nilpotent sur F et un automorphisme de G se recollent en un unique endomorphisme de E. Ainsi, on dispose de la relation :

$$\varepsilon_n = \sum_{i=0}^{n} \binom{n}{i}_q \alpha_i \nu_{n-i},$$

en regroupant les décompositions selon la dimension de G.

Exprimons Φ_ε en fonction de Φ_α et Φ_ν. C'est un simple produit de Cauchy :

$$\begin{aligned}
\Phi_\varepsilon(x) &= \sum_{i=0}^{+\infty} \varepsilon_i \frac{x^i}{i!_q} = \sum_{i=0}^{+\infty} \sum_{j=0}^{i} \binom{i}{j}_q \alpha_j \nu_{i-j} \frac{x^i}{i!_q} \\
&= \sum_{i=0}^{+\infty} \sum_{j=0}^{i} \frac{\alpha_j}{j!_q} x^j \frac{\nu_{i-j}}{(i-j)!_q} x^{i-j} \\
&= \Phi_\alpha(x)\Phi_\nu(x).
\end{aligned}$$

Conclusion. D'après l'expression trouvée précédemment pour Φ_α, on obtient $\Phi_\nu(x) = (1-x)\Phi_\varepsilon(x)$, c'est-à-dire en identifiant les coefficients :

$$\frac{\nu_n}{n!_q} = \frac{\varepsilon_n}{n!_q} - \frac{\varepsilon_{n-1}}{(n-1)!_q}.$$

Autrement dit :

$$\begin{aligned}
\nu_n &= \varepsilon_n - n_q \varepsilon_{n-1} \\
&= q^{n^2} - q^{n-1}(q^n - 1)q^{(n-1)^2} \\
&= q^{n^2} - q^{n+n-1+(n-1)^2} + q^{n-1+(n-1)^2} \\
&= q^{n^2} - q^{2n-1+n^2-2n+1} + q^{n(n-1)} \\
&= q^{n(n-1)}.
\end{aligned}$$

□

Remarques.

— Les séries génératrices sont des séries formelles. Il faut comprendre de quels objets on parle ! Les x écrits ici sont des indéterminées formelles, et tous les calculs sont licites.

— Il faut se rappeler que la décomposition de Fitting vient des noyaux/images itérées. Les suites $\mathrm{Ker}(u^m)$ et $\mathrm{Im}(u^m)$ stationnent à partir du même m par le théorème du rang. Encore avec le théorème du rang, pour montrer la décomposition il suffit de montrer que l'intersection est nulle. Si $x \in \mathrm{Ker}(u^m) \cap \mathrm{Im}(u^m)$ alors en notant $x = u^m(y)$, on a :

$$0 = u^m(x) = u^m(u^m(y)) = u^{2m}(y)$$

donc $y \in \mathrm{Ker}(u^{2m}) = \mathrm{Ker}(u^m)$, d'où $x = 0$. Le fait que les deux espaces soient stables par u et que la restriction au noyau soit nilpotente sont évidents. La resctriction à l'image est un automorphisme car $\mathrm{Im}(u|_{\mathrm{Im}(u^m)}) = u(\mathrm{Im}(u^m)) = \mathrm{Im}(u^{m+1}) = \mathrm{Im}(u^m)$ donc c'est surjectif.

— Il faut aussi absolument savoir expliquer pourquoi le cardinal de $\mathrm{GL}_n(\mathbf{F}_q)$ est $(q^n-1)(q^n-q)\cdot\ldots\cdot(q^n-q^{n-1})$. Choisir une matrice inversible revient à choisir des vecteurs colonnes qui sont linéairement indépendants dans $\mathbf{F}_q^n$. Ainsi le premier facteur vient du choix du premier vecteur (on peut tout choisir sauf le vecteur nul), le deuxième facteur vient du choix du deuxième vecteur (on peut tout choisir sauf ce qui est dans la droite engendrée par le premier), et ainsi de suite jusqu'à choisir le dernier dans le complémentaire d'un hyperplan.

Recasages.

— 106 : On utilise le cardinal des groupes linéaires sur les corps finis. Attention par contre, les matrices nilpotentes ne sont jamais inversibles, donc il faut que le résultat soit placé au bon endroit. Ce n'est clairement pas la meilleure leçon pour ce développement, ni le meilleur développement pour cette leçon.

— 123 : C'est très bien si l'on fait une partie sur la combinatoire des corps finis par exemple. On pourrait penser à une partie ou une sous-partie qui traite du dénombrement des ensembles de matrices ou des espaces projectifs, et alors ce développement rentre parfaitement. Sinon, ça peut vite glisser hors-sujet.

— 148 : Il faut faire très très attention, la décomposition de Fitting n'est pas une décomposition de matrices au sens usuel du terme. C'est une décomposition de l'espace sur lequel agit la matrice. On peut quand même expliquer que dans une base adaptée à la décomposition de Fitting, la matrice est diagonale par blocs, le premier bloc nilpotent et le second inversible.

— 153 : Si l'on met l'accent sur la justification de la décomposition de Fitting. Mais alors ça devient assez faible, et ça sent le recasage abusif.

— 154 : C'est très bien, encore une fois il faut bien appuyer sur le fait que c'est la décomposition de Fitting qui permet de relier ν à ε et α.

— 157 : C'est l'habitat naturel de ce développement.

— 190 : Et ça, c'est sa résidence secondaire.

2.1.19 Pavages du plan

Leçons 101, 103, 106, 160, 161, 191

Référence Géométrie (Berger)

Prérequis. Aucun.

Définition 2.1.23. Étant donnée une tuile (un compact connexe d'intérieur non vide) $P \subset \mathbf{R}^2$, un sous-groupe $G < \mathrm{Isom}^+(\mathbf{R}^2)$ est dit *paveur*, *cristallographique* ou encore *de papier peint* pour P si :
— $\mathbf{R}^2 = \bigcup_{g\in G} gP$;
— $g\mathring{P} \cap h\mathring{P} \neq \emptyset$ implique $g = h$.

Théorème 2.1.24. *Il y a cinq groupes de papier peint à conjugaison près.*

Démonstration. On commence par le lemme suivant.

Lemme 2.1.25. *On note T le sous-groupe distingué des translations ($\mathrm{Isom}^+(\mathbf{R}^2) = T \rtimes \mathrm{SO}(2,\mathbf{R})$) et $\Gamma = T \cap G$. Alors Γ est un réseau : il existe $u, v \in \mathbf{R}^2$ tel que Γ soit formé exactement des translations de vecteur $\in \mathbf{Z}u \oplus \mathbf{Z}v$.*

Démonstration. Si $\Gamma = 1$ alors G n'a que des rotations. S'il y a deux rotations de centres distincts alors leur commutateur est une translation non triviale, ce qui est impossible. Donc elles ont toutes le même centre et il est impossible de paver le plan avec P puisque $\bigcup_{g\in G} gP$ reste compact.

Si toutes les translations dans Γ étaient de directions parallèles, alors pour un élément $r \in G \setminus \Gamma$ et $t \in \Gamma$, l'élément $rtr^{-1} \in \Gamma$ serait une translation parallèle à t. Comme r est une isométrie différente de l'identité, c'est alors un retournement. Maintenant si deux retournements r et r' ont pour centres respectifs a et a', alors $r'r$ est une translation de vecteur $2\overrightarrow{aa'}$. Ainsi les centres des éléments de $G \setminus \Gamma$ (qui sont tous des retournements) sont alignés sur une droite D parallèle à la direction des translations de Γ. Encore une fois, cela contredit $\bigcup_{g\in G} gP = \mathbf{R}^2$ puisque la réunion est entièrement contenue dans une bande centrée en D.

Ainsi, Γ contient deux translations linéairement indépendantes. La difficulté à présent est de montrer l'existence d'une $\mathbf{Z}$-base à deux éléments pour Γ. On choisit dans Γ une translation non triviale de vecteur u de norme minimale. C'est possible avec le deuxième axiome de G : si (w_n) est une suite de vecteurs non nuls dont les normes tendent vers l'infimum, alors la suite des translations g_n de vecteurs $w_{n+1} - w_n$ converge simplement vers l'identité. Pour n assez grand on a alors $g_n\mathring{P} \cap \mathring{P}$ non vide, et donc g_n stationne à l'identité.

De même, on peut choisir un vecteur v de norme minimale parmi $\Gamma \setminus \mathbf{Z}u$. On montre que les vecteurs u et v conviennent. Soient a un point fixé du plan affine et Q le parallélogramme :

$$Q = \{a + tu + sv \colon t, s \in [0,1]\}.$$

Comme $\mathbf{Z}u+\mathbf{Z}v \subset \Gamma$, les images gQ avec $g \in \Gamma$ recouvrent le plan affine. Ainsi s'il existe un point de l'orbite de a sous Γ qui n'est pas dans $a+\mathbf{Z}u+\mathbf{Z}v$, alors il y en a aussi un, disons y, qui se trouve dans Q. Mais alors la distance de y à l'un des sommets de Q est strictement plus petite que $\|u\|$ ou $\|v\|$, ce qui contredit le choix de u et v. □

On peut maintenant conclure. Soit $g \in G$. Notons t_u et t_v les translations de vecteurs respectifs u et v. On a $gt_ug^{-1} = t_{\vec{g}(u)}$, de même pour v. On écrit alors :

$$\begin{cases} \vec{g}(u) = n_1u + m_1v \\ \vec{g}(v) = n_2u + m_2v \end{cases}$$

avec $n_1, n_2, m_1, m_2 \in \mathbf{Z}$. La matrice de $\vec{g}$ dans la base (u, v) est alors $\begin{pmatrix} n_1 & n_2 \\ m_1 & m_2 \end{pmatrix}$. Ainsi la trace de $\vec{g}$, qui est une rotation, est à la fois de la forme $2\cos\theta$, et un entier $n_1 + m_2$. Il ne reste alors plus que cinq cas possibles :

— si $\cos\theta = -1$ alors $\vec{g} = -\mathrm{id}$;
— si $\cos\theta = -1/2$ alors $\vec{g}$ est d'angle $2\pi/3$;
— si $\cos\theta = 0$ alors $\vec{g}$ est d'angle $\pi/2$;
— si $\cos\theta = 1/2$ alors $\vec{g}$ est d'angle $\pi/3$;
— si $\cos\theta = 1$ alors $\vec{g} = \mathrm{id}$.

L'image de G par l'homomorphisme flèche est alors cyclique engendré par une de ces cinq rotations (car c'est un sous-groupe de $\mathrm{SO}(2, \mathbf{R})$ qui est abélien, et si l'on compose deux rotations non triviales parmi ces cinq, on obtient une rotation impossible, par exemple d'angle $5\pi/6$). □

Remarques.

— C'est peut-être un peu long si l'on veut tout faire dans les détails. On peut passer un peu vite sur le fait que le deuxième axiome force l'infimum des normes à être atteint.
— C'est bien de connaître un pavage correspondant pour chacun des cinq groupes de papier peint. Ils sont dessinés dans Berger, page 13.
— C'est bien aussi de savoir que si l'on autorise les isométries non directes, alors il y a 17 groupes. Et en dimension trois, il y en a 230.

Recasages.

— 101 : Le cas des espaces affines est un cas particulier important d'actions simplement transitives. Dans le développement, on utilise plusieurs fois cette action par translations. Et le résultat concerne les groupes de papier peint, qui sont par essence des groupes faits pour agir.
— 103 : On conjugue des translations par des isométries à deux reprises, illustrant au passage le principe de conjugaison dans les actions de groupes. Et l'on utilise aussi le fait que Γ est distingué. Les pavages sont une manière concrète et visuelle de comprendre ces concepts, donc le développement illustre parfaitement la leçon.

— 106 : Au moins comme remarque dans le plan, ce développement permet de remplir par exemple une partie ou une sous-partie qui traite de l'incarnation des groupes linéaires en géométrie affine. On peut y préciser le produit semi-direct (et celui analogue pour le groupe affine par exemple), faire le lien entre la géométrie affine et la géométrie vectorielle, ...
— 160 : C'est une idée, même si ça commence à être un peu limite. Le lien ici est l'intérêt du groupe spécial orthogonal, et de l'utilisation de la trace pour classifier les rotations possibles.
— 161 : C'est parfaitement dans le thème. Il n'y a pas mieux pour parler d'isométries que de classifier les pavages !
— 191 : Les techniques d'algèbre sont assez nombreuses dans ce développement. On utilise même une technique d'analyse pour trouver une base du réseau, mais ça, il faut pas le dire.

2.1.20 Polygones réguliers constructibles

Leçons 102, 121, 125, 144, 151, 191

Référence Théorie des corps (Carrega)

Prérequis. Irréductibilité des polynômes cyclotomiques ; points et nombres constructibles. La définition suivante au grand minimum est à mettre dans le plan :

Définition 2.1.26. Un point P du plan euclidien est *constructible (à la règle et au compas)* s'il est égal à $(0,0)$, à $(1,0)$, ou s'il est intersection de deux objets distincts parmi l'ensemble des droites qui passent par deux points constructibles distincts et l'ensemble des cercles de centre un point constructible passant par un autre point constructible. Un nombre réel a est *constructible* s'il est l'abscisse d'un point constructible.

Théorème 2.1.27 (Wantzel)**.** *Un nombre réel a est constructible si et seulement s'il existe une suite finie de corps $(\mathbf{L}_i)_{0\leqslant i\leqslant n}$ tels que :*
- $\mathbf{L}_0 = \mathbf{Q}$ *;*
- $\mathbf{L}_{i+1}$ *est une extension quadratique de* $\mathbf{L}_i$ *pour tout* $i < n$ *;*
- $a \in \mathbf{L}_n$.

Démonstration. Si a est constructible, il est l'abscisse d'un point constructible M que l'on peut supposer situé sur l'axe Ox. Soit $M_1 = (0,0), M_2 = (1,0), M_3, \dots, M_n = M$ la suite des points construits pour obtenir M. Pour tout i, on note $M_i = (x_i, y_i)$. On pose alors :

$$K_i = \mathbf{Q}(x_1, y_1, \dots, x_i, y_i).$$

On a clairement $K_1 \subset \cdots \subset K_n$, $K_1 = K_2 = \mathbf{Q}$ et $a = x_n \in K_n$. Il reste à montrer que K_{i+1}/K_i est soit triviale soit de degré deux.

C'est évident pour $i = 1$. Ensuite, si M_{i+1} est l'intersection de deux droites, les nombres x_{i+1} et y_{i+1} sont solutions d'un système de la forme :

$$\begin{cases} \alpha x + \beta y + \gamma = 0 \\ \alpha' x + \beta' y + \gamma' = 0 \end{cases}$$

avec $\alpha, \beta, \gamma, \alpha', \beta', \gamma' \in K_i$. Ce système se résout dans K_i, et alors $K_{i+1} = K_i(x_{i+1}, y_{i+1}) = K_i$.

Si M_{i+1} est l'intersection d'une droite et d'un cercle, cette fois x_{i+1} et y_{i+1} sont solutions d'un système de la forme :

$$\begin{cases} \alpha x + \beta y + \gamma = 0 \\ x^2 + y^2 - 2\alpha' x - 2\beta' y + \gamma' = 0. \end{cases}$$

Si β n'est pas nul, on peut exprimer y en fonction de x et injecter dans la deuxième équation qui devient de degré deux en x. De même si α n'est pas nul.

Enfin, si M_{i+1} est l'intersection de deux cercles, le système est de la forme :

$$\begin{cases} x^2+y^2-2\alpha x-2\beta y+\gamma=0 \\ x^2+y^2-2\alpha' x-2\beta' y+\gamma'=0 \end{cases}$$

qui équivaut à :

$$\begin{cases} x^2+y^2-2\alpha x-2\beta y+\gamma=0 \\ 2(\alpha-\alpha')x+2(\beta-\beta')y-(\gamma-\gamma')=0 \end{cases}$$

et l'on se ramène au cas de l'intersection entre une droite et un cercle.

Réciproquement, supposons qu'il existe une suite de corps $\mathbf{Q}=\mathbf{L}_1\subset\mathbf{L}_2\subset\cdots\subset\mathbf{L}_p\subset\mathbf{R}$, chaque extension $\mathbf{L}_{i+1}/\mathbf{L}_i$ étant de degré deux, avec $a\in\mathbf{L}_p$. On montre par récurrence sur i que chaque élément de $\mathbf{L}_i$ est constructible.

C'est clairement vrai pour $i=1$ car tous les nombres rationnels sont constructibles. Supposons alors que $\mathbf{L}_i$ soit un corps de nombres constructibles, et montrons qu'il en est de même pour $\mathbf{L}_{i+1}$. Pour $x\in\mathbf{L}_{i+1}$, la famille $(1,x,x^2)$ est liée donc x est solution d'une équation polynomiale de degré deux de la forme $\alpha x^2+\beta x+\gamma=0$. Si α est nul alors $x\in\mathbf{L}_i$ et il n'y a rien à démontrer. Sinon, on a $x=\frac{-\beta\pm\sqrt{\beta^2-4\alpha\gamma}}{2\alpha}$ et il est alors clair que x est constructible (on sait construire des racines carrées à la règle et au compas). □

Corollaire 2.1.28. *La quadrature du cercle, la trisection de l'angle et la duplication du cube sont impossibles à réaliser.*

Démonstration. Respectivement parce que π est transcendant, parce que $\cos(\pi/9)$ est de degré trois (donc on ne peut pas trisecter $\pi/3$), et parce que $\sqrt[3]{2}$ est de degré trois. □

Théorème 2.1.29 (Gauss)**.** *Si un polygone régulier à n côtés est constructible à la règle et au compas, alors la décomposition de n en facteurs premiers est :*

$$n=2^\alpha p_1\cdot\ldots\cdot p_r$$

où $\alpha\in\mathbf{N}$ et les p_i sont des nombres premiers de Fermat (de la forme $2^{2^k}+1$).

Démonstration. On commence par un lemme.

Lemme 2.1.30. *Si m et n sont deux entiers premiers entre eux, alors l'angle $\widehat{\frac{2\pi}{mn}}$ est constructible si et seulement si les angles $\widehat{\frac{2\pi}{m}}$ et $\widehat{\frac{2\pi}{n}}$ le sont.*

Démonstration. Si $\widehat{\frac{2\pi}{mn}}$ est constructible, on peut le reporter respectivement n et m fois pour construire ses multiples que l'on cherche. Réciproquement, on écrit une relation de Bézout $\lambda n+\mu m=1$ et alors on sait facilement construire :

$$\lambda\widehat{\frac{2\pi}{m}}+\mu\widehat{\frac{2\pi}{n}}=\widehat{\frac{2\pi}{mn}}.$$

□

Ainsi, en écrivant n en produit de puissances de nombres premiers p^s, l'angle $\widehat{\frac{2\pi}{n}}$ est constructible si et seulement si chaque $\widehat{\frac{2\pi}{p^s}}$ l'est. Le cas $p = 2$ est facile à construire, soit alors p un nombre premier impair. On montre que si $\widehat{\frac{2\pi}{p^\alpha}}$ est constructible alors p est de Fermat et $\alpha = 1$. Posons $q = p^\alpha$.

Comme cet angle est constructible, le nombre $\cos\frac{2\pi}{q}$ est constructible, et d'après Wantzel on a :

$$[\mathbf{Q}(\cos\frac{2\pi}{q}) : \mathbf{Q}] = 2^m.$$

Le nombre complexe $\omega = \cos\frac{2\pi}{q} + \mathrm{i}\sin\frac{2\pi}{q}$ est en particulier algébrique sur $\mathbf{Q}$, avec $[\mathbf{Q}(\omega) : \mathbf{Q}] = \varphi(q) = p^{\alpha-1}(p-1)$.

D'autre part, ω est solution de l'équation $\omega^2 - 2\omega\cos\left(\frac{2\pi}{q}\right) + 1$ donc $[\mathbf{Q}(\omega) : \mathbf{Q}(\cos\frac{2\pi}{q})] = 2$. Finalement :

$$p^{\alpha-1}(p-1) = [\mathbf{Q}(\omega) : \mathbf{Q}] = 2^{m+1}.$$

Comme p est impair on en déduit que $\alpha = 1$, et que $p = 1 + 2^{m+1}$. Il reste à montrer que $m+1$ est une puissance de 2. On écrit alors $m + 1 = 2^\beta\lambda$ avec λ impair, et l'on remarque que :

$$p = 1 + 2^{m+1} = 1 + 2^{2^\beta\lambda} = 1 + (2^{2^\beta})^\lambda.$$

Comme λ est impair, $1 + X^\lambda$ est multiple de $1 + X$ donc p est multiple de $1 + 2^{2^\beta}$. Comme il est premier, il y a égalité et le résultat est démontré. □

Remarques.

— Tout cela est trop long à démontrer en quinze minutes. Il faut alors choisir ce que l'on fait. En général, il vaut mieux choisir en fonction de la leçon dans laquelle on place le développement. Dans les leçons sur les corps on préférera passer du temps sur le théorème de Gauss, et dans les leçons géométriques on préférera passer du temps sur celui de Wantzel.

— Il est inutile d'écrire, ni même de justifier le corollaire au milieu. Il est là pour la culture, et c'est bien de le connaître car il peut apparaître comme question du jury. En tout cas, il montre qu'on a du recul sur le théorème de Wantzel. D'ailleurs, Wantzel a démontré son théorème avant la démonstration de la transcendance de π, donc on ne savait pas à l'époque déduire du théorème de Wantzel que la quadrature du cercle était impossible. La transcendance de π a d'ailleurs été démontrée en 1882, c'est-à-dire 36 ans après la mort en 1848 de Wantzel.

— Pour la démonstration du théorème de Wantzel, si on n'a pas le temps car on veut passer au théorème de Gauss, on peut se garder d'écrire les équations en entier. On écrit juste une équation de droite et une équation de cercle, et on explique à l'oral que les solutions sont soit dans le corps de base, soit dans une extension de degré deux.

— C'est bien, pour montrer qu'on a du recul, de savoir en déduire une construction d'un polygone non trivial comme par exemple le pentagone (même si ce n'est pas

la construction optimale). Cela revient à écrire $\cos\frac{2\pi}{5}$ avec des radicaux, et à en déduire une construction.

— Dans le même ordre d'idée, il faut savoir que le théorème de Gauss est une équivalence. On n'en démontre qu'une implication, car la réciproque est trop difficile à faire à l'agreg (elle demande trop de théorie de Galois pour une leçon).

— Il faut savoir démontrer que les rationnels sont constructibles. On sait facilement tracer l'axe des abscisses à la règle puis tous les entiers avec un compas, et le théorème de Thalès nous permet de découper un segment en n parties égales, donc on a tous les rationnels.

— Il faut aussi savoir pourquoi la formule $x = \frac{-\beta\pm\sqrt{\beta^2-4\alpha\gamma}}{2\alpha}$ permet de dire que x est constructible. Les scalaires α, β et γ sont par hypothèse constructibles, donc il suffit de savoir mettre au carré, faire des multiplications, des additions et soutractions, des racines carrées et des divisions. Pour faire une multiplication (et donc une mise au carré) ou une division, on utilise le théorème de Thalès. Pour tracer la racine carrée d'un nombre constructible ξ, on trace un demi-cercle de diamètre $\xi+1$, et on trace la perpendiculaire au diamètre qui se trouve à distance 1 de l'un de ses sommets. La distance entre l'angle droit et l'intersection de la perpendiculaire avec le demi-cercle est $\sqrt{\xi}$.

— Il faut enfin savoir pourquoi $[\mathbf{Q}(\omega) : \mathbf{Q}] = \phi(q)$. C'est parce que le polynôme minimal d'une racine primitive q-ème de l'unité est Φ_q, qui est de degré $\varphi(q)$. On doit admettre ceci dans le développement, car montrer que Φ_q est le polynôme minimal de ω est à lui seul à développement, lorsque l'on montre que les polynômes cyclotomiques sont irréductibles sur $\mathbf{Q}$ (cf Perrin, page 83).

— Le théorème de Gauss est extrêmement restrictif, car l'on ne connaît aujourd'hui que cinq nombres premiers de Fermat (les cinq premiers : 3, 5, 17, 257 et 65 537). Fermat pensait que tous les nombres de la forme $1 + 2^{2^n}$ étaient premiers, mais il n'avait vérifié que les cinq premiers, et le sixième est composé. Le septième aussi, le huitième aussi, [...], le trente-troisième aussi. À partir de $1 + 2^{2^{33}}$, on ne sait pas. Le plus grand nombre de Fermat connu comme étant composé en 2013 est $1 + 2^{2^{2747497}}$. Bref, les polygones réguliers à un nombre impair de côtés étant actuellement démontrés constructibles sont donc au nombre de $2^5 = 32$.

Recasages.

— 102 : On utilise bien dans le théorème de Gauss les propriétés des racines de l'unité. C'est super si on rajoute juste avant dans le plan le fait que leur polynôme minimal est cyclotomique, et si l'on expose le théorème de Gauss comme une conséquence. Il faut un peu éclipser Wantzel ici.

— 121 : Le théorème de Gauss permet d'introduire les nombres premiers de Fermat, et la discussion sur l'existence ou non d'autres nombres de Fermat qui sont premiers. Encore une fois, il faut passer un peu Wantzel.

— 125 : C'est parfait, à la fois pour le théorème de Wantzel et pour le théorème de Gauss. C'est bien alors de faire les parties de chaque démonstration qui sont vraiment à propos d'extensions de corps, et de passer rapidement sur le reste

(mais il ne faut pas non plus trop négliger la géométrie).

— 144 : C'est limite, mais les polynômes minimaux des nombres algébriques jouent un rôle crucial ici, que ce soit dans la démonstration ou encore mieux dans la recherche d'un procédé de construction à la règle et au compas d'un polygone régulier donné. Et le théorème de Wantzel est une application assez spectaculaire de la formule pour les racines d'un polynôme de degré deux.

— 151 : On fait apparaître des conditions nécessaires pour un problème géométrique à travers des degrés d'extensions de corps, c'est-à-dire des dimensions d'espaces vectoriels. Si on appuie sur ce point, le développement est vraiment joli dans cette leçon.

— 191 : Tout est dans le titre de la leçon, c'est même difficile de faire mieux que ce développement ici.

2.1.21 Réciprocité quadratique

Leçons 120, 121, 123, 126, 190

Référence Petit compagnon des nombres

Prérequis. Symbole de Legendre.

Lemme 2.1.31 (critère d'Euler). *Soient p un nombre premier impair et $a \in \mathbf{Z}$. On a alors :*

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \ [p].$$

Démonstration. Si p divise a c'est évident. Sinon, le petit théorème de Fermat donne $a^{p-1} = 1$ dans $\mathbf{F}_p^\times$, donc $a^{(p-1)/2} = \pm 1$. Si $a = b^2$ alors $a^{(p-1)/2} = b^{p-1} = 1$, et seuls les carrés non nuls vérifient ceci puisqu'il y en a $(p-1)/2$, degré du polynôme $X^{(p-1)/2} - 1$. □

Théorème 2.1.32 (loi de réciprocité quadratique). *Pour tous nombres premiers impairs $p \neq q$, on a :*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}.$$

Démonstration. Notons $A = \{2, 4, 6, \ldots, p-1\}$ et $\overline{A}$ sa réduction modulo p. Pour $a \in A$, on note aussi r_a le reste de la division euclidienne de qa par p.

Étape 1. L'application

$$\begin{aligned} A &\longrightarrow \overline{A} \\ a &\longmapsto (-1)^{r_a}\overline{r_a} \end{aligned}$$

est bien définie et bijective. En effet, si r_a est pair alors $r_a \in A$, sinon $p - r_a \in A$ avec $p - r_a \equiv -r_a \ [p]$. Elle est injective car si $(-1)^{r_a} r_a \equiv (-1)^{r_b} r_b \ [p]$ alors $qa \equiv \pm qb \ [p]$ donc $a \equiv \pm b \ [p]$. Comme $0 < a + b < 2p$, si $a \neq b$ alors $a + b = p$. C'est impossible car p est impair. Enfin, puisque $|\overline{A}| \leqslant |A|$, l'application est bien bijective.

Étape 2. On a $\left(\frac{q}{p}\right) = (-1)^{\sum_{a\in A} r_a}$. En effet, par définition :

$$\prod_{a\in A} r_a \equiv q^{\frac{p-1}{2}} \prod_{a\in A} a \ [p]$$

et l'étape 1 montre que :

$$\prod_{a\in A} a \equiv (-1)^{\sum_{a\in A} r_a} \prod_{a\in A} r_a \ [p].$$

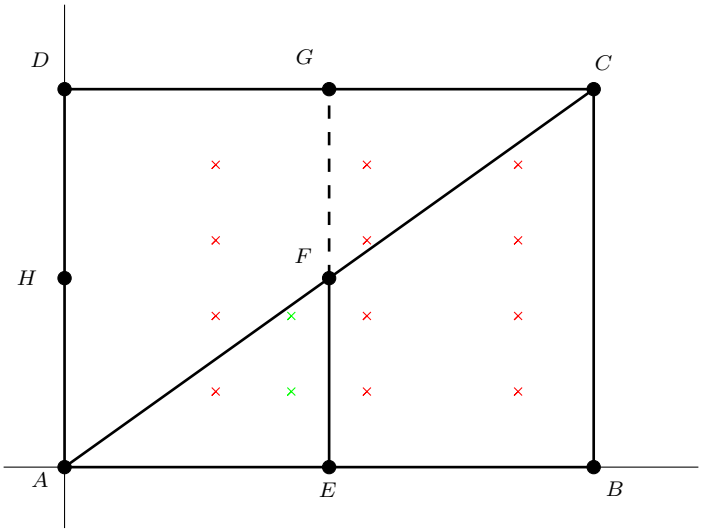
Ainsi $q^{\frac{p-1}{2}} \equiv (-1)^{\sum_{a\in A} r_a} \ [p]$, et le critère d'Euler conclut.

Étape 3. Comme r_a est le reste de la division euclidienne de qa par p, on peut écrire $qa = p\left\lfloor \frac{qa}{p} \right\rfloor + r_a$ et en déduire que :

$$\sum_{a\in A} r_a \equiv \sum_{a\in A} \left\lfloor \frac{qa}{p} \right\rfloor \ [2].$$

Il reste seulement à calculer ce membre de droite, ce que l'on fait par dénombrement.

On construit les points suivants dans $\mathbf{R}^2$: A est l'origine, $B = (p, 0)$, $C = (p, q)$ et $D = (0, q)$. E, F, G et H sont les milieux respectifs de $[AB]$, $[AC]$, $[DC]$ et $[AD]$. On s'intéresse au nombre (modulo 2) de points entiers d'abscisse paire à l'intérieur (strict) du rectangle $ABCD$.



Soit $a \in A$. Le nombre de points d'abscisse a dans $ABCD$ est pair (il y en a $q-1$). Comme p et q sont premiers entre eux, il n'y a aucun tel point sur le segment $[AC]$. Ainsi, il y a autant de points d'abscisse paire dans le triangle FCG que dans le trapèze $EBCF$, modulo 2. Comme les points d'abscisse paire de FCG correspondent bijectivement avec les points d'abscisse impaire de AEF, le nombre de points d'abscisse paire dans le grand triangle ABC est égal modulo 2 au nombre total de points dans le triangle AEF.

Ce premier nombre vaut exactement $\sum_{a\in A} \left\lfloor \frac{qa}{p} \right\rfloor$ car pour chaque abscisse paire $a \in A$, il y a bien $\left\lfloor \frac{qa}{p} \right\rfloor$ points sous le segment $[AC]$ qui est de pente q/p. En notant μ le second nombre, on a donc $\left(\frac{q}{p}\right) = (-1)^\mu$. De même symétriquement, $\left(\frac{p}{q}\right) = (-1)^\nu$ où ν est le nombre de points entiers dans le triangle AFH. Au total, $\mu + \nu = \frac{p-1}{2} \cdot \frac{q-1}{2}$ d'où le résultat. □

Remarques.

— C'est trop long si l'on démontre le critère d'Euler, alors autant laisser celui-ci dans le plan (ou l'ignorer complètement, mais le dire quand on l'utilise). Et c'est plus facile de donner les idées de la preuve que de tout écrire.

— Il faut bien s'entraîner à donner à l'oral les justifications géométriques. Elles sont faciles, mais après avoir répété cinquante fois le mot abscisse, on peut perdre le jury. Et il ne faut pas s'emmêler entre les abscisses paires ou non.
— Le dessin fait très bonne figure au tableau. En plus il permet de voir un peu magiquement le symbole de Legendre : c'est -1 exposant le nombre de points dans le petit triangle en bas à gauche. Une fois qu'on a vu ça, la réciprocité quadratique devient évidente ! C'est une bonne idée d'insister un peu sur cette visualisation.
— Il faut absolument savoir justifier pourquoi il n'y a aucun point sur $]AC[$. Un point entier non nul (x, y) sur (AC) vérifie $x/y = p/q$, donc $qx = py$, et comme p et q sont premiers entre eux, q divise y et p divise x, avec le même quotient. Donc c'est impossible que (x, y) soit sur le segment $]AC[$.
— Une fois que l'on a bien compris la démonstration, la partie difficile à retenir sont l'étape 2 et le début de l'étape 3. L'argument géométrie à la fin est le cœur de la démonstration, mais c'est aussi la partie la plus facile à retenir.
— Il faut bien retenir où apparaissent le fait que p et q doivent être impairs. Pour p, c'est quand on explique que les points d'abscisse paire de FCG sont en bijection avec les points d'abscisse impaire de AEF, par rotation centrale de centre F. C'est l'imparité de p qui fait changer la parité des abscisses. Pour q, c'est le fait que chaque colonne contienne un nombre pair de points, ce qui permet l'égalité modulo 2 entre les points d'abscisse paire dans FCG et ceux dans $EBCF$. On remarque que si l'on pose $p = 2$ ou $q = 2$, le dessin se casse la figure.

Recasages.

— 120 : L'étude des carrés dans les anneaux $\mathbf{Z}/n\mathbf{Z}$ est vraiment bien dans le thème, le théorème chinois y est obligatoire et les symboles de Legendre/Jacobi sont bienvenus. Le rapport de jury demande en plus de faire une section à propos du cas où n est premier, super nickel.
— 121 : On utilise le fait que $\mathbf{Z}/p\mathbf{Z}$ est un corps quand p est premier, on utilise le petit théorème de Fermat, tout tourne autour de nombres premiers. On peut même mettre l'accent sur le segment $]AC[$ qui ne contient aucun point entier, si l'on a vraiment peur de ce recasage.
— 123 : C'est vraiment parfait si on fait une partie sur les carrés dans les corps finis. Sinon, c'est un peu hors sujet !
— 126 : Les symboles de Legendre donnent une réponse directe à l'existence de solutions aux équations de la forme $x^2 \equiv a$ modulo n. En plus, le rapport de jury en parle en disant que c'est naturel.
— 190 : Difficile de faire mieux qu'une approche combinatoire à un théorème abstrait. On montre que l'on sait compter des choses, et l'on en déduit un joli théorème.

2.1.22 Simplicité de $\mathfrak{A}_n$ pour $n \geqslant 5$

Leçons 103, 104, 105

Référence Perrin

2.1.23 Table de caractères de $\mathfrak{S}_4$

Leçons 103, 104, 105, 160

Référence Le groupe $\mathfrak{S}_4$ et ses métamorphoses

2.1.24 Théorème de Sophie Germain

Leçons 120, 121, 126, 142

Référence Oraux Algèbre 1

Prérequis. Petit théorème de Fermat.

Théorème 2.1.33. *Soit p un nombre premier impair tel que $q = 2p + 1$ le soit aussi. Alors il n'existe aucune solution $(x, y, z) \in \mathbf{Z}^3$ à l'équation $x^p + y^p + z^p = 0$ avec $xyz \not\equiv 0$ modulo p.*

Démonstration. Par l'absurde, soit (x, y, z) une telle solution.

Montrons que l'on peut supposer les entiers premiers entre eux deux à deux. Si d est le pgcd de x, y et z alors $(x/d, y/d, z/d)$ est encore une solution, donc on peut supposer $d = 1$. Ensuite si un premier p_0 divise x et y alors il divise $x^p + y^p$ donc aussi z^p donc z, d'où p_0 divise d. Donc on peut supposer x et y premier entre eux, de même pour les deux autres paires.

Montrons que $y+z$ est une puissance p-ème. Soit p_0 un diviseur premier commun de $y + z$ et de :

$$\beta = \sum_{k=0}^{p-1} (-z)^{p-k-1} y^k.$$

Alors :

$$\begin{aligned}(y+z)\beta &= \sum_{k=0}^{p-1} \left((-z)^{p-k-1} y^{k+1} - (-z)^{p-k} y^k \right) \\ &= y^p + z^p = -x^p = (-x)^p.\end{aligned}$$

Donc p_0^2 divise $y^p + z^p = (-x)^p$, en particulier p_0 divise x. Comme $y \equiv -z$ modulo p_0, on a aussi :

$$0 \equiv \beta \equiv \sum_{k=0}^{p-1} y^{p-1} \equiv p y^{p-1} \ [p_0]$$

donc p_0 divise py^{p-1}.

Si p_0 divise p alors $p_0 = p$ donc p divise x, ce qui est supposé faux. Donc p_0 divise y^{p-1} donc y, d'où p_0 divise à la fois x, y et z : c'est absurde.

Conclusion : $y+z$ et β doivent être premiers entre eux. Comme leur produit $(-x)^p$ est une puissance p-ème, ce sont eux-mêmes des puissances p-èmes. Le même raisonnement

marche pour les autres sommes, et l'on écrit alors :

$$\begin{aligned} \beta &= \alpha^p, \\ y+z &= a^p, \\ x+z &= b^p, \\ x+y &= c^p. \end{aligned}$$

Montrons que l'un des trois nombres x, y ou z est multiple de q. Soit $m \in \mathbf{Z}$ non multiple de q. Par le petit théorème de Fermat, $m^{q-1} \equiv (m^p)^2 \equiv 1$ modulo q donc $m^p \equiv \pm 1$.

Si q ne divisait ni x, ni y, ni z, on aurait alors cette congruence pour $m = x$, y et z. En sommant, aucun cas n'amène à 0 modulo q car $q \geqslant 5$. Sans perdre de généralité, disons que c'est x qui est multiple de q.

Conclusion. On a :

$$c^p + b^p - a^p = 2x \equiv 0 \ [q]$$

et $y \equiv x + y \equiv c^p \ [q]$. Comme q ne peut pas diviser y, on a comme au paragraphe précédent $y \equiv c^p \equiv \pm 1$ modulo q. De même, $z \equiv \pm 1$.

Si q ne divisait pas a alors on aurait $a^p \equiv \pm 1$ et la somme $c^p + b^p - a^p$ ne pourrait pas valoir 0 modulo q. Donc q divise a, d'où $y + z \equiv a^p \equiv 0$ modulo q donc :

$$\beta \equiv \sum_{k=0}^{p-1} y^{p-1} \equiv p y^{p-1} \equiv p(\pm 1)^{p-1} \equiv p \ [q]$$

ce qui est absurde car $\beta = \alpha^p$ est congru soit à 0 soit à ± 1 modulo q, et certainement pas à p. □

Remarques.

- La démonstration est assez longue et technique à mémoriser. Mais en la découpant en étapes, on s'en sort mieux.
- Il faut savoir démontrer le petit théorème de Fermat : si p est un nombre premier et a n'est pas multiple de p, alors a^{p-1} est congru à 1 modulo p. Pour le démontrer, a est dans le groupe multiplicatif du corps $\mathbf{Z}/p\mathbf{Z}$, groupe qui est d'ordre $p-1$.
- C'est dommage de faire ce développement sans connaître un peu le grand théorème de Fermat. Il a été énoncé au début du dix-septième siècle, démontré par Wiles (avec l'aide de Frey, Serre, Ribet, Hellegouarch, Shimura, Taniyama, ...) en 1994, et entre-temps attaqué par à peu près tout le monde, dont Sophie Germain qui a démontré son théorème en 1823 (donc à mi-chemin, 200 ans après l'énoncé et 170 ans avant la démonstration complète). Le théorème de Sophie Germain est évidemment un cas particulier du dernier théorème de Fermat.
- Un nombre premier p tel que $q = 2p + 1$ soit aussi premier est appelé *nombre premier de Sophie Germain*. L'existence d'une infinité de tels nombres est encore une conjecture (A005384 sur l'OEIS). Les nombres premiers de Sophie Germain

peuvent être utilisés en cryptographie parce que $(\mathbf{Z}/(2p+1)\mathbf{Z})^\times$ a un sous-groupe d'ordre premier très grand.

Recasages.

— 120 : On utilise à fond l'arithmétique modulaire, et le petit théorème de Fermat qui est vraiment un théorème à propos de l'anneau $\mathbf{Z}/q\mathbf{Z}$. Le théorème de Sophie Germain est vraiment une application de l'étude de ces anneaux.

— 121 : On utilise plein de propriétés des nombres premiers, et on démontre un cas particulier du dernier théorème de Fermat, avec des liens forts avec la primalité.

— 126 : Rien à ajouter, tout est dans le titre de la leçon. C'est probablement l'équation diophantienne la plus renommée.

— 142 : Un peu limite. L'argument utilisé au début pour dire que les nombres peuvent être supposés premiers entre eux est souvent pratique, mais il n'est pas au centre du développement.

2.1.25 Théorème de Springer

Leçons 125, 141, 144, 170

Référence Carnet de voyage en Algébrie

Prérequis. Aucun.

Théorème 2.1.34. *Soient* $\mathbf{L}/\mathbf{K}$ *une extension finie de degré impair et* q *une forme quadratique sur* $\mathbf{K}^n$*. Si* q *admet un vecteur isotrope dans* $\mathbf{L}^n$ *alors elle en admet un dans* $\mathbf{K}^n$*.*

Démonstration. On démontre le résultat par récurrence sur le degré $m = [\mathbf{L} : \mathbf{K}]$. Le cas $m = 1$ ne demande aucune démonstration, supposons le résultat vrai jusqu'à $m - 2$.

Remarquons d'abord que $\mathbf{L}$ est de la forme $\mathbf{K}[\alpha_1, \dots, \alpha_k]$, et que chaque extension $\mathbf{K}[\alpha_1, \dots, \alpha_{s+1}]/\mathbf{K}[\alpha_1, \dots, \alpha_s]$ est de degré impair. Par hypothèse de récurrence, on peut alors supposer que $k = 1$, c'est-à-dire que $\mathbf{L} = \mathbf{K}[\alpha]$. On notera μ le polynôme minimal de α (qui est donc de degré m).

Soit donc $v \in \mathbf{L}^n$ un vecteur isotrope. On écrit alors $v_i = g_i(\alpha)$ avec $g_i \in \mathbf{K}[X]$ de degré au plus $m - 1$. De l'égalité :

$$0 = q(g_1(\alpha), \dots, g_n(\alpha)) \in \mathbf{K}[\alpha] \cong \mathbf{K}[X]/\mu,$$

on déduit que le reste de la division euclidienne de $q(g_1, \dots, g_n)$ par μ est nul. On peut alors écrire $q(g_1, \dots, g_n) = \mu h$ avec $h \in \mathbf{K}[X]$.

Montrons que l'on peut prendre les g_i premiers entre eux. Soit δ leur pgcd. Alors comme q est une forme quadratique :

$$\delta^2 q(g_1/\delta, \dots, g_n/\delta) = \mu h,$$

et comme $\deg \delta < m$, il est impossible que μ divise δ. Comme μ est irréductible, ces deux polynômes sont premiers entre eux, donc δ^2 divise h. On peut alors écrire :

$$q(g_1/\delta, \dots, g_n/\delta) = \mu \tilde{h}$$

avec $\tilde{h}$ de degré plus petit que h et cette fois-ci, les g_i/δ premiers entre eux.

Si $h = 0$, alors par coprimalité des g_i, pour $x \in \mathbf{K}$ il existe i tel que $g_i(x) \neq 0$ donc le vecteur $(g_1(x), \dots, g_n(x)) \in \mathbf{K}^n$ n'est pas nul et est isotrope.

Supposons alors $h \neq 0$. Comme $q(g_1, \dots, g_n)$ est de degré pair $< 2m$ (chaque g_i est de degré $< m$) et μ de degré m impair, le degré de h doit être impair et $< m$. Ainsi, h admet un facteur irréductible de degré impair h_0.

On pose alors $\mathbf{L}_0 = \mathbf{K}[X]/h_0 \cong \mathbf{K}[\beta]$ avec β la classe de X. Comme $\deg h_0 \leqslant \deg h < m$, on a $[\mathbf{L}_0 : \mathbf{K}] < [\mathbf{L} : \mathbf{K}]$. Et comme $h_0(\beta) = 0$, on a aussi :

$$q(g_1(\beta), \ldots, g_n(\beta)) = 0.$$

Si tous les $g_i(\beta)$ sont nuls alors h_0 divise g_i ce qui est impossible. Donc le vecteur $(g_1(\beta), \ldots, g_n(\beta))$ est isotrope dans une extension de degré impair $< m$: l'hypothèse de récurrence conclut. □

Remarques.

— Pour éclaircir, le schéma de la preuve est le suivant : on fait une récurrence sur le degré (impair) de $\mathbf{L}/\mathbf{K}$, l'initialisation étant évidente et l'hérédité se faisant en construisant une extension de $\mathbf{K}$ de degré impair et plus petit que celui de $\mathbf{L}/\mathbf{K}$ dans laquelle on a un vecteur isotrope. En fait, on obtient en quelque sorte un algorithme pour passer d'un vecteur isotrope sur $\mathbf{L}$ à un vecteur isotrope sur $\mathbf{K}$, sous réserve de savoir calculer α, son polynôme minimal μ, et l'isomorphisme entre $\mathbf{K}[\alpha]$ et $\mathbf{K}[X]/\mu$. Si l'on sait faire ça, on calcule facilement les polynômes g_i puis les facteur irréductibles de $q(g_1, \ldots, g_n)/\mu$. On calcule alors le vecteur isotrope $(g_1(\beta), \ldots, g_n(\beta))$ et l'on recommence jusqu'à ce que l'extension soit triviale.

— Il est important d'avoir les idées claires sur l'endroit où l'on utilise le fait que l'extension est de degré impair. C'est quand on dit que h a un facteur irréductible de degré impair, ce qui peut être faux si h est de degré pair. Il faut absolument un contre-exemple dans le cas pair. On peut prendre $\mathbf{K} = \mathbf{R}$ et $\mathbf{L} = \mathbf{C}$, la forme quadratique $x^2 + y^2$ n'a aucun vecteur isotrope sur $\mathbf{R}$ pourtant le vecteur $(1, i)$ est isotrope sur $\mathbf{C}$.

— Ce théorème est important mais au niveau de l'agrégation c'est difficile de donner une motivation compréhensible. Il démontre par exemple l'injectivité du morphisme $W(\mathbf{K} \to \mathbf{L}) : W(\mathbf{K}) \to W(\mathbf{L})$ entre les anneaux de vecteurs de Witt correspondants, ou bien est utile dans l'étude des algèbres de Jordan.

Recasages.

— 125 : C'est en plein dans le mille. On utilise des théorèmes usuels sur les extensions de corps, et on fait un pont avec d'autres leçons (puisque d'habitude, c'est rare de parler de formes quadratiques dans cette leçon).

— 141 : Un peu limite, mais ça passe parce que l'on utilise plusieurs fois l'irréductibilité de μ de manières différentes. Une fois comme polynôme minimal, une fois avec le lemme de Gauss/Euclide, on choisit une fois un facteur irréductible de h, ... Attention alors à bien motiver le théorème en tant que conséquence de quelque chose dans le sujet.

— 144 : C'est borderline. Le théorème s'inscrit dans la même famille que les propriétés du résultant ou le théorème de Gauss-Wantzel ; ici, on obtient des racines dans un petit corps à partir de racines dans un grand corps. Attention par contre à ne pas faire de hors-sujet parce que la leçon doit surtout parler de polynômes à une seule indéterminée.

— 170 : Parfait. En plus, le jury n'aime pas que l'isotropie soit oubliée : ce développement est un cadeau du ciel pour eux (et probablement que ça leur plaira assez pour qu'ils choisissent ce développement à coup sûr, rendant l'autre presque inutile à réviser pendant la préparation).

2.2 Développements d'analyse

2.2.1 Approximation de Korovkin

Leçons 201, 203, 209, 241

Référence Hirsch-Lacombe

2.2.2 Bohr-Mollerup et intégrale de log Gamma

Leçons 228, 229, 236, 239, 253, 265

Référence Rudin - Analyse réelle et complexe

Prérequis. Propriétés de la fonction Γ, caractérisations de la convexité, dérivabilité des intégrales à paramètre, intégrale de Gauss.

Théorème 2.2.1. *La fonction Γ est log-convexe et pour tout $x > 0$, on a $\Gamma(x+1) = x\Gamma(x)$.*

Démonstration. Pour la log-convexité, Γ est infiniment dérivable et strictement positive, donc $\ln\Gamma$ est bien définie et infiniment dérivable. On a comme d'habitude :

$$(\ln\Gamma)'' = \frac{\Gamma\Gamma'' - \Gamma'^2}{\Gamma^2}$$

et étudier la convexité de $\ln\Gamma$ revient à étudier le signe de $\Gamma\Gamma'' - \Gamma'^2$. On calcule alors :

$$\begin{aligned}\Gamma'(x) &= \int_0^{+\infty} \ln(t) t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t \\ &\leqslant \sqrt{\int_0^{+\infty} \ln(t)^2 t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t}\sqrt{\int_0^{+\infty} t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t} \\ &= \sqrt{\Gamma''(x)}\sqrt{\Gamma(x)}\end{aligned}$$

d'après Cauchy-Schwarz, ce qui conclut. Pour l'équation fonctionnelle, il suffit d'une intégration par parties :

$$\Gamma(x+1) = \int_0^{+\infty} t^x\mathrm{e}^{-t}\,\mathrm{d}t = \lim\left(\left[-t^x\mathrm{e}^{-t}\right]_\varepsilon^A + x\int_\varepsilon^A t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t\right) = x\Gamma(x).$$

□

Théorème 2.2.2 (Bohr-Mollerup)**.** *Si $f : \mathbf{R}_+^* \to \mathbf{R}_+^*$ est une application log-convexe telle que $f(x+1) = xf(x)$ pour tout $x > 0$, alors $f(x) = f(1)\Gamma(x)$. Autrement dit, Γ est à homothétie près la seule fonction log-convexe vérifiant cette équation fonctionnelle.*

Démonstration. Comme f et Γ vérifient l'équation fonctionnelle, f/Γ est 1-périodique. On montre le résultat pour $x \in]0,1]$. Comme $\ln f$ est convexe, on dispose de l'égalité des trois pentes, disons pour un entier $n \geqslant 1$:

$$\ln f(n+1) - \ln f(n) \leqslant \frac{\ln f(n+1+x) - \ln f(n+1)}{x} \leqslant \ln f(n+2) - \ln f(n+1)$$

ce qui se réécrit :

$$\ln n \leqslant \frac{1}{x}\ln\left(\frac{f(n+1+x)}{f(n+1)}\right) \leqslant \ln(n+1).$$

On multiplie par x et prend l'exponentielle pour avoir finalement :

$$1 \leqslant \frac{f(n+1+x)}{n^x f(n+1)} \leqslant \left(1+\frac{1}{n}\right)^x$$

ce qui montre que le membre du milieu tend vers 1 lorsque $n \to +\infty$. En utilisant une dernière fois l'équation fonctionnelle, on obtient :

$$1 = \lim \frac{f(n+1+x)}{n^x f(n+1)} = \lim \frac{(n+x)(n-1+x)\cdot \ldots \cdot xf(x)}{n^x n! f(1)} = \frac{f(x)}{f(1)} \lim(\text{truc indé. de } f)$$

ce qui est aussi vrai pour Γ. Ainsi, $f(x)/f(1) = \Gamma(x)/\Gamma(1) = \Gamma(x)$. □

Corollaire 2.2.3 (formule de duplication)**.** *Pour tout $x > 0$:*

$$2^{x-1}\Gamma\left(\frac{x}{2}\right)\Gamma\left(\frac{x+1}{2}\right) = \sqrt{\pi}\Gamma(x).$$

Démonstration. Le membre de gauche est log-convexe et vérifie l'équation fonctionnelle, et sa valeur en 1 est $\sqrt{\pi}$. □

Corollaire 2.2.4 (intégrale de $\ln\Gamma$)**.**

$$\int_0^1 \ln\Gamma = \ln\sqrt{2\pi}.$$

Démonstration. La fonction $\ln\Gamma$ est équivalente à $-\ln$ au voisinage de 0 donc l'intégrale est bien convergente. En passant au logarithme dans la formule de duplication et en intégrant, on obtient :

$$-\frac{1}{2}\ln 2 + \int_0^1 \ln\Gamma\left(\frac{x}{2}\right)\,\mathrm{d}x + \int_0^1 \ln\Gamma\left(\frac{x+1}{2}\right)\,\mathrm{d}x = \ln\sqrt{\pi} + \int_0^1 \ln\Gamma(x)\,\mathrm{d}x.$$

Avec les changements de variables évidents pour se ramener à des intégrales de $\ln\Gamma$, on retrouve :

$$\int_0^1 \ln\Gamma = \frac{1}{2}\ln 2 + \ln\sqrt{\pi}$$

ce qui est le résultat recherché. □

Remarques.

— Il faut savoir démontrer que Γ est strictement positive. On peut dire par exemple que $\Gamma(x) \geqslant \int_0^1 t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t \geqslant \mathrm{e}^{-1}\int_0^1 t^{x-1}\,\mathrm{d}t$, la première inégalité étant par restriction de l'intégrande qui est positif, et la seconde par minoration de l'exponentielle sur $[0,1]$.

— De même, il faut savoir démontrer que Γ est infiniment dérivable, et que dériver k fois fait apparaître un $(\ln t)^k$ dans l'intégrande. On fait cela avec le théorème de dérivation des intégrales à paramètres, avec la domination technique :

$$\begin{cases} \left|\left(\dfrac{\partial}{\partial x}\right)^j (t^{x-1}\mathrm{e}^{-t})\right| \leqslant |\ln t|^j t^{\varepsilon-1} & \text{si } 0 < t \leqslant 1, \\ \left|\left(\dfrac{\partial}{\partial x}\right)^j (t^{x-1}\mathrm{e}^{-t})\right| \leqslant (\ln t)^j t^{A-1}\mathrm{e}^{-t} & \text{si } t \geqslant 1. \end{cases}$$

sur tout compact $x \in [\varepsilon, A]$ de $\mathbf{R}_+^*$, et pour tout $0 \leqslant j \leqslant k$.

— L'égalité $\Gamma(1/2) = \sqrt{\pi}$ que l'on utilise pour démontrer la formule de duplication, c'est exactement l'intégrale de Gauss après le changement de variables habituel $u = \sqrt{t}$. Si l'on est dans une partie de leçon qui parle de Γ, on pourra précéder ce développement de la formule des compléments qui a cette égalité pour corollaire.

— On peut rapprocher ce théorème de celui de Wielandt qui énonce que Γ est la seule fonction holomorphe sur le demi-plan $\{\Re z > 0\}$ qui vaut 1 en 1, qui vérifie l'équation fonctionnelle, et qui est bornée sur $\{1 \leqslant \Re z \leqslant 2\}$.

Recasages.

— 228 : On peut par exemple commencer par démontrer que Γ est infiniment dérivable, ce qui montre que l'on a bien compris l'un des théorèmes les plus importants de la leçon. Ensuite, on peut aussi insister sur la démonstration de la log-convexité, et faire le lien avec les théorèmes de convexité qui peuvent rentrer dans la leçon.

— 229 : Le fait que Γ soit monotone à partir d'un certain temps n'est pas forcément intéressant et n'a rien à voir avec le développement. D'ailleurs, le minimum de Γ se trouve en :

$$\frac{3}{2}+\frac{2}{\pi^2-8}(\gamma-2+\ln 4)+\frac{8(7\zeta(3)-8)}{(\pi^2-8)^3}(\gamma-2+\ln 4)^2+\left(\frac{64(7\zeta(3)-8)^3}{(\pi^2-8)^5}-\frac{8(\pi^2-96)}{3(\pi^2-8)^4}\right)(\gamma-2+\ln 4)^4+\dots$$

ce qui vaut environ $1,46163$. Le développement sert donc uniquement du côté de la convexité, et là il utilise plusieurs résultats à propos des fonctions convexes, qu'il est bon de souligner pendant la présentation.

— 236 : Le calcul de l'intégrale de $\ln\Gamma$ est assez joli et miraculeux, et en plus on utilise le calcul de l'intégrale de Gauss. Il y a plein de manières de faire rentrer ce développement ici, mais c'est bien d'éclipser un peu le théorème de Bohr-Mollerup pour passer plus de temps sur les deux corollaires.

— 239 : On illustre bien le théorème de dérivabilité des intégrales à paramètres. Il faudrait alors le mettre comme exemple après ce théorème, ou carrément faire une sous-partie à propos de la fonction Γ.

— 253 : La convexité n'est pas à proprement parler *utilisée* dans la démonstration de Bohr-Mollerup, c'est plutôt une hypothèse du théorème. Par contre, on utilise réellement la convexité par exemple pour démontrer la formule de duplication, donc c'est bien d'insister plutôt sur ce point.

— 265 : C'est très bien si l'on veut faire l'habituelle partie sur la fonction Γ. Il faut garder en tête que ce n'est pas très très original (mais bon, on fait comme on peut).

2.2.3 Densité des polynômes orthogonaux

Leçons 201, 209, 213, 234, 239, 245, 250

Référence Objectif agrégation

Prérequis. Théorème d'holomorphie sous l'intégrale

Définition 2.2.5. Soient I un intervalle de $\mathbf{R}$. Une fonction mesurable $\rho : I \to \mathbf{R}$ strictement positive telle que $\int_I |x|^n \rho(x)\,\mathrm{d}x < +\infty$ pour tout $n \in \mathbf{N}$ s'appelle une *fonction de poids*. On obtient alors un espace de Hilbert :

$$\mathrm{L}^2_\rho = \{f : I \to \mathbf{C} \text{ mesurable} \mid \int_I |f|^2\rho < +\infty\}/(= \text{ p.p.}) = \mathrm{L}^2(I, \rho(x)\,\mathrm{d}x)$$

muni du produit hermitien $\langle f, g\rangle_\rho = \int_I f\overline{g}\rho$.

Théorème 2.2.6. *Si ρ est une fonction de poids et s'il existe $a \in \mathbf{R}$ tel que :*

$$\int_I \mathrm{e}^{a|x|}\rho(x)\,\mathrm{d}x < +\infty$$

alors les polynômes orthogonaux $(P_n)_{n\in\mathbf{N}}$, obtenus en appliquant le procédé de Gram-Schmidt à la famille $(X^n)_{n\in\mathbf{N}}$ dans L^2_ρ, est une base hilbertienne de cet espace.

Démonstration. C'est déjà une famille orthonormée, il suffit donc de montrer qu'elle est totale. Par théorème du supplémentaire orthogonal, il suffit donc de montrer que si $f \in \mathrm{L}^2_\rho$ est telle que $\langle f, P\rangle_\rho = 0$ pour tout polynôme P alors $f = 0$ dans L^2_ρ.

Posons $\varphi = f\rho\mathbf{1}_I$ définie sur $\mathbf{R}$. Alors φ est intégrable car f et $\mathbf{1}_I$ sont dans L^2_ρ. On va calculer sa transformée de Fourier et montrer qu'elle est nulle.

Posons $B_a = \{z \in \mathbf{C} \mid |\mathfrak{Im}(z)| < a/2\}$, et $g(z, x) = \mathrm{e}^{-\mathrm{i}zx}f(x)\rho(x)$ définissant g sur $B_a \times I$. Pour $z \in B_a$, on a alors :

$$\begin{aligned}\int_I |g(z,x)|\,\mathrm{d}x = \int_I \mathrm{e}^{\mathfrak{Im}(z)x}|f(x)|\rho(x)\,\mathrm{d}x &\leqslant \int_I \mathrm{e}^{a|x|/2}|f(x)|\rho(x)\,\mathrm{d}x \\ &\leqslant \sqrt{\int_I \mathrm{e}^{a|x|}\rho(x)\,\mathrm{d}x}\sqrt{\int_I |f(x)|^2\rho(x)\,\mathrm{d}x}\end{aligned}$$

qui est fini par hypothèse sur ρ et car $f \in \mathrm{L}^2_\rho$. Ainsi, $g(z, -)$ est intégrable pour tout $z \in B_a$ ce qui nous permet de considérer :

$$F(z) = \int_I g(z, x)\,\mathrm{d}x.$$

On a clairement $F|_{\mathbf{R}} = \widehat{\varphi}$, montrons alors que F est holomorphe puis nulle.

On applique le théorème d'holomorphie sous l'intégrale :

— Pour tout $z \in B_a$, la fonction $g(z, -)$ est bien mesurable (elle est même intégrable).

— Pour tout $x \in I$, la fonction $g(-,x)$ est bien holomorphe.
— Pour tous $z \in B_a$ et $x \in I$, on a la domination :

$$|g(z,x)| \leqslant \mathrm{e}^{a|x|/2}|f(x)|\rho(x)$$

comme précédemment, qui est intégrable.

Ainsi, F est holomorphe, et l'on a pour tout $n \in \mathbf{N}$:

$$F^{(n)}(z) = \int_I (-\mathrm{i}x)^n \mathrm{e}^{-\mathrm{i}zx} f(x)\rho(x)\ \mathrm{d}x$$

d'où en particulier :

$$F^{(n)}(0) = (-\mathrm{i})^n \int_I x^n f(x)\rho(x)\ \mathrm{d}x = 0$$

ce qui montre que $F = 0$ sur B_a. Ainsi $\widehat{\varphi} = 0$ puis $\varphi = 0$, d'où $f = 0$. □

Remarques.

— Il faut bien sûr connaître le procédé de Gram-Schmidt et des exemples de familles de polynômes orthogonaux. On peut par exemple citer :
  — Les polynômes de Legendre, avec $I = [-1,1]$ et $\rho = 1$;
  — Les polynômes de Tchebychev, avec $I =]-1,1[$ et $\rho(x) = \frac{1}{\sqrt{1-x^2}}$;
  — Les polynômes de Hermite, avec $I = \mathbf{R}$ et $\rho(x) = \mathrm{e}^{-x^2}$;
  — Les polynômes de Laguerre, avec $I = [0,+\infty[$ et $\rho(x) = \mathrm{e}^{-x}$.
— De même, il faut savoir montrer que L^2_ρ est un espace de Hilbert. La démonstration est exactement la même que pour montrer que L^2 est un espace de Hilbert (d'ailleurs c'est la même, puisque L^2_ρ est un espace L^2), donc elle n'est pas très intéressante.
— Il faut bien se rappeler que pour le théorème d'holomorphie sous l'intégrale, on ne demande pas de majoration de la dérivée, mais bien une majoration de la fonction elle-même. C'est bien de savoir pourquoi : on utilise le théorème de Morera pour démontrer ce théorème, mais pour obtenir l'expression de la dérivée, on applique le théorème de dérivation des intégrales à paramètres, la domination venant des inégalités de Cauchy.

Recasages.

— 201 : Le développement n'est pas le plus original pour cette leçon, mais il rentre bien dans une partie sur les espaces L^p. Attention quand même à ne pas trop étaler dessus parce que la démonstration s'écarte un peu du thème de la leçon.
— 209 : L'approximation de fonctions est très liée aux bases hilbertiennes, puisqu'une base hilbertienne fournit une approximation de tous les éléments de l'espace préhilbertien par des combinaisons linéaires de fonctions simples. Les théorèmes de densité sont souvent bienvenus dans cette leçon.
— 213 : On illustre bien le théorème du supplémentaire orthogonal (enfin, son corollaire qui est le critère de densité) qui n'est valable que dans les espaces de

Hilbert, et l'on introduit un espace de Hilbert intéressant. Malheureusement le seul raisonnement hilbertien que l'on fait est au tout début, le reste n'a plus grand chose à voir. Attention à la redondance si l'on choisit aussi le développement sur l'espace de Bergman.

— 234 : On obtient un espace de fonctions L^2 dont on sait donner une jolie base hilbertienne autre que les polynômes trigonométriques. On utilise le fait que le produit de deux fonctions de carré intégrable est intégrable, ce qui est assez dans le thème.

— 239 : C'est une très jolie illustration du théorème d'holomorphie sous l'intégrale, surtout qu'on ne s'y attend pas trop en lisant l'énoncé du théorème. Attention à la manière d'incruster le développement dans le plan par contre. Le mieux, c'est de le mettre juste après le théorème d'holomorphie sous l'intégrale.

— 245 : On utilise le théorème d'holomorphie sous l'intégrale et le fait que les fonctions holomorphes sont développables en série entière, tout ça pour démontrer un théorème qui pourrait n'avoir aucun rapport avec $\mathbf{C}$. Par contre le développement n'est pas très original, surtout qu'il y a beaucoup de développements possibles et meilleurs dans cette leçon.

— 250 : Le développement y rentre très bien parce qu'on utilise l'injectivité de la transformée de Fourier pour démontrer un résultat qui a priori n'a aucun rapport. Il se trouve juste qu'il est plus facile de montrer que la transformée de Fourier de f est nulle, que de démontrer directement que f l'est. Donc c'est une jolie application.

2.2.4 Développement asymptotique d'une suite récurrente

Leçons 223, 224, 226, 230

Référence Bernis Bernis

Prérequis. Sommation des relations de comparaison, équivalent de la série harmonique

Théorème 2.2.7. *Soient $b > 0$ et $f : [0, b] \to \mathbf{R}$ tels que :*

1. *f est continue ;*
2. *f est croissante ;*
3. *$f(0) = 0$ et pour tout $x \in]0, b]$, on a $f(x) < x$;*
4. *Il existe $\lambda > 0$ et $R > 1$ tels que lorsque $x \to 0$:*

$$f(x) = x - \lambda x^R + o(x^R).$$

Pour tout $c \in]0, b[$, la suite définie par $u_0 = c$ et $u_{n+1} = f(u_n)$ reste à valeurs dans $]0, b[$, tend vers 0 et plus précisément :

$$u_n \sim \frac{K}{n^{\frac{1}{R-1}}}, \qquad K = (\lambda(R-1))^{\frac{1}{1-R}}.$$

Démonstration. Par croissance de f, celle-ci ne peut pas s'annuler ailleurs qu'en 0 (sinon le développement asymptotique 4 serait faux). Ainsi pour tout $x \in]0, b[$, on a $0 < f(x) < x < b$, ce qui montre que la suite $(u_n)_{n \in \mathbf{N}}$ est bien définie.

D'après 3, la suite est décroissante en plus d'être positive, donc elle converge vers un $\ell \in [0, b]$, qui par continuité doit être un point fixe de f. D'après ce qui précède, cela entraîne $\ell = 0$. Il reste donc juste à démontrer l'équivalent pour u_n.

D'après 4, on a $\frac{u_{n+1} - u_n}{u_n^R} \to -\lambda$. Un dessin avec la courbe $y = x^{-R}$ suggère que ce quotient soit équivalent à $\int_{u_{n+1}}^{u_n} \frac{\mathrm{d}t}{t^R} = \frac{u_n^{1-R} - u_{n+1}^{1-R}}{1-R}$, ce que l'on va vérifier. On calcule alors :

$$\begin{aligned}
\frac{u_n^{1-R} - u_{n+1}^{1-R}}{1-R} &= \frac{1}{1-R}\left(u_n^{1-R} - (u_n - \lambda u_n^R + o(u_n^R))^{1-R}\right) \\
&= \frac{u_n^{1-R}}{1-R}(1 - (1 - \lambda \underbrace{u_n^{R-1}}_{\to 0} + o(u_n^{R-1}))^{1-R}) \\
&= \frac{u_n^{1-R}}{1-R}\left(1 - (1 - (1-R)\lambda u_n^{R-1} + o(u_n^{R-1})\right) \\
&= \lambda + o(1) \sim \lambda.
\end{aligned}$$

Comme $\sum_{n=0}^{+\infty} \lambda = +\infty$, la sommation des relations de comparaison montre que :

$$\sum_{k=0}^{n-1} \frac{u_k^{1-R} - u_{k+1}^{1-R}}{1-R} \sim \sum_{k=0}^{n-1} \lambda = n\lambda$$

ce qui par téléscopage montre que :

$$\frac{u_0^{1-R} - u_n^{1-R}}{1-R} \sim n\lambda$$

ce qui donne le résultat annoncé. □

Proposition 2.2.8. *Si l'on pose $f : x \mapsto \ln(1+x)$ et $u_0 > 0$ quelconque, alors on obtient d'après ce qui précède $u_n \sim 2/n$. On peut pousser le développement asymptotique plus loin pour obtenir :*

$$u_n = \frac{2}{n} + \frac{2\ln n}{3n^2} + o\left(\frac{\ln n}{n^2}\right).$$

Démonstration. Il suffit de reprendre la méthode précédente, avec plus de précision :

$$\begin{aligned}
\frac{u_n^{-1} - u_{n+1}^{-1}}{-1} &= u_n^{-1}\left(\frac{u_n}{u_{n+1}} - 1\right) \\
&= u_n^{-1}\left(\frac{u_n}{\ln(1+u_n)} - 1\right) \\
&= u_n^{-1}\left(\frac{1}{1 - u_n/2 + u_n^2/3 + o(u_n^2)} - 1\right) \\
&= u_n^{-1}\left(1 + \left(\frac{u_n}{2} - \frac{u_n^2}{3} + o(u_n^2)\right) + \left(\frac{u_n}{2} - \frac{u_n^2}{3} + o(u_n^2)\right)^2 + o(u_n^2) - 1\right) \\
&= u_n^{-1}\left(\frac{u_n}{2} - \frac{u_n^2}{3} + \frac{u_n^2}{4} + o(u_n^2)\right) = \frac{1}{2} - \frac{u_n}{12} + o(u_n).
\end{aligned}$$

En posant alors $v_n = u_{n+1}^{-1} - u_n^{-1} - \frac{1}{2}$, on a $v_n \sim \frac{-u_n}{12} \sim \frac{-1}{6n}$. Comme la série $\sum \frac{1}{6n}$ diverge, on obtient :

$$u_n^{-1} - u_0^{-1} - \frac{n}{2} = \sum_{k=0}^{n-1} v_k \sim \sum_{k=0}^{n-1} -\frac{1}{6k} \sim -\frac{1}{6}\ln(n).$$

Finalement on obtient $u_n = \left(\frac{n}{2} - \frac{1}{6}\ln n + o(\ln n)\right)^{-1}$ ce qui après un développement limité de $\frac{1}{1-x}$ donne le résultat annoncé. □

Remarques.

- Le développement est technique et calculatoire, mais le résultat est très général donc c'est plutôt joli. On peut aussi l'appliquer au sinus par exemple. Même sans pousser le raisonnement jusqu'au bout, on obtient l'idée à partir d'un développement asymptotique de f d'estimer les différences de la suite à une certaine puissance pour obtenir un équivalent simple.
- L'équivalence de la série harmonique avec le logarithme se fait par comparaison série-intégrale.

— On commence à voir dans la seconde proposition qu'il peut être difficile d'aller plus loin dans le développement asymptotique, car il faut s'assurer de pouvoir sommer les relations de comparaison.
— Il faut faire des dessins ! Un premier pour donner une allure du graphe de f au tout début, puis au moins un autre pour donner l'idée de comparer avec l'intégrale.

Recasages.
— 223 : C'est très bien dans cette leçon, on donne une manière d'étudier les comportements asymptotiques de toute une famille de suites numériques, et l'on illustre ses connaissances sur les relations de comparaison.
— 224 : Pas mieux, on passe littéralement du développement asymptotique d'une fonction à celui des suites qu'elle définit, et on donne même un exemple (on peut en donner beaucoup d'autres pour remplir le plan).
— 226 : Là aussi, on est en plein dans le thème, parmi les suites définies par récurrence on sait contrôler toutes celles définies par une fonction assez bien connue autour de 0.
— 230 : C'est surtout pour la sommation des relations de comparaison qui arrive plusieurs fois, ici dans le comportement des sommes partielles. On utilise ainsi la divergence de certaines séries et des équivalents de sommes partielles. Attention, l'énoncé du développement peut faire croire que c'est hors-sujet, il faut donc bien l'annoncer comme application du théorème de sommation des relations de comparaison.

2.2.5 Espace de Bergman

Leçons 201, 205, 208, 213, 234, 243, 245

Référence Bernis Bernis

Prérequis. Formule de la moyenne, théorème de Riesz-Fischer, théorème de Weierstrass, critère de densité dans les espaces de Hilbert, formule de Cauchy, théorème de Riesz, égalité de Bessel

Théorème 2.2.9. *Notons* $\mathbf{D}$ *le disque unité (ouvert) complexe, et* $\mathrm{L}^2_a(\mathbf{D}) = \mathrm{L}^2(\mathbf{D}) \cap \mathcal{H}(\mathbf{D})$ *l'espace* de Bergman *des fonctions holomorphes et de carré intégrable sur* $\mathbf{D}$*, muni du produit hermitien de* L^2*,* $\langle f, g\rangle = \int_{\mathbf{D}} f\overline{g}$*. Alors :*

1. $\mathrm{L}^2_a(\mathbf{D})$ *est un espace de Hilbert ;*
2. *La famille* $e_n : z \mapsto \sqrt{\frac{n+1}{\pi}} z^n$ *en est une base hilbertienne ;*
3. *La fonction* $K : (z, w) \mapsto \frac{1}{\pi(1-z\overline{w})^2}$ *en est un noyau reproduisant : pour tout* $f \in \mathrm{L}^2_a(\mathbf{D})$*,*
$$f = \int_{\mathbf{D}} f(w) K(-, w).$$

Démonstration. On commence par un lemme crucial qui relie les topologies des espaces $\mathrm{L}^2(\mathbf{D})$ et $\mathcal{H}(\mathbf{D})$.

Lemme 2.2.10. *Pour tout* $f \in \mathrm{L}^2_a(\mathbf{D})$ *et tout compact* $K \subset \mathbf{D}$*, on a :*
$$\|f\|_{\infty,K} \leqslant \frac{1}{\sqrt{\pi} d(K, \mathbf{S}^1)} \|f\|_2.$$

Démonstration. Soient $a \in K$ et $r > 0$ tel que $\overline{D}(a, r) \subset \mathbf{D}$. Pour tout $\rho < r$, on a la formule de la moyenne :
$$f(a) = \frac{1}{2\pi} \int_0^{2\pi} f(a + \rho \mathrm{e}^{\mathrm{i}\theta}) \, \mathrm{d}\theta$$
ce qui permet de calculer l'intégrale en polaire :
$$\begin{aligned} \int_{\overline{D}(a,r)} f &= \int_0^r \int_0^{2\pi} f(a + \rho \mathrm{e}^{\mathrm{i}\theta}) \, \mathrm{d}\theta \rho \, \mathrm{d}\rho \\ &= \int_0^r 2\pi\rho f(a) \, \mathrm{d}\rho = \pi r^2 f(a). \end{aligned}$$
Grâce à cela on peut majorer avec Cauchy-Schwarz :
$$|f(a)| \leqslant \frac{1}{\pi r^2} \sqrt{\int_{\overline{D}(a,r)} |f|^2} \sqrt{\int_{\overline{D}(a,r)} 1} \leqslant \frac{1}{\pi r^2} \|f\|_2 \sqrt{\pi} r = \frac{\|f\|_2}{\sqrt{\pi} r}.$$
On fait alors tendre r vers $1 - |a| = d(a, \mathbf{S}^1) \geqslant d(K, \mathbf{S}^1)$ pour obtenir le résultat. □

On peut maintenant contrôler les deux topologies en même temps, et démontrer le théorème.

$\mathrm{L}^2_a(\mathbf{D})$ est un espace de Hilbert. Soit $(f_n)_{n\in\mathbf{N}}$ une suite de Cauchy dans $\mathrm{L}^2_a(\mathbf{D})$, elle est de Cauchy dans $\mathrm{L}^2(\mathbf{D})$ donc elle converge vers une certaine fonction f qui est de carré intégrable. Le lemme permet alors de contrôler la norme uniforme : pour tout compact K de $\mathbf{D}$ et tous $m, n \in \mathbf{N}$,

$$\|f_m - f_n\|_{\infty,K} \leqslant \frac{1}{\sqrt{\pi} d(K, \mathbf{S}^1)} \|f_m - f_n\|_2$$

donc $(f_n)_{n\in\mathbf{N}}$ est aussi de Cauchy pour la topologie métrisable de la convergence uniforme sur tout compact, et par théorème de Weierstrass on dispose d'une fonction holomorphe g vers laquelle les f_n convergent uniformément sur tout compact. Comme f_n converge à extraction près vers f presque partout, on a $g = f$ presque partout et donc $f \in \mathrm{L}^2_a(\mathbf{D})$ ce qui conclut.

Les e_n forment une base hilbertienne. On a juste à calculer par passage en coordonnées polaires :

$$\begin{aligned}\langle e_p, e_q\rangle &= \int_{\mathbf{D}} \sqrt{\frac{p+1}{\pi}}\sqrt{\frac{q+1}{\pi}} z^p \overline{z}^q \,\mathrm{d}z \\ &= \sqrt{\frac{(p+1)(q+1)}{\pi^2}} \int_0^1 \int_0^{2\pi} \rho^{p+q+1} \mathrm{e}^{\mathrm{i}(p-q)} \,\mathrm{d}\theta \,\mathrm{d}\rho\end{aligned}$$

qui vaut 0 si $p \neq q$. Si $p = q$, alors :

$$\langle e_p, e_p\rangle = 2\pi \frac{p+1}{\pi} \int_0^1 \rho^{2p+1} \,\mathrm{d}\rho = 1.$$

Reste donc à montrer que la famille est totale. On applique le critère de densité, donc on choisit f orthogonale à tous les e_n et l'on montre que $f = 0$. Pour cela on développe f en série entière au voisinage de 0, disons $f(z) = \sum_{n=0}^{+\infty} a_n z^n$. On a alors la formule de Cauchy pour $r \in [0, 1[$:

$$a_n = \frac{1}{2\pi r^n} \int_0^{2\pi} f(r\mathrm{e}^{\mathrm{i}\theta}) \mathrm{e}^{-\mathrm{i}n\theta} \,\mathrm{d}\theta,$$

qui permet de calculer :

$$\begin{aligned}0 = \langle f, e_n\rangle &= \sqrt{\frac{n+1}{\pi}} \int_{\mathbf{D}} f(z)\overline{z}^n \,\mathrm{d}z = \sqrt{\frac{n+1}{\pi}} \int_0^1 \int_0^{2\pi} f(r\mathrm{e}^{\mathrm{i}\theta}) \mathrm{e}^{-\mathrm{i}n\theta} r^n r \,\mathrm{d}r \,\mathrm{d}\theta \\ &= \sqrt{\frac{n+1}{\pi}} \int_0^1 2\pi r^n a_n r^n r \,\mathrm{d}r = \sqrt{\frac{n+1}{\pi}} \frac{2\pi}{2n+2} a_n \\ &= \sqrt{\frac{\pi}{n+1}} a_n\end{aligned}$$

d'où $a_n = 0$ pour tout n et donc $f = 0$.

La fonction K est un noyau reproduisant. Pour $z \in \mathbf{D}$, posons :

$$\ell_z : \begin{array}{ccc} \mathrm{L}^2_a(\mathbf{D}) & \longrightarrow & \mathbf{C} \\ f & \longmapsto & f(z) \end{array}$$

la forme d'évaluation, qui est linéaire, et continue grâce au lemme. D'après le théorème de Riesz on dispose de $k_z \in \mathrm{L}^2_a(\mathbf{D})$ tel que $\ell_z = \langle -, k_z \rangle$. On note alors pour tout $z \in \mathbf{D}$:

$$K(z, -) = \overline{k_z}.$$

D'une part, l'égalité de Bessel donne :

$$K(z,-) = \overline{k_z} = \overline{\sum_{n=0}^{+\infty} \langle k_z, e_n \rangle e_n} = \sum_{n=0}^{+\infty} \langle e_n, k_z \rangle \overline{e_n} = \sum_{n=0}^{+\infty} e_n(z) \overline{e_n}$$

d'où la propriété de noyau reproduisant :

$$\int_{\mathbf{D}} f(w) K(z,w) \, \mathrm{d}w = \langle f, \overline{K(z,-)} \rangle = \langle f, k_z \rangle = f(z),$$

et d'autre part comme il y a convergence L^2 de la somme donc aussi sur tout compact d'après le lemme, on a :

$$K(z,w) = \sum_{n=0}^{+\infty} e_n(z) \overline{e_n}(w) = \sum_{n=0}^{+\infty} \frac{n+1}{\pi} z^n \overline{w}^n = \frac{1}{\pi(1 - z\overline{w})^2}.$$

□

Remarques.

- C'est très clairement trop long de tout faire dans les détails en quinze minutes. Il faut donc faire une petite recette en fonction de la leçon dans laquelle on place le développement. Attention à toujours démontrer le lemme, qui contient l'essentiel de la magie du développement. La majoration de la norme uniforme par la norme L^2 est exceptionnelle, et provient du miracle qu'est la théorie des fonctions holomorphes. On en déduit tout le reste sans trop d'efforts.
- C'est vraiment mieux si l'on fait des dessins pour expliquer le raisonnement, notamment lors de la démonstration du lemme. On peut faire un dessin du disque unité, d'un compact K, du point a et du rayon r que l'on fait tendre vers la distance de a au cercle. C'est vraiment très important pour ne pas perdre l'auditoire.
- On voit dans la démonstration que les racines carrées dans la définition des e_n sont un peu inutiles, elles servent juste à ce que $\|e_n\|_2 = 1$. C'est important pour le troisième point car on utilise l'égalité de Bessel, mais dans la démonstration du second point on peut écrire des $\sqrt{\cdots}$ pour aller plus vite. L'important est que les z^n sont orthogonales.

— Pour justifier que (e_n) est une famille totale, on pourrait vouloir écrire :

$$0 = \langle f, e_n \rangle = \langle \sum_{m=0}^{+\infty} a_m z^m, e_n \rangle = \sum_{m=0}^{+\infty} a_m \sqrt{\frac{\pi}{m+1}} \langle e_m, e_n \rangle$$

d'où $a_m = 0$ pour tout m en choisissant $m = n$, en justifiant que l'on peut sortir la somme du produit hermitien par continuité. Ce raisonnement est faux, et c'est un piège qui apparaît plusieurs fois dans le développement sans crier gare. Il faut absolument être au courant de son existence : les sommes infinies en jeu sont parfois des sommes au sens L^2, et parfois des sommes au sens de la convergence uniforme sur tout compact. Ici la somme à l'intérieur est convergente sur tout compact, mais pas nécessairement L^2 donc la continuité de $\langle -, - \rangle$ pour la topologie L^2 ne permet pas de sortir la somme. On retrouve cela quand on calcule $K(z, w)$, en disant que la somme est convergente dans L^2 donc, grâce au lemme, est convergente uniformément sur tout compact.

— Dans la même lignée, il peut être instructif de souligner (au moins en entraînement, et même peut-être à l'oral) pour chaque somme si elle est convergente dans L^2 ou uniformément sur tout compact. Plus généralement, on peut souligner chaque apparition du monde L^2 (Riesz-Fischer, Riesz, Bessel) et chaque apparition du monde holomorphe (formule de la moyenne, Weierstrass, formule de Cauchy). On observe vraiment une utilisation intensive de chacun des deux domaines, et il faut absolument savoir lequel est utilisé à quel moment pour que le développement tienne la route aux yeux du jury.

— Le théorème de Weierstrass énonce que si U est un ouvert de $\mathbf{C}$, alors l'espace $\mathcal{H}(U)$ avec la topologie de la convergence uniforme sur tout compact est métrisable, et cette métrique le rend complet. On utilise cette complétude dans le développement, donc mieux vaut avoir une idée de ce théorème.

— La question risque d'arriver : le a en indice de $\mathrm{L}^2_a(\mathbf{D})$ signifie *analytique*, c'est-à-dire que l'on ne garde de l'espace $\mathrm{L}^2(\mathbf{D})$ que les fonctions qui sont analytiques.

— On peut de même analyse les espaces de Bergman sur d'autres domaines que $\mathbf{D}$, les résultats sont sensiblement les mêmes.

— Au sujet du noyau reproduisant : c'est le noyau d'un opérateur à noyau, et un théorème d'analyse fonctionnelle dit que les opérateurs à noyau de L^2 dans lui-même sont compacts (et quand ils sont hermitiens comme K, ils sont diagonalisables). Attention à ne pas tomber dans ce piège et dire que ça s'applique : le noyau K n'est pas de carré intégrable, le théorème ne s'applique donc pas. Heureusement, parce que l'opérateur que l'on obtient est l'identité, qui ne risque pas d'être un opérateur compact en dimension infinie.

— Les espaces de Hilbert à noyau reproduisant (RKHS) comme $\mathrm{L}^2_a(\mathbf{D})$ ont été développés par Bergman et Aronszajn au milieu du vingtième siècle. Ils sont utilisés en statistiques, en analyse complexe, en mécanique quantique et en machine learning.

Recasages.

— 201 : On introduit un espace de fonctions et on utilise les propriétés de deux autres espaces, l'un qui est incontournable dans la leçon, qui est un espace de Hilbert séparable (L^2) et l'autre qui est un peu plus original mais tout aussi intéressant, complètement métrisable ($\mathcal{H}$). L'intersection des deux donne lieu à une explosion de propriétés qui proviennent de la magie de chacun des deux premiers espaces. Dans cette leçon, on peut par exemple passer rapidement sur la fin pour se concentrer sur le début du développement.
— 205 : Les deux espaces que l'on intersecte sont complets, et l'on utilise leur complétude pour montrer celle de L^2_a. On utilise ensuite la théorie des espaces de Hilbert (donc la complétude, de manière un peu plus cachée) pour le noyau reproduisant. On peut donc passer un peu vite sur la base hilbertienne.
— 208 : L'espace L^2_a est de Hilbert (donc en particulier normé), et l'on utilise des applications linéaires continues dans le dernier point du théorème. C'est pertinent dans cette leçon surtout si l'on fait une partie sur les espaces de Hilbert ou les espaces préhilbertiens. On peut encore une fois passer rapidement sur le deuxième point.
— 213 : C'est en plein dans le thème de la leçon et on utilise tout plein de théorèmes à propos des espaces de Hilbert (Riesz-Fischer L^2, le critère de densité, l'égalité de Bessel, ...) donc c'est parfait pour montrer que l'on a tout compris. La difficulté ici est que tout le développement est très pertinent, donc c'est difficile d'éclipser certaines parties. On peut par exemple essayer de tout faire rentrer en passant plus rapidement sur les calculs et en insistant sur les gros théorèmes hilbertiens que l'on utilise.
— 234 : C'est très bien juste après le théorème de Riesz-Fischer par exemple. On introduit un espace de fonctions de carré intégrable, donc c'est pertinent dans l'énoncé. Attention la démonstration n'utilise pas beaucoup de théorie de l'intégrabilité, surtout vers la fin. Il faut donc bien insister sur le lemme et sur la première partie du développement.
— 243 : C'est un bon développement pour illustrer les propriétés d'analyticité des fonctions holomorphes. Par exemple la formule intégrale de Cauchy joue un rôle assez crucial. Et tout à la fin on calcule l'expression explicite d'une somme de série entière, donc c'est pertinent. On pourra passer un peu plus rapidement sur les deux premiers points.
— 245 : Il est difficile de trouver un développement qui utilise autant de théorèmes d'analyse complexe (formule de la moyenne, théorème de Weierstrass, formule intégrale de Cauchy, analyticité, ...). Le théorème de Montel en est un autre un peu plus original, mais celui-ci est déjà très solide. On pourra passer un peu plus rapidement sur les calculs et sur le premier point pour insister plus fortement sur le lemme et les utilisations des théorèmes d'analyse complexe.

2.2.6 Formule de Poisson

Leçons 241, 246, 250, 265

Références Bernis Bernis, Gourdon Analyse 2e Ed

Prérequis. Théorème de Dirichlet sur les séries de Fourier, transformée de Fourier de la gaussienne

Théorème 2.2.11. *Soit $f \in \mathcal{S}(\mathbf{R})$. Alors pour tout $x \in \mathbf{R}$ on a :*

$$\sum_{n=-\infty}^{+\infty} f(x+2n\pi) = \frac{1}{2\pi} \sum_{n=-\infty}^{+\infty} \widehat{f}(n) \mathrm{e}^{\mathrm{i}nx},$$

avec la convention $\widehat{f}(\xi) = \int_{\mathbf{R}} f(x)\mathrm{e}^{-\mathrm{i}x\xi}\,\mathrm{d}x$.

Démonstration. La série du membre de gauche converge normalement sur tout compact de $\mathbf{R}$. En effet on sait qu'il existe $M > 0$ tel que $|f(x)| \leqslant M/(1+x^2)$ pour tout $x \in \mathbf{R}$, et alors pour tout x dans un compact $[-K, K]$ et tout $|n| \geqslant K$:

$$|f(x+2\pi n)| \leqslant \frac{M}{1+(x+2\pi n)^2} \leqslant \frac{M}{1+(2\pi|n|-K)^2} \in \ell^1 \text{ indépendant de } x.$$

On peut alors poser :

$$F(x) = \sum_{n=-\infty}^{+\infty} f(x+2\pi n)$$

pour tout $x \in \mathbf{R}$. Cette fonction est de classe C^1, puisque le raisonnement ci-dessus s'applique aussi à f' et donc le théorème de dérivation des intégrales à paramètre s'applique.

De plus, F est 2π-périodique. En effet, pour $N \in \mathbf{N}$ et $x \in \mathbf{R}$ on a :

$$\underbrace{\sum_{n=-N}^{N} f(x+2\pi+2n\pi)}_{\to F(x+2n\pi)} = \underbrace{\sum_{n=-N-1}^{N+1} f(x+2n\pi)}_{\to F(x)} - \underbrace{f(x-2N\pi) - f(x-2(N+1)\pi)}_{\to 0}.$$

On peut donc appliquer le théorème de Dirichlet à F :

$$F(x) = \sum_{n=-\infty}^{+\infty} c_n(F)\mathrm{e}^{\mathrm{i}nx}$$

avec les coefficients de Fourier valant :

$$\begin{aligned} c_n(F) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} F(x)\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=-\infty}^{+\infty} f(x+2k\pi)\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x \\ &= \frac{1}{2\pi}\sum_{k=-\infty}^{+\infty}\int_{-\pi}^{\pi} f(x+2k\pi)\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x \\ &= \frac{1}{2\pi}\sum_{k=-\infty}^{+\infty}\int_{(2k-1)\pi}^{(2k+1)\pi} f(u)\mathrm{e}^{-\mathrm{i}n(u-2k\pi)}\,\mathrm{d}u \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(u)\mathrm{e}^{-\mathrm{i}nu}\,\mathrm{d}u = \frac{1}{2\pi}\widehat{f}(n). \end{aligned}$$

□

Corollaire 2.2.12. *Pour tout $s > 0$, on définit :*

$$\theta(s) = \sum_{n=-\infty}^{+\infty} \mathrm{e}^{-\pi n^2 s}$$

et l'on déduit de la formule de Poisson que :

$$\theta(-1/s) = \sqrt{s}\theta(s).$$

Démonstration. On pose $f : x \mapsto \mathrm{e}^{-x^2 s/4\pi}$ et l'on calcule sa transformée de Fourier : en la dérivant (avec théorème de dérivation des intégrales à paramètres) et en résolvant l'équation différentielle avec le calcul de l'intégrale de Gauss, on tombe finalement sur :

$$\widehat{f}(\xi) = \frac{2\pi}{\sqrt{s}}\mathrm{e}^{\frac{\xi^2\pi}{s}}.$$

On applique alors la formule de Poisson en $x = 0$ pour obtenir l'équation fonctionnelle. □

Corollaire 2.2.13 (Cohn-Elkies)**.** *S'il existe une fonction $f : \mathbf{R}^n \to \mathbf{R}^n$ non identiquement nulle pour laquelle on dispose d'un $\delta > 0$ tel que f et $\widehat{f}$ soient dominées par $(1+|x|)^{-n-\delta}$, et vérifiant $f(x) \leqslant 0$ pour $|x| \geqslant 1$ et $\widehat{f}(\xi) \geqslant 0$ pour tout ξ, alors la densité de tout empilement de sphères en dimension n est majorée par $f(0)/2^n\widehat{f}(0)$.*

Idées de la démonstration. Il suffit de le démontrer pour les empilements périodiques, c'est-à-dire pour lesquels les centres des sphères sont situés sur les points d'un réseau Λ. On applique alors la formule de Poisson à f (écrite dans le théorème en dimension 1, elle est valable en dimension quelconque en remplaçant la somme sur les entiers et les multiples de 2π par des sommes sur les éléments de Λ) pour minorer $f(0)$ par une quantité proportionnelle à $\widehat{f}(0)/|\Lambda|$. □

Remarques.

— Faire seulement la formule de Poisson est trop court, alors il faut en général rajouter un petit corollaire. Un autre corollaire possible est le calcul de $\zeta(2)$, qui se fait en transformant la somme en une somme géométrique grâce à la formule de Poisson.

— L'équation fonctionnelle obtenue sur la fonction θ de Jacobi dans le corollaire n'est pas anodine. Premièrement elle permet d'établir une équation fonctionnelle sur la fonction ζ, mais plus généralement elle place θ au centre de la théorie des nombres, puisque cette relation trahit le fait que θ est une forme modulaire. Cette remarque est profondément liée à la remarque suivante.

— Le dernier corollaire n'a pas forcément vocation à être écrit dans le plan, mais peut être énoncé pendant la défense ou à la fin du développement. C'est un théorème miraculeux qui permet d'établir des majorations de la densité d'un empilement de sphères optimal rien qu'avec l'existence d'une certaine *fonction magique* (c'est leur nom). Maryna Viazovska, qui a reçu la médaille Fields en 2022 pour cela, a trouvé de telles fonctions magiques en dimensions 8 puis 24, qui ont donné une majoration égale à la minoration déjà connue d'un empilement optimal. Ainsi elle a démontré que l'empilement optimal en dimension 8 est donné par le réseau E_8, et celui en dimension 24 est donné par le réseau de Leech. C'est une remarque historique très récente qui permet d'ancrer le théorème dans les mathématiques d'aujourd'hui, donc c'est bienvenu à l'oral !

Recasages.

— 241 : On applique la convergence normale et le théorème de dérivation des intégrales à paramètre pour obtenir la formule. On utilise aussi les séries de Fourier et leur convergence simple pour les fonctions continues et C^1 par morceaux : tout le contenu mathématique de la démonstration est une illustration de la leçon.

— 246 : La démonstration établit un léger lien entre les séries de Fourier et la transformation de Fourier. On illustre le théorème de Dirichlet à des problématiques concrètes, pour changer de l'habituel calcul de $\zeta(2)$.

— 250 : Il n'y a pas vraiment de connaissances à propos de la transformation de Fourier à mobiliser pour ce développement, en tout cas pour la démonstration de la formule de Poisson. Par contre, les nombreux corollaires à cette formule sont riches en transformation de Fourier, par exemple ici on calcule (une nouvelle fois) la transformée de Fourier de la gaussienne (ce qui peut prendre du temps, mais c'est probablement le plus intéressant dans ce développement pour cette leçon).

— 265 : Non seulement la formule de Poisson permet d'étudier des fonctions usuelles comme les fonctions trigonométriques et hyperboliques, mais elle montre toute sa force dans l'étude des fonctions spéciales. Le premier exemple est celui de la fonction θ de Jacobi, suivi de l'étude de la fonction ζ et éventuellement des empilements de sphères. De quoi remplir la leçon avec autre chose que la fonction Γ.

2.2.7 Formule des compléments

Leçons 235, 236, 239, 245, 265, 267

Référence Stein-Shakarchi

Prérequis. Théorème des résidus, théorème de changement de variables, fonction Gamma

Théorème 2.2.14. *Pour tout $z \in \mathbf{C}$ tel que $0 < \mathfrak{Re}(z) < 1$, on a :*

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}.$$

Démonstration. Comme les deux membres de l'égalité sont holomorphes, il suffit de le démontrer pour $s \in]0,1[$. On calcule alors :

$$\begin{aligned}
\Gamma(s)\Gamma(1-s) &= \int_0^{+\infty} x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x \int_0^{+\infty} y^{-s}\mathrm{e}^{-y}\,\mathrm{d}y \\
&= \int_0^{+\infty}\int_0^{+\infty} x^{s-1}y^{-s}\mathrm{e}^{-(x+y)}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^{+\infty}\int_0^{+\infty} \left(\frac{uv}{v+1}\right)^{s-1}\left(\frac{u}{v+1}\right)^{-s}\mathrm{e}^{-u}\frac{u}{(v+1)^2}\,\mathrm{d}u\,\mathrm{d}v \\
&= \int_0^{+\infty}\int_0^{+\infty} \frac{v^{s-1}\mathrm{e}^{-u}}{v+1}\,\mathrm{d}u\,\mathrm{d}v \\
&= \int_0^{+\infty} \mathrm{e}^{-u}\,\mathrm{d}u \int_0^{+\infty} \frac{v^{s-1}}{v+1}\,\mathrm{d}v \\
&= \int_0^{+\infty} \frac{v^{s-1}}{v+1}\,\mathrm{d}v \\
&= \int_{\mathbf{R}} \frac{\mathrm{e}^{ts}}{1+\mathrm{e}^t}\,\mathrm{d}t
\end{aligned}$$

le premier changement de variable étant donné par $(u,v) = (x+y, x/y)$, le second par $t = \ln v$. Reste simplement à évaluer cette dernière intégrale, avec le théorème des résidus.

Posons donc $f(z) = \frac{\mathrm{e}^{zs}}{1+\mathrm{e}^z}$ qui est méromorphe sur $\mathbf{C}$, ses pôles étant en $(2k+1)\mathrm{i}\pi$ pour $k \in \mathbf{Z}$. On applique le théorème des résidus à f avec le lacet γ décrivant un rectangle autour du pôle $\mathrm{i}\pi$:

— γ_1 décrit le segment $[-R, R]$;
— γ_2 décrit le segment $[R, R+2\mathrm{i}\pi]$;
— γ_3 décrit le segment $[R+2\mathrm{i}\pi, -R+2\mathrm{i}\pi]$;
— γ_4 décrit le segment $[-R+2\mathrm{i}\pi, -R]$.

Par théorème des résidus on a :

$$\int_\gamma f = 2\mathrm{i}\pi \mathrm{Res}_f(\mathrm{i}\pi) = 2\mathrm{i}\pi \lim_{z\to\mathrm{i}\pi}(z-\mathrm{i}\pi)\frac{\mathrm{e}^{zs}}{1+\mathrm{e}^z} = 2\mathrm{i}\pi\frac{\mathrm{e}^{\mathrm{i}\pi s}}{\mathrm{e}^{\mathrm{i}\pi}} = -2\mathrm{i}\pi\mathrm{e}^{\mathrm{i}\pi s}.$$

D'autre part, sur γ_2 on a :

$$\left|\int_{\gamma_2} f\right| = \left|\int_0^{2\pi} \frac{\mathrm{e}^{sR}\mathrm{e}^{sit}}{1+\mathrm{e}^{R+it}}\mathrm{i}\ \mathrm{d}t\right| \leqslant \frac{2\pi\mathrm{e}^{sR}}{e^R-1} \to 0,$$

sur γ_4 on a de même :

$$\left|\int_{\gamma_4} f\right| = \left|\int_0^{2\pi} \frac{\mathrm{e}^{-sR}\mathrm{e}^{sit}}{1+\mathrm{e}^{-R+it}}\mathrm{i}\ \mathrm{d}t\right| \leqslant \frac{2\pi\mathrm{e}^{-sR}}{1-\mathrm{e}^{-R}} \to 0.$$

Sur γ_3 on a :

$$\int_{\gamma_3} f = -\int_{-R}^{R} \frac{\mathrm{e}^{ts}\mathrm{e}^{2\mathrm{i}\pi s}}{1+\mathrm{e}^{t+2\mathrm{i}\pi}}\ \mathrm{d}t = -\mathrm{e}^{2\mathrm{i}\pi s}\int_{-R}^{R}\frac{\mathrm{e}^{ts}}{1+\mathrm{e}^t}\ \mathrm{d}t = -\mathrm{e}^{2\mathrm{i}\pi s}\int_{-R}^{R} f$$

d'où finalement :

$$-2\mathrm{i}\pi\mathrm{e}^{\mathrm{i}\pi s} = \int_\gamma f = (1-\mathrm{e}^{2\mathrm{i}\pi s})\int_{-R}^{R} f \to \int_{\mathbf{R}} f$$

ce qui se réécrit :

$$\int_{\mathbf{R}} f = \frac{2\mathrm{i}\pi\mathrm{e}^{\mathrm{i}\pi s}}{\mathrm{e}^{2\mathrm{i}\pi s}-1} = \frac{\pi}{\sin(\pi s)}.$$

□

Corollaire 2.2.15. *On obtient le calcul de l'intégrale de Gauss :*

$$\int_{\mathbf{R}} \mathrm{e}^{-x^2}\ \mathrm{d}x = \sqrt{\pi}.$$

Démonstration. Il suffit d'évaluer la formule des compléments en $z = 1/2$, et voir avec un changement de variables que $\Gamma(1/2)$ est égal à l'intégrale de Gauss. □

Remarques.

- Le développement est très calculatoire et assez peu original. Mais s'il est bien maîtrisé, il est très bien car il montre une aisance avec le calcul d'intégrales, les changements de variables et le théorème des résidus.
- La formule des compléments sert principalement à deux choses. La première est le corollaire $\Gamma(1/2) = \sqrt{\pi}$, mais la formule ne joue pas un rôle très important puisque l'on peut obtenir cette égalité d'autres manières. L'application principale de cette formule est de pouvoir prolonger Γ à tout le plan complexe privé des pôles (les entiers négatifs), là où on n'avait avant que le demi-plan $\{\mathfrak{Re}(z) > 0\}$. Comme la fonction ζ de Riemann est solution d'une équation fonctionnelle faisant apparaître Γ, on peut ainsi prolonger analytiquement la fonction ζ aux complexes de partie réelle négative.
- Il est important de bien maîtriser le changement de variable en deux dimensions qui est fait au début. En particulier, savoir expliquer pourquoi c'est un C^1-difféomorphisme (par inversion globale), et savoir calculer le jacobien. Le plus dur est probablement d'arriver à exprimer x et y en fonction de u et v.

— Il faut absolument faire un dessin du chemin γ et y faire figurer les pôles de f. On explique à l'oral que l'on va faire tendre R vers $+\infty$ comme d'habitude, pour motiver le calcul.

Recasages.

— 235 : On utilise deux fois le théorème de Fubini-Tonelli, qui n'est assurément pas la meilleure manière d'illustrer la leçon, mais qui est quand même dans le thème.

— 236 : C'est un très bon développement si l'on insiste un peu sur le corollaire, puisque c'est une manière assez originale de calculer l'intégrale de Gauss. En plus, on illustre le théorème des résidus qui est central dans la leçon.

— 239 : L'étude de la fonction Γ est presque incontournable, et l'on commence par montrer que $\Gamma(s)\Gamma(1-s)$ est une intégrale à paramètre. Attention le développement n'utilise aucun théorème de régularité des intégrales à paramètre (à part si l'on n'a pas déjà expliqué que Γ est méromorphe), donc il faut vraiment insister sur le fait que l'on étudie Γ.

— 245 : On illustre bien le théorème des résidus dans un cas assez concret et loin d'être dénué de sens. Surtout que pour une fois, il y a un vrai résidu à calculer, ce qui est assez rare dans les calculs simples. On utilise aussi de manière cruciale le théorème des zéros isolés, pour simplifier le problème. C'est donc un beau résumé de la leçon.

— 265 : C'est surtout si l'on fait une partie sur l'étude de la fonction Γ. C'est très souvent le cas, et ce développement apparaît très souvent dans cette leçon, donc rien de très marquant.

— 267 : Une application du théorème des résidus dans cette leçon est toujours la bienvenue, attention de ne pas trop en mettre non plus.

2.2.8 Inégalité isopérimétrique

Leçons 219, 236, 246, 267

Référence Zuily-Queffélec

Prérequis. Théorème de Stokes, Séries de Fourier, théorème de Jordan.

Théorème 2.2.16. *Soit $\gamma : [0,1] \to \mathbf{C}$ un lacet simple continu et C^1 par morceaux de longueur L, entourant une surface d'aire S. Alors :*

$$S \leqslant \frac{L^2}{4\pi}.$$

Les cas d'égalité se font exactement lorsque γ décrit un cercle.

Démonstration. Par homogénéité, on peut supposer que $L = 1$. Aussi, l'inégalité à démontrer ne dépend pas de la manière dont on a paramétrisé γ, donc on peut très bien le reparamétriser de sorte à avoir $|\gamma'(t)| = 1$ pour tout t, et à ce qu'il soit orienté positivement.

On développe γ en série de Fourier avec $(c_n)_{n\in\mathbf{Z}}$ pour coefficients : la dérivée γ' a pour coefficients de Fourier $(c'_n = 2\mathrm{i}\pi n c_n)_{n\in\mathbf{Z}}$. D'une part, on peut calculer la longueur de l'arc avec Parseval :

$$L = \int_0^1 |\gamma'(t)| \,\mathrm{d}t = \int_0^1 |\gamma'(t)|^2 \,\mathrm{d}t = \sum_{n=-\infty}^{+\infty} |c'_n|^2 = \sum_{n=-\infty}^{+\infty} 4\pi^2 n^2 |c_n|^2$$

et d'autre part, la formule de Stokes donne, en écrivant $\gamma(t) = x(t) + \mathrm{i}y(t)$:

$$\begin{aligned} S &= \frac{1}{2}\int_0^1 (xy' - x'y) = \frac{1}{2}\mathfrak{Im}\int_0^1 \gamma'\overline{\gamma} = \frac{1}{2}\mathfrak{Im}\sum_{n=-\infty}^{+\infty} c'_n\overline{c_n} \\ &= \frac{1}{2}\mathfrak{Im}\sum_{n=-\infty}^{+\infty} 2\mathrm{i}\pi n|c_n|^2 = \sum_{n=-\infty}^{+\infty} \pi n|c_n|^2. \end{aligned}$$

Ainsi comme $L = L^2$ on obtient $L^2 - 4\pi S = 4\pi^2 \sum_{n=-\infty}^{+\infty}(n^2 - n)|c_n|^2 \geqslant 0$. Le cas d'égalité se produit exactement lorsque $c_n = 0$ pour $n \notin \{0, 1\}$, ce qui correspond au cas où γ décrit un cercle. □

Remarques.

- Le développement en l'état est très court, mais le principe est de remplir le temps restant avec des explications de tout ce que l'on a laissé de côté. En particulier :
  - La reparamétrisation par longueur d'arc pour avoir $|\gamma'| = 1$ est possible parce que la longueur du début de l'arc $t \mapsto \int_0^t |\gamma'|$ est un C^1-diffféomorphisme. La longueur de l'arc ne change pas après reparamétrisation par théorème de changement de variables.

- On n'utilise pas vraiment le fait que γ est égal partout à la somme de sa série de Fourier (ce qui est vrai car γ est continu et C^1 par morceaux), mais l'on utilise les hypothèses de régularité pour dire que $c_n' = 2\mathrm{i}\pi n c_n$.
- Le théorème de Stokes utilisé ici est aussi appelé *formule de Green-Riemann*. Si l'on veut détailler dans la généralité, on peut expliquer que l'on applique le théorème de Stokes à la 1-forme différentielle $\omega = \frac{1}{2}(x\,\mathrm{d}y - y\,\mathrm{d}x)$. La différentielle extérieure vaut $\mathrm{d}\omega = \mathrm{d}x \wedge \mathrm{d}y$, et le théorème de Stokes appliqué au compact entouré par γ (il est compact par le théorème de Jordan, et son bord est C^1 par morceaux ; il faudrait C^1 mais on peut toujours dire qu'au pire on découpe) donne la formule écrite ci-dessus. Le théorème de Stokes se démontre en introduisant une partition de l'unité adaptée à un atlas choisi, pour se ramener via les cartes au cas plat, qui se fait avec le théorème de Fubini et des intégrations par parties.
- Justement, le théorème de Jordan doit être connu, car il permet à l'énoncé d'avoir un sens. Il énonce que si γ est un lacet simple continu découpe le plan en deux composantes connexes, l'une bornée et l'autre non. Si l'on garde γ de classe C^1 par morceaux, alors on peut utiliser l'indice d'un lacet utilisé en analyse complexe. Si l'on suppose seulement γ continu, alors les démonstrations sont techniques, une pas trop compliquée utilise le théorème de point fixe de Brouwer.
- Les seuls coefficients de Fourier non nuls de γ sont c_0 et c_1 si et seulement si γ décrit un cercle. En effet, si c'est le cas alors comme γ est égal à la somme de sa série de Fourier (il est continu et C^1 par morceaux) il est de la forme $c_0 + c_1\mathrm{e}^{2\mathrm{i}\pi t}$ qui est clairement un cercle. Réciproquement, si γ décrit un cercle alors comme on a supposé qu'il était orienté positivement et que c'était un lacet simple, il est de la forme $c_0 + c_1\mathrm{e}^{2\mathrm{i}\pi t}$ et alors ses coefficients de Fourier sont évidents.

— On peut relaxer l'hypothèse C^1 par morceaux qui est très forte, le théorème est en fait vrai dès que γ est continu (donc dès l'application du théorème de Jordan). La démonstration avec les séries de Fourier et le théorème de Stokes tombe alors en ruine, et il faut faire autrement. Une démonstration assez simple utilise le théorème de Brunn-Minkowski, qui énonce que si μ est la mesure de Lebesgue sur $\mathbf{R}^n$ et si A et B sont deux compacts non vides, alors :

$$\mu(A+B)^{1/n} \geqslant \mu(A)^{1/n} + \mu(B)^{1/n}.$$

On le démontre en partant du cas où A et B sont des pavés simples, puis on augmente pas à pas la généralité. Pour obtenir l'inégalité isopérimétrique (et en toutes dimensions !), on applique cette inégalité au compact K dont on veut majorer la mesure, et à une boule de rayon ε. On met alors l'inégalité à la puissance n, on soustrait $\mu(K)$, on divise par ε et on le fait tendre vers 0. À gauche on obtient l'aire de la surface de K, et à droite avec un développement limité on obtient à un facteur près $\mu(K)^{1/n}$.

— Il y a des inégalités plus précises que celle-ci. Par exemple, l'inégalité de Bonnesen énonce que si la surface entourée par γ est convexe et si r (resp. R) est le rayon

du cercle inscrit (resp. circonscrit), alors $L^2 - 4\pi S$ est plus grand que $\pi^2(R^2 - r^2)$.

Recasages.

— 219 : Même si les techniques utilisées ne sont pas du tout au cœur de la leçon, on cherche quand même à minimiser le périmètre pouvant entourer une surface d'aire donnée (ou maximiser l'aire de la surface entourée par une courbe de longueur donnée). Et l'on obtient une illustration d'une recherche d'extrema d'une fonction définie sur un espace assez sauvage et très différent d'un espace vectoriel normé comme on en a trop l'habitude dans cette leçon.

— 236 : C'est pas forcément idéal, mais on calcule quand même une intégrale (l'aire de la surface entourée) avec le théorème de Stokes. Il faudrait alors mettre le théorème de Stokes dans le plan...

— 246 : C'est une belle application de la théorie des séries de Fourier, mais il ne faut pas oublier que la démonstration dans le cas général ne les utilise pas. On fait quand même apparaître plusieurs fois l'égalité de Parseval et l'on visualise bien le lien avec les cercles, donc c'est très pertinent.

— 267 : Les problèmes isopérimétriques sont explicitement dans le rapport de jury sur cette leçon, donc c'est très bien. Attention quand même dans la défense, parce que le titre de la leçon comporte le mot *utilisations*.

2.2.9 Lax-Milgram et $-(pu')' + qu' + u = f$

Leçons 205, 213, 220

Références Bernis Bernis, Hirsch-Lacombe

Prérequis. Théorème de Riesz, critère de densité, espace de Sobolev $H_0^1(0,1)$

Théorème 2.2.17 (de Lax-Milgram). *Soient H un espace de Hilbert, $\phi \in H'$ et $a : H \times H \to \mathbf{R}$ une forme bilinéaire continue et coercive (il existe $\alpha > 0$ tel que $a(x,x) \geqslant \alpha\|x\|^2$). Alors il existe un unique $u \in H$ tel que $\phi = a(u,-)$. Si de plus a est symétrique, alors u est l'unique minimum de :*

$$\frac{1}{2}a(u,u) - \phi(u).$$

Démonstration. Soit $x \in H$; le théorème de Riesz appliqué à $a(x,-)$ donne un unique $Tx \in H$ tel que $a(x,y) = \langle Tx, y\rangle$ pour tout $y \in H$. La bilinéarité de a et l'unicité dans le théorème de Riesz montrent que T est linéaire, et la continuité de a montre que T est continu. Si $Tx = 0$ alors :

$$0 = \langle Tx, x\rangle = a(x,x) \geqslant \alpha\|x\|^2$$

donc T est injectif. Enfin, montrons que T est surjectif en montrant que son image est fermée. Pour toute suite $(y_n = Tx_n)_{n\in\mathbf{N}}$ dans l'image de T qui converge vers $y \in H$, on a par Cauchy-Schwarz :

$$\alpha\|x\|^2 \leqslant a(x,x) = \langle Tx, x\rangle \leqslant \|Tx\|\|x\|$$

pour tout $x \in H$, d'où pour tous $p, q \in \mathbf{N}$:

$$\|y_p - y_q\| = \|Tx_p - Tx_q\| \geqslant \alpha\|x_p - x_q\|,$$

ce qui montre que $(x_n)_{n\in\mathbf{N}}$ est de Cauchy. Elle converge vers un x dont l'image par T doit être y par continuité. Maintenant si y est orthogonal à l'image de T, alors on a :

$$\alpha\|y\|^2 \leqslant a(y,y) = \langle Ty, y\rangle = 0$$

donc $y = 0$, ce qui démontre que l'image de T, en plus d'être fermée, est dense dans H. Donc T est surjectif.

Il suffit alors de poser u' le représentant de Riesz de ϕ, puis $u = T^{-1}u'$. C'est clairement l'unique u qui convient. Si maintenant a est symétrique, alors pour tout $v \in H$ on a :

$$\frac{1}{2}a(u+v,u+v) - \phi(u+v) = \frac{1}{2}a(u,u) - \phi(u) + \frac{1}{2}a(v,v)$$

et le tout dernier terme est positif, et nul exactement lorsque $v = 0$. □

Théorème 2.2.18. *On considère le problème elliptique :*

$$\begin{cases} -(\alpha u')' + \beta u' + u = f \\ u(0) = u(1) = 0. \end{cases}$$

Soient $\alpha \in \mathrm{L}^\infty([0,1])$ *que l'on peut encadrer :*

$$0 < \alpha_{\min} \leqslant \alpha \leqslant \alpha_{\max},$$

$f \in \mathrm{L}^2([0,1])$ *et* $\beta \in C^1([0,1])$ *vérifiant* $\beta' \leqslant 2$. *Alors il existe une unique solution faible* $u \in H_0^1$ *au problème elliptique, c'est-à-dire un unique* u *tel que pour tout* $v \in H_0^1$,

$$\langle \alpha u', v' \rangle_2 + \langle \beta u', v \rangle_2 + \langle u, v \rangle_2 = \langle f, v \rangle_2.$$

Démonstration. On appelle $a(u,v)$ le membre de gauche. Comme H_0^1 est de Hilbert et $\langle f, - \rangle_2$ est continue, il suffit de démontrer que a est bilinéaire continue et coercive. La bilinéarité est évidente, et la continuité est facile :

$$\begin{aligned} |a(u,v)| &\leqslant \alpha_{\max} \|u'\|_2 \|v'\|_2 + \|\beta\|_\infty \|u'\|_2 \|v\|_2 + \|u\|_2 \|v\|_2 \\ &\leqslant C \|u\| \|v\|_2 \end{aligned}$$

d'après l'inégalité de Poincaré. Pour la coercivité, soit $u \in H_0^1$. Alors :

$$\begin{aligned} a(u,u) &= \langle \alpha u', u' \rangle_2 + \langle \beta u', u \rangle_2 + \|u\|_2^2 \\ &\geqslant \alpha_{\min} \|u'\|_2^2 + \langle \beta u', u \rangle_2 + \|u\|_2^2. \end{aligned}$$

Si u est lisse à support compact dans $]0,1[$ alors par intégration par parties :

$$\langle \beta u', u \rangle_2 = \int_0^1 \beta u' u = -\frac{1}{2} \int_0^1 \beta' u^2 \geqslant -\|u\|_2^2.$$

Par continuité du produit scalaire et densité des fonctions lisses à support compact dans H_0^1, c'est vrai pour tout u. Finalement :

$$\begin{aligned} a(u,u) &\geqslant \alpha_{\min} \|u'\|_2^2 - \|u\|_2^2 + \|u\|_2^2 \\ &\geqslant \alpha_{\min} \|u'\|_2^2 \\ &\geqslant C' \|u\|_{H^1}^2. \end{aligned}$$

Le théorème de Lax-Milgram conclut. □

Remarques.

— Si l'on rajoute des hypothèses (notamment la continuité de f), alors on peut montrer que l'unique solution faible est en fait de classe C^2, puis en fait une solution forte. Ainsi on a résolu le problème elliptique en passant par des solutions faibles dans des espaces de Hilbert ! C'est une bonne application de la théorie des distributions et de la théorie des espaces de Hilbert à l'étude d'équations différentielles.

— Il existe d'autres hypothèses possibles sur α, β et f pour que le résultat soit vrai, mais celles-ci ont l'air d'être les plus simples à retenir et à utiliser.
— Le théorème de Lax-Milgram peut être précisé en le théorème de Stampacchia (c'est une lignée évolutive de Pokémon : Riesz $\to$ Lax-Milgram $\to$ Stampacchia). Il énonce que sous les mêmes hypothèses, pour tout compact K de H il existe un unique $u \in K$ tel que pour tout $v \in K$:

$$a(u, v-u) \geqslant \phi(v-u).$$

Et de même, si a est symétrique alors u est l'unique minimum de la même fonctionnelle que dans l'énoncé de Lax-Milgram. Le théorème de Stampacchia peut par exemple servir en physique pour établir des théorèmes à propos de l'énergie d'un système mécanique.

Recasages.
— 205 : On utilise la complétude de H pour montrer que l'image de T est fermée. Sinon c'est un théorème assez usuel dans les parties sur les espaces de Hilbert, et il est très possible de faire une partie sur les espaces de Hilbert dans cette leçon.
— 213 : C'est bien de mettre ce théorème après celui de Riesz par exemple, en ajoutant la définition et l'étude de l'espace de Sobolev $H_0^1(0,1)$. Pour les résultats de base qui sont utilisés ici, voir Hirsch-Lacombe, chapitre 10.
— 220 : Il est assez impressionnant que des théorèmes abstraits à propos d'espaces de Hilbert comme le théorème de Riesz ou celui de Lax-Milgram permettent de résoudre des problèmes aussi concrets qu'une équation différentielle. Bien sûr presque toute la théorie des équations différentielles se base sur l'analyse fonctionnelle de tels espaces, mais ce développement en est une bonne illustration.

2.2.10 Les isométries locales sont des isométries affines

Leçons 204, 214, 215

Référence Gourdon Analyse 2e Ed

Prérequis. Théorème d'inversion locale, théorème des accroissements finis

Théorème 2.2.19. *Soit $f : \mathbf{R}^n \to \mathbf{R}^n$ de classe C^1 telle que $\mathrm{d}f(x)$ est une isométrie pour tout $x \in \mathbf{R}^n$. Alors f est une isométrie affine.*

Démonstration. Par hypothèse on a $\|\mathrm{d}f(x)\| = 1$ pour tout $x \in \mathbf{R}^n$, ainsi l'inégalité des accroissements finis montre que f est 1-lipschitzienne.

Soit $a \in \mathbf{R}^n$. Comme $\mathrm{d}f(a)$ est inversible, l'inversion locale donne un voisinage $V_a \ni a$ tel que $f|_{V_a}$ soit un C^1-difféomorphisme $V_a \to W_a = f(V_a)$. Pour tout $y = f(x) \in W_a$, le théorème des fonctions composées montre que $\mathrm{d}(f^{-1})(y) = \mathrm{d}f(x)^{-1}$ est une isométrie, donc $\|\mathrm{d}(f^{-1})(y)\| = 1$. On peut alors appliquer le même raisonnement qu'au début : l'inégalité des accroissements finis montre que f^{-1} est 1-lipschitzienne sur W_a (il faut que W_a soit convexe mais on peut le supposer quitte à réduire V_a). Finalement, pour tous $x, x' \in V_a$:

$$\|x - x'\| = \|f^{-1}f(x) - f^{-1}f(x')\| \leqslant \|f(x) - f(x')\|$$

d'où $\|f(x) - f(x')\| = \|x - x'\|$.

Cette égalité s'écrit :

$$\langle f(x) - f(x'), f(x) - f(x')\rangle = \langle x - x', x - x'\rangle$$

qui se différentie par rapport à x :

$$2\langle \mathrm{d}f(x)(h), f(x) - f(x')\rangle = 2\langle h, x - x'\rangle$$

puis par rapport à x' :

$$-\langle \mathrm{d}f(x)(h), \mathrm{d}f(x')(k)\rangle = -\langle h, k\rangle$$

Ce qui permet finalement de montrer que :

$$\|\mathrm{d}f(x)(h) - \mathrm{d}f(x')(h)\|^2 = \|\mathrm{d}f(x)(h)\|^2 - 2\langle \mathrm{d}f(x)(h), \mathrm{d}f(x')(h)\rangle + \|\mathrm{d}f(x')(h)\|^2 = 0$$

Maintenant que l'on sait que $\mathrm{d}f$ est constante sur V_a pour tout a, on peut conclure, comme dans tous les résultats de propagation de propriétés sur des espaces connexes. L'ensemble $\Gamma = \{a \in \mathbf{R}^n \mid \mathrm{d}f(a) = \mathrm{d}f(0)\}$ est ouvert d'après ce qui précède, mais c'est aussi un fermé puisque f est de classe C^1. Donc $\mathrm{d}f$ est constante sur $\mathbf{R}^n$, et $f - \mathrm{d}f(0)$ est de différentielle nulle donc constante : f est une isométrie affine. □

Remarques.

— Le développement est très court, il faut donc bien prendre son temps et faire des dessins. On peut par exemple illustrer l'inversion locale, les accroissements finis, le fait que $\langle \mathrm{d}f(x)(h), \mathrm{d}f(x')(k)\rangle = \langle h, k\rangle$, le fait que $\mathrm{d}f$ soit localement constante et pourquoi on peut en déduire qu'elle est constante, ... De même on peut expliciter le théorème des accroissements finis.

— C'est un théorème de passage local-global, qui se fait par connexité comme beaucoup d'autres en analyse. On peut faire un commentaire sur ce point, et ainsi motiver la démonstration. Ce n'est pas un développement long ni difficile, donc c'est mieux si on l'accompagne d'une preuve de recul sur les notions utilisées.

— Tout à la fin on utilise le fait que f est C^1 pour conclure que $\mathrm{d}f$ est constante, mais c'est trop fort. La fonction $\mathrm{d}f$ est localement constante, donc elle est constante puisque $\mathbf{R}^n$ est connexe. On utilise la continuité de $\mathrm{d}f$ pour dire que Γ est fermé, mais on peut faire plus simple puisque les $(\mathrm{d}f)^{-1}(y)$ pour $y \in \mathbf{R}^n$ sont des ouverts disjoints de $\mathbf{R}^n$ qui le recouvrent.

Recasages.

— 204 : C'est une démonstration par connexité comme il en existe beaucoup d'autres. Elle est assez originale pour l'agrégation, mais pas très difficile.

— 214 : C'est une application directe du théorème d'inversion locale. On peut aussi citer le théorème d'inversion globale, le théorème de Hadamard-Lévy, et plein d'autres. Celui-ci a l'avantage d'être assez simple et d'être un résultat intéressant en lui-même.

— 215 : On illustre une propriété de rigidité sur les fonctions différentiables, et le résultat fait joli dans un plan.

2.2.11 Montel et Osgood

Leçons 201, 203, 205, 241, 245

Références Zuily-Queffélec, Bernis Bernis

Prérequis. Théorème d'Ascoli, inégalité des accroissements finis, théorème de Baire, théorème de Weierstrass

Théorème 2.2.20 (de Montel)**.** *Soit U un ouvert de $\mathbf{C}$. Alors de toute suite $(f_n)_{n\in\mathbf{N}}$ de fonctions holomorphes sur U qui est bornée (elle est bornée pour $\|.\|_{\infty,K}$ pour tout compact K de U), on peut extraire une sous-suite qui converge uniformément sur tout compact.*

Démonstration. Soit K un compact de U. On veut extraire une sous-suite de $(f_n)_{n\in\mathbf{N}}$ qui converge uniformément sur K, donc on peut appliquer le théorème d'Ascoli.

(f_n) est ponctuellement bornée. Comme (f_n) est bornée pour $\|.\|_{\infty,K}$, elle est bornée en tout point.

(f_n) est équicontinue. Soient $z_0 \in K$ et $\varepsilon > 0$. On inclut un disque fermé $\overline{D}(z_0, r)$ dans U qui est ouvert, et l'inégalité des accroissements finis donne pour $z \in \overline{D}(z_0, r/2)$:

$$|f_n(z) - f_n(z_0)| \leqslant \|f_n'\|_{\infty,\overline{D}(z_0,r/2)}|z - z_0|.$$

On conclut avec le lemme suivant.

Lemme 2.2.21. *Il existe $M > 0$ tel que pour toute fonction holomorphe f sur U :*

$$\|f'\|_{\infty,\overline{D}(z_0,r/2)} \leqslant M\|f\|_{\infty,\overline{D}(z_0,r)}.$$

Démonstration. Pour $z \in \overline{D}(z_0, r/2)$, on a la formule de Cauchy :

$$f'(z) = \frac{1}{2\mathrm{i}\pi}\int_{\partial D(z_0,r)} \frac{f(\zeta)}{(\zeta - z)^2}\,\mathrm{d}\zeta = \frac{1}{2\mathrm{i}\pi}\int_0^{2\pi} \frac{f(z_0 + r\mathrm{e}^{\mathrm{i}t})}{(z_0 + r\mathrm{e}^{\mathrm{i}t} - z)^2}\mathrm{i}r\mathrm{e}^{\mathrm{i}t}\,\mathrm{d}t$$

d'où :

$$\begin{aligned}|f'(z)| &\leqslant \frac{1}{2\pi}\int_0^{2\pi} \frac{|f(z_0 + r\mathrm{e}^{\mathrm{i}t})|}{(r/2)^2} r\,\mathrm{d}t \\ &\leqslant \frac{4}{r}\|f\|_{\infty,\overline{D}(z_0,r)}\end{aligned}$$

ce qui conclut. □

Le lemme permet de conclure puisque la suite $\|f_n\|_{\infty,\overline{D}(z_0,r)}$ est bornée.

Le théorème d'Ascoli s'applique alors, et permet d'extraire une sous-suite qui converge uniformément sur K. Pour conclure, il suffit de trouver une suite exhaustive de compacts pour U, par exemple :

$$K_n = \{z \in U \mid |z| \leqslant n, d(z, U^c) \geqslant 1/n\}$$

puis de faire une extraction diagonale. □

Théorème 2.2.22 (d'Osgood)**.** *Soient U un ouvert de $\mathbf{C}$ et $(f_n)_{n\in\mathbf{N}}$ une suite de fonctions holomorphes sur U qui converge simplement vers une fonction (quelconque) f. Alors f est holomorphe sur un ouvert dense dans U.*

Démonstration. Soit $\overline{D}$ un disque fermé inclus dans U, il suffit de montrer que f est holomorphe sur un ouvert V inclus dans $\overline{D}$. En posant :

$$F_k = \{z \in \overline{D} \mid \sup_{n\in\mathbf{N}} |f_n(z)| \leqslant k\},$$

on remarque que $\bigcup_{k\in\mathbf{N}} F_k = \overline{D}$. Comme $\overline{D}$ est complet on peut appliquer le lemme de Baire, et trouver un F_{k_0} d'intérieur non vide. On choisit V un ouvert inclus dans F_{k_0}. Mais alors la suite (f_n) est bornée sur tout compact de V, et le théorème de Montel donne une sous-suite qui converge uniformément sur tout compact de V. Par théorème de Weierstrass, la limite, qui coïncide avec f, doit être holomorphe sur V. □

Remarques.

- Le théorème de Montel sert à d'autres choses qu'à démontrer le théorème d'Osgood. Par exemple, on peut l'utiliser pour démontrer le théorème de l'application conforme de Riemann, qui énonce que les seuls ouverts simplement connexes de $\mathbf{C}$ (à biholomorphisme près) sont $\mathbf{C}$ et le disque unité (cf. Rudin, théorème 14.8 page 330). Une autre application est le théorème de Henri Cartan en dynamique discrète (cf. Zuily-Queffélec, théorème III.7 page 161).
- Mieux vaut savoir détailler le passage de K à tous les compacts avec l'extraction diagonale. La suite exhaustive de compacts permet de recouvrir l'ouvert U avec un nombre dénombrable de compacts, ce qui permet l'extraction diagonale.
- Le théorème d'Osgood sert peut-être (?) à démontrer des résultats théoriques, mais en pratique on ne connaît pas l'ouvert sur lequel la limite est holomorphe, et donc c'est à peu près inutilisable.

Recasages.

- 201 : L'espace des fonctions holomorphes sur un ouvert de $\mathbf{C}$ est, comme les espaces de Fréchet C^∞, un espace souvent un peu oublié dans cette leçon. C'est bien de ne pas se restreindre aux espaces vectoriels normés, surtout si l'on a des choses à dire sur les espaces en eux-mêmes. Ici on montre que les parties bornées de l'espace des fonctions holomorphes sur un ouvert de $\mathbf{C}$ sont séquentiellement compactes, grâce à la rigidité de l'analyse complexe.

— 203 : On utilise le théorème d'Ascoli, et on démontre un résultat de compacité (séquentielle) à la Bolzano-Weierstrass, que l'on utilise ensuite pour démontrer un théorème assez fort sur la convergence simple des fonctions holomorphes.
— 205 : La complétude sert ici à appliquer le théorème d'Ascoli, qui est en plein dans la leçon. On peut par exemple mettre le développement juste après Ascoli dans le plan.
— 241 : Le théorème illustre la rigidité bien connue des fonctions holomorphes. C'est bien de faire le parallèle avec les fonctions réelles, où la convergence simple est très loin de conserver la dérivabilité. Attention quand même, la leçon demande de donner beaucoup d'exemples.
— 245 : L'essentiel de la démonstration du théorème de Montel se trouve dans le lemme, qui est une inégalité miraculeuse provenant de la formule de Cauchy (comme d'habitude). On pourrait citer tous les autres miracles que cette formule a pu créer, pour illustrer le début de la leçon.

2.2.12 Nombre de zéros d'une solution d'équation différentielle

Leçons 220, 221, 224, 267

Référence Zuily-Queffélec

Prérequis. Théorème de Cauchy-Lipschitz, sommation des relations de comparaison.

Théorème 2.2.23. *Soient $a \in \mathbf{R}$, et $q \in C^1([a,+\infty[)$ strictement positive telle que $\int_a^{+\infty} \sqrt{q} = +\infty$ et $q' = o(q^{3/2})$. Notons y une solution non nulle de l'équation de Hill-Matthieu $y'' + qy = 0$ sur $[a,+\infty[$, et $N(x)$ le nombre de zéros de y sur $[a,x]$. Alors :*

$$N(x) \sim \frac{1}{\pi} \int_a^x \sqrt{q}.$$

Démonstration. On commence par reparamétriser y, en posant $\tau(x) = \int_a^x \sqrt{q}$. C'est un C^1-difféomorphisme strictement croissant $[a,+\infty[\to [0,+\infty[$. On peut alors poser $Y = y \circ \tau^{-1}$, et en dérivant on obtient :

$$\begin{aligned} y'(x) &= \sqrt{q(x)} Y'(\tau(x)), \\ y''(x) &= \frac{q'(x)}{2\sqrt{q(x)}} Y'(\tau(x)) + \sqrt{q(x)} Y''(\tau(x)). \end{aligned}$$

Ainsi de l'équation définissant y on déduit que :

$$Y'' + \phi Y' + Y = 0, \qquad \phi(t) = \frac{q'(\tau^{-1}(t))}{2q^{3/2}(\tau^{-1}(t))}.$$

On remarque que Y et Y' ne peuvent pas avoir de zéro commun, sinon par Cauchy-Lipschitz on aurait $Y = 0$ puis $y = 0$. On applique alors le lemme de relèvement suivant :

Lemme 2.2.24. *Soient $y_1, y_2 \in C^1([0,+\infty[)$ sans zéro commun. En écrivant $y_1(0) + \mathrm{i}y_2(0) = r_0 \mathrm{e}^{\mathrm{i}\theta_0}$, on peut alors trouver $r, \theta \in C^1([0,+\infty[)$ tels que $y_1 = r\cos(\theta)$ et $y_2 = r\sin(\theta)$.*

On obtient ainsi r et θ tels que $Y = r\sin(\theta)$ et $Y' = r\cos(\theta)$. En dérivant on obtient :

$$\begin{aligned} Y' &= r'\sin(\theta) + r\theta'\cos(\theta) = r\cos(\theta), \\ Y'' &= r'\cos(\theta) - r\theta'\sin(\theta) = -\phi r\cos(\theta) - r\sin(\theta). \end{aligned}$$

En calculant $\cos(\theta)$ fois la première égalité moins $\sin(\theta)$ fois la seconde, on obtient $\theta' = 1 + \phi\cos(\theta)\sin(\theta)$ (car r ne s'annule pas). Ainsi $|\theta' - 1| \leqslant \frac{1}{2}|\phi|$. L'expression de ϕ montre avec les hypothèses sur q que $\phi \to 0$, donc $\theta' \to 1$, puis par sommation des relations de comparaison, $\theta(x) \sim x$ lorsque $x \to +\infty$.

On peut alors compter le nombre $M(t)$ de zéros de Y sur $[0,t]$. Pour un t_0 assez grand on a $\theta' > 0$ sur tout $[t_0, +\infty[$. On a alors :

$$\begin{aligned} M(t) - |\{u \in [0,t_0] \mid Y(u) = 0\}| &= |\{u \in [t_0,t] \mid sin(\theta(u)) = 0\}| \\ &= |\{k \in \mathbf{Z} \mid \theta(t_0) \leqslant k\pi \leqslant \theta(t)\} \sim \frac{\theta(t)}{\pi} \sim \frac{t}{\pi}. \end{aligned}$$

Le cardinal $|\{u \in [0,t_0] \mid Y(u) = 0\}|$ est fini car sinon les zéros de Y sur $[0,t_0]$ auraient un point d'accumulation et on y trouverait un zéro commun de Y et Y'. Donc $M(t) \sim t/\pi$. Enfin, $N = M \circ \tau$ ce qui donne l'équivalent annoncé.

Démonstration du lemme. On pose $\alpha = y_1 + \mathrm{i}y_2$ qui ne s'annule pas, puis :

$$\beta(x) = \int_0^x \frac{\alpha'}{\alpha} + \ln(r_0) + \mathrm{i}\theta_0.$$

Les deux fonctions α et β sont clairement de classe C^1. On a alors $(\alpha \mathrm{e}^{-\beta})' = 0$ et $\alpha(0)\mathrm{e}^{-\beta(0)} = 1$ d'où $y_1(x) + \mathrm{i}y_2(x) = \alpha(x) = \mathrm{e}^{\beta(x)}$ pour tout $x \geqslant 0$. On a alors $r = |\alpha|$ et $\theta = \mathfrak{Im}(\beta)$ qui sont bien de classe C^1. □

□

Remarques.

— Le livre de Zuily et Queffélec donne une expression explicite de r et θ dans le lemme, ce qui rallonge inutilement la démonstration car on n'utilise pas du tout ces expressions.

— Il est bien d'expliquer à l'oral pourquoi on fait chaque étape. On reparamétrise y pour se ramener à une équation où l'on voit $Y'' + Y$, et heureusement le terme $\phi Y'$ est très petit par hypothèse sur q. Ainsi Y ressemble fortement à une solution de $f'' + f = 0$, c'est pourquoi on applique le lemme. Ce lemme ne dit rien d'autre que la possibilité de trouver une fonction *module* et une fonction *argument* pour un chemin C^1 qui ne passe pas par 0.

— L'hypothèse $\int_a^{+\infty} \sqrt{q} = +\infty$ sert à faire une sommation des relations de comparaison. Cauchy-Lipschitz s'applique pour $Y'' + \phi Y' + Y = 0$ parce que ϕ est de classe C^1.

Recasages.

— 220 : C'est un très joli développement pour cette leçon, qui donne des résultats qualitatifs sur des équations que l'on ne sait pas résoudre en général. On peut le mettre juste après l'étude des équations d'ordre deux à coefficients constants, pour le voir comme une généralisation, et faire les liens avec cos et sin.

— 221 : Pareil c'est parfait, ce sont des équations linéaires mais assez complexes car les coefficients ne sont pas constants. On illustre en particulier le théorème de Cauchy-Lipschitz et la vectorialisation.

— 224 : C'est assez original puisque l'on donne un développement asymptotique d'une fonction qui est loin d'être usuelle. Et on fait un pont avec les équations différentielles qui ne sont pas toujours présentes dans la leçon.
— 267 : C'est pas forcément le meilleur endroit où mettre ce développement, mais on utilise le lemme de relèvement et le fait que (Y, Y') dessine une courbe C^1 dans le plan des phases pour obtenir des informations très précises. Il est très bien si l'on fait une partie à propos des portraits de phase par exemple.

2.2.13 Prokhorov et Lévy

Leçons 203, 228, 229, 261, 262

Référence Kurtzmann

Prérequis. Caractérisation de la convergence en loi par les fonctions de répartition, la fonction caractéristique caractérise la loi

Définition 2.2.25. Une suite $(\mu_n)_{n\in\mathbf{N}}$ de probabilités sur $\mathbf{R}^n$ *converge étroitement* vers μ si pour tout $\phi : \mathbf{R}^n \to \mathbf{R}$ continue bornée on a :

$$\int_{\mathbf{R}^n} \phi \, \mathrm{d}\mu_n \to \int_{\mathbf{R}^n} \phi \, \mathrm{d}\mu.$$

Autrement dit, la convergence étroite des lois de variables aléatoires correspond à la convergence en loi de ces variables.

Définition 2.2.26. Une suite $(\mu_n)_{n\in\mathbf{N}}$ de probabilités sur $\mathbf{R}^n$ est *tendue* si pour tout $\varepsilon > 0$ il existe un compact K tel que pour tout $n \in \mathbf{N}$:

$$\mu_n(K) \geqslant 1 - \varepsilon.$$

Dans la suite on se place dans $\mathbf{R}$.

Théorème 2.2.27 (de Helly)**.** *Soit (F_n) une suite de fonctions de répartition. Alors il existe une extraction $(\alpha_n)_{n\in\mathbf{N}}$ et une fonction $F : \mathbf{R} \to [0,1]$ croissante et continue à droite telle qu'en tout point x de continuité de F on ait :*

$$F_{\alpha_n}(x) \to F(x).$$

Démonstration. Pour $q \in \mathbf{Q}$, la suite $(F_n(q))_{n\in\mathbf{N}}$ est dans $[0,1]$, on peut donc en extraire une sous-suite qui converge vers un $F_{\mathbf{Q}}(q) \in [0,1]$. Par extraction diagonale, il existe une extraction $(\alpha_n)_{n\in\mathbf{N}}$ telle que pour tout $q \in \mathbf{Q}$, $F_{\alpha_n}(q)$ converge vers $F_{\mathbf{Q}}(q)$. On pose alors pour $x \in \mathbf{R}$:

$$F(x) = \inf_{q\in Q, x<q} F_{\mathbf{Q}}(q).$$

On définit ainsi une fonction $F : \mathbf{R} \to [0,1]$ croissante. Pour la continuité à droite, soient $x \in \mathbf{R}$ et $\varepsilon > 0$. On trouve un $q > x$ tel que $F_{\mathbf{Q}}(q) \leqslant F(x) + \varepsilon$, de sorte que pour tout $y \in [x, q]$ on ait :

$$F(x) \leqslant F(y) \leqslant F_{\mathbf{Q}}(q) \leqslant F(x) + \varepsilon.$$

Il reste à démontrer le dernier point, on fixe alors x un point de continuité de F, et $\varepsilon > 0$. Par continuité on trouve $y < x$ tel que :

$$F(x) - \varepsilon \leqslant F(y).$$

On n'a plus qu'à intercaler $y < r < x < s$ avec $r, s \in \mathbf{Q}$ et $F_{\mathbf{Q}}(s) \leqslant F(x) + \varepsilon$ pour obtenir :

$$\begin{aligned} F(x) - \varepsilon \leqslant F(y) \leqslant F_{\mathbf{Q}}(r) = \lim F_{\alpha_n}(r) \leqslant \liminf F_{\alpha_n}(x) \\ \leqslant \limsup F_{\alpha_n}(x) \leqslant \limsup F_{\alpha_n}(s) = F_{\mathbf{Q}}(s) \leqslant F(x) + \varepsilon. \end{aligned}$$

Le résultat est démontré en faisant tendre ε vers 0. □

Théorème 2.2.28 (de Prokhorov). *De toute suite tendue $(\mu_n)_{n\in\mathbf{N}}$ sur $\mathbf{R}$ on peut extraire une sous-suite qui converge étroitement.*

Démonstration. Soit F_n la fonction de répartition de μ_n. On applique le théorème de Helly pour obtenir une fonction F, et il suffit alors de montrer que c'est une fonction de répartition (la convergence étroite des lois est la convergence en loi des variables aléatoires), puisque l'on sait déjà qu'après extraction il y a convergence en tout point de continuité. Comme F est croissante et continue à droite, il manque seulement de montrer que F tend vers 0 (resp. 1) en $-\infty$ (resp. en $+\infty$).

On fixe $\varepsilon > 0$ et l'on obtient par tension un $M > 0$ tel que pour tout $n \in \mathbf{N}$:

$$\mu_n([-M, M]) \geqslant 1 - \varepsilon.$$

Comme F a un nombre au plus dénombrable de discontinuités, on suppose que M est un point de continuité de F, de sorte à pouvoir écrire en passant à la limite :

$$F(M) - F(-M) \geqslant 1 - \varepsilon.$$

Il ne reste plus qu'à faire tendre ε vers 0 pour obtenir le résultat. □

Théorème 2.2.29 (de Lévy). *Soit $(\mu_n)_{n\in\mathbf{N}}$ une suite de probabilités sur $\mathbf{R}$. On note ϕ_n la fonction caractéristique de μ_n. Si $(\phi_n)_{n\in\mathbf{N}}$ converge simplement vers une fonction ϕ continue en 0 alors celle-ci est la fonction caractéristique d'une probabilité μ vers laquelle $(\mu_n)_{n\in\mathbf{N}}$ converge étroitement.*

Démonstration. On veut appliquer le théorème de Prokhorov, donc on montre que $(\mu_n)_{n\in\mathbf{N}}$ est tendue. Pour $u > 0$, on a par Fubini :

$$\frac{1}{2u}\int_{-u}^{u}(1-\phi_n) = \int_{\mathbf{R}}\left(1 - \frac{\sin(ux)}{ux}\right)\mathrm{d}\mu_n(x) \geqslant \left(1 - \frac{2}{\pi}\right)\mu_n\left(\mathbf{R} \setminus \left[-\frac{\pi}{2u}, \frac{\pi}{2u}\right]\right).$$

Or, $\frac{1}{2u}\int_{-u}^{u}(1-\phi_n)$ converge vers $\frac{1}{2u}\int_{-u}^{u}(1-\phi)$ qui elle-même converge, lorsque $u \to 0$, vers 0 (par continuité de ϕ en 0). Ainsi pour tout $\varepsilon > 0$ il existe $u > 0$ tel qu'à partir d'un certain rang :

$$\mu_n\left(\mathbf{R} \setminus \left[\frac{-\pi}{2u}, \frac{\pi}{2u}\right]\right) \leqslant \varepsilon.$$

Donc la suite $(\mu_n)_{n\in\mathbf{N}}$ est tendue, on applique le théorème de Prokhorov pour obtenir une valeur d'adhérence μ. Mais comme toute valeur d'adhérence de $(\mu_n)_{n\in\mathbf{N}}$ a pour fonction caractéristique ϕ et que la fonction caractéristique caractérise la loi, μ est l'unique valeur d'adhérence de $(\mu_n)_{n\in\mathbf{N}}$, donc il y a convergence étroite. □

Remarques.

— C'est trop long pour quinze minutes, mais on peut se contenter de donner les idées du théorème de Prokhorov qui ne contient pas vraiment de raisonnement intéressant. La preuve intéressante se trouve dans le théorème de Helly et le théorème de Lévy, et on peut adapter en fonction de la leçon.

— Dans le théorème de Helly on utilise la séparabilité de $\mathbf{R}$ et une extraction diagonale pour obtenir une convergence à extraction près sur tout l'espace. Le théorème de Prokhorov est vrai de manière générale sur les espaces métriques séparables, mais la démonstration est bien sûr différente.

— Le théorème de Lévy sert à démontrer le théorème central limite : en posant $Y_i = \frac{X_i - \mu}{\sigma}$, on a le développement limité :

$$\phi_Y(t) = 1 + \frac{t^2}{2} + o(t^2)$$

et la moyenne $Z_n = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$ vaut $\sum_{i=1}^n Y_i/\sqrt{n}$, d'où :

$$\phi_{Z_n}(t) = \left(\phi_Y\left(\frac{t}{\sqrt{n}}\right)\right)^n = \left(1 + \frac{t^2}{2n} + o(t^2)\right)^n \to \mathrm{e}^{-t^2/2}$$

et le théorème de Lévy conclut.

— Les fonctions $F : \mathbf{R} \to [0,1]$ croissantes continues à droite et de limite nulle en $-\infty$ et 1 en $+\infty$ sont des fonctions de répartition. Plus précisément, c'est la fonction de répartition de $F^{-1}(U)$ où F^{-1} est l'inverse généralisé de F et où U est une variable aléatoire de loi uniforme sur $]0,1[$.

— Le théorème de Prokhorov sert aussi en théorie du transport optimal, pour démontrer qu'il existe une solution au problème de Kantorovich.

Recasages.

— 203 : Le théorème de Prokhorov est très clairement un théorème de compacité à la Bolzano-Weierstrass. Pour l'utilisation de la compacité, on est en plein dans le thème : on utilise le théorème de Prokhorov pour démontrer le théorème de Lévy qui est très utile en probabilités.

— 228 : Il faut mettre l'accent sur le théorème de Helly, mais c'est vraiment dans le thème (avec plusieurs fois l'utilisation de la définition de la continuité et des limites inférieure et supérieure), surtout si l'on fait une petite partie à propos des fonctions de répartition par exemple.

— 229 : Les fonctions de répartition sont monotones, et on l'utilise intensivement dans la démonstration du théorème de Helly : c'est comme ça que l'on passe de $F_{\mathbf{Q}}$ à F. On utilise aussi la monotonie dans la démonstration du théorème de Prokhorov pour dire que M peut être choisi comme un point de continuité : les fonctions monotones n'ont qu'un nombre au plus dénombrable de discontinuités.

— 261 : Rien de mieux que de démontrer le théorème de Lévy dans cette leçon, puisqu'il énonce justement une caractérisation de la convergence en loi. On pourra passer plus rapidement sur le théorème de Helly et sur celui de Prokhorov.

— 262 : De même qu'en 261, c'est un théorème de convergence en loi. On peut même énoncer et démontrer le théorème central limite dans la foulée s'il reste du temps, quitte à éclipser encore plus les théorèmes de Helly et Prokhorov.

2.2.14 Raikov

Leçons 241, 243, 245, 261, 264, 266

Référence

Prérequis. Relèvement exponentiel, formule de la moyenne, formule de Cauchy

Théorème 2.2.30. *Soit Z une variable aléatoire de Poisson de paramètre λ qui s'écrit $X+Y$ avec X et Y des variables aléatoires indépendantes à valeurs dans $\mathbf{N}$. Alors X et Y sont aussi de Poisson.*

Démonstration. On note G les séries génératrices, de manière à avoir :

$$G_Z(s) = \mathrm{e}^{\lambda(s-1)}, \quad \text{et} \quad G_Z = G_X G_Y.$$

On montre que G_X et G_Y sont entières. Pour cela, on remarque que par indépendance :

$$\mathbb{P}(X=n)\mathbb{P}(Y=0) \leqslant \mathbb{P}(Z=n) = \mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}$$

et $\mathbb{P}(Y=0)$ n'est pas nul puisque $\mathbb{P}(X=0)\mathbb{P}(Y=0) = \mathbb{P}(Z=0) \neq 0$. Ainsi $\mathbb{P}(X=n) = O(\lambda^n/n!)$ et G_X est de rayon infini. De même pour G_Y.

Comme G_Z ne s'annule pas il en est de même pour G_X et G_Y, qui peuvent alors se relever en $G_X = \exp(f)$ et $G_Y = \exp(g)$. Pour tout $s \in \mathbf{C}$ on a :

$$\exp(f(s)+g(s)) = \exp(\lambda(s-1))$$

et comme $s \mapsto f(s)+g(s)-\lambda(s-1)$ est continue à valeurs dans un espace discret elle est constante. Or G_X et G_Y sont réelles positives sur $[0,+\infty[$ donc on peut supposer que $f(x), g(x) \in \mathbf{R}$ pour $x \in \mathbf{R}$, et en déduire que $f(s)+g(s) = \lambda(s-1)$ pour tout $s \in \mathbf{C}$.

Comme $X \leqslant Z$ on dispose alors de la majoration :

$$\mathrm{e}^{\mathfrak{Re}(f(s))} = \left|\mathbb{E}\left[s^X\right]\right| \leqslant \mathbb{E}\left(|s|^Z\right) = \mathrm{e}^{\lambda(|s|-1)}$$

d'où $\mathfrak{Re}(f(s)) \leqslant \lambda(|s|-1) \leqslant \lambda|s|$. Il en va de même pour g, donc les parties réelles de ces deux fonctions entières sont majorées par des polynômes de degré 1. Le résultat vient alors du lemme suivant.

Théorème 2.2.31 (de la partie réelle de Hadamard)**.** *Si f est une fonction entière telle qu'il existe $m \in \mathbf{N}$ et $K > 0$ avec :*

$$\mathfrak{Re}(f(s)) \leqslant K|z|^m$$

pour tout $z \in \mathbf{C}$, alors f est polynomiale de degré $\leqslant m$.

Démonstration. On pose $f = u + \mathrm{i}v$ d'une part et $f(z) = \sum_{k=0}^{+\infty} a_k z^k$ d'autre part. On a d'une part l'égalité de la moyenne :

$$a_k = \frac{1}{2\pi r^k} \int_0^{2\pi} f(r\mathrm{e}^{\mathrm{i}\theta})\mathrm{e}^{-\mathrm{i}k\theta}\ \mathrm{d}\theta$$

et d'autre part :

$$0 = \int_{|\zeta|=r} f(\zeta)\zeta^{k-1}\ \mathrm{d}\zeta = \mathrm{i}r^k \int_0^{2\pi} f(r\mathrm{e}^{\mathrm{i}\theta})\mathrm{e}^{\mathrm{i}k\theta}\ \mathrm{d}\theta.$$

En ajoutant les deux, on obtient :

$$a_k r^k = \frac{1}{\pi} \int_0^{2\pi} u(r\mathrm{e}^{\mathrm{i}\theta})\mathrm{e}^{-\mathrm{i}k\theta}\ \mathrm{d}\theta.$$

Par formule de Cauchy on a $u(0) = \frac{1}{2\pi}\int_0^{2\pi} u(r\mathrm{e}^{\mathrm{i}\theta})\ \mathrm{d}\theta$, d'où :

$$|a_k|r^k + 2u(0) \leqslant \frac{1}{\pi} \int_0^{2\pi} (|u(r\mathrm{e}^{\mathrm{i}\theta})| + u(r\mathrm{e}^{\mathrm{i}\theta}))\ \mathrm{d}\theta.$$

On note alors $A(r) = \max_{|z|=r} u(z)$. Si $A(r) \leqslant 0$ alors $|u|+u = 0$ et alors $|a_k|r^k+2u(0) \leqslant 0$. Sinon, on a quand même $|u| + u \leqslant 2A(r)$ et :

$$|a_k|r^k + 2u(0) \leqslant 4A(r).$$

Dans tous les cas, $|a_k|r^k \leqslant \max(4A(r), 0) - 2u(0)$ ce qui montre en passant à la limite sur r que $a_k = 0$ pour tout $k > m$. □

□

Remarques.

- Le théorème est vrai avec d'autres lois que la loi de Poisson. Pour la loi normale c'est le théorème de Cramér, et pour la loi convolée d'une loi de Poisson et d'une loi normale, c'est un théorème de Linnik.
- Le théorème a une extension aux groupes abéliens localements compacts G : si $x_0 \in G$ et $\lambda > 0$ sont fixés, on peut considérer la probabilité :

  $$\mu = \mathrm{e}^{-\lambda}(\delta_{x_0} + \lambda\delta_{x_1} + \frac{\lambda^2}{2}\delta_{x_2} + \ldots).$$

  Si $\mu = \mu_1 * \mu_2$ et si x_0 est d'ordre 2 ou infini, alors μ_1 et μ_2 sont aussi des lois de Poisson. Si x_0 n'est pas d'ordre 2 ou infini, alors on ne peut rien dire en général.
- Le théorème de Cramér (qui dit la même chose mais pour les lois gaussiennes) s'applique en théorie du signal par exemple : si un bruit est supposé gaussien et qu'on le décompose en deux contributions indépendantes, alors ces deux bruits sont aussi gaussiens.

Recasages.

— 241 : Si on n'a rien à mettre dans cette leçon, ce développement utilise les fonctions génératrices des probabilités (donc des séries entières) pour démontrer un résultat de probabilités qui n'a *a priori* rien à voir. Mais il y a probablement mieux à faire dans cette leçon.

— 243 : cf. 241.

— 245 : On illustre la formule de la moyenne et la formule de Cauchy dans la démonstration du théorème de la partie réelle de Hadamard qui peut être un théorème du plan (avec tous les théorèmes de rigidité des fonctions entières). En plus on utilise le théorème de l'existence d'un logarithme qui peut amener à des discussions topologiques.

— 261 : Très bonne leçon pour ce développement, puisque l'on utilise le fait que la loi est caractérisée par la série génératrice des moments (ce qui est évident, en fait) et même l'énoncé parle de lois de Poisson.

— 264 : C'est encore mieux, puisque les lois de Poisson sont vraiment importantes dans cette leçon.

— 266 : L'indépendance permet souvent d'obtenir des informations intéressantes, mais ce développement le montre particulièrement bien. L'indépendance de X et Y permet de montrer que G_X et G_Y sont entières, ce qui est très fort.

2.2.15 Stone-Weierstrass / Bernstein

Leçons 201, 203, 209, 241, 264, 266

Référence Hirsch-Lacombe

Prérequis. Inégalité de Bienaymé-Tchebychev, théorème de Heine

Théorème 2.2.32. *Soit $f : [0,1] \to \mathbf{R}$ une fonction continue. On pose pour tout $n \in \mathbf{N}$:*

$$B_n f(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right).$$

Alors $B_n f \to f$ uniformément sur $[0,1]$, avec vitesse en $\omega_f(1/\sqrt{n})$ où ω_f est le module de continuité de f.

Démonstration. Soit $x \in [0,1]$. On considère une suite $(X_j)_{j\in\mathbf{N}}$ de variables aléatoires indépendantes et identiquement distribuées de loi de Bernoulli de paramètre x. On note $S_n = \sum_{j=1}^{n} X_j$ qui suit une loi binomiale de paramètres n et x. On a alors par formule de transfert $B_n f(x) = \mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right]$.

Comme f est uniformément continue, pour $\varepsilon > 0$ il existe $\delta > 0$ tel que $|x-y| \leqslant \delta$ implique $|f(x) - f(y)| \leqslant \varepsilon$. On a alors :

$$\begin{aligned}
|B_n f(x) - f(x)| &\leqslant \mathbb{E}\left[\left|f\left(\frac{S_n}{n}\right) - f(x)\right|\right] \\
&= \mathbb{E}\left[\left|f\left(\frac{S_n}{n}\right) - f(x)\right| \mathbf{1}_{|\frac{S_n}{n}-x|\leqslant\delta}\right] + \mathbb{E}\left[\left|f\left(\frac{S_n}{n}\right) - f(x)\right| \mathbf{1}_{|\frac{S_n}{n}-x|>\delta}\right] \\
&= \varepsilon + 2\|f\|_\infty \mathbb{P}\left(\left|\frac{S_n}{n} - x\right| > \delta\right).
\end{aligned}$$

La probabilité à la fin est majorée par Bienaymé-Tchebychev :

$$\mathbb{P}\left(\left|\frac{S_n}{n} - x\right| > \delta\right) \leqslant \frac{V(S_n/n)}{\delta^2} = \frac{x(1-x)}{n\delta^2} \leqslant \frac{1}{4n\delta^2}.$$

Ainsi pour $n \geqslant \|f\|_\infty / 2\varepsilon\delta^2$ on obtient :

$$\|B_n f - f\|_\infty \leqslant \varepsilon + \frac{\|f\|_\infty}{2n\delta^2} \leqslant 2\varepsilon.$$

On montre maintenant que la convergence se fait en vitesse $1/\sqrt{n}$. On pose :

$$\omega_f(\delta) = \sup_{|x-y|\leqslant\delta} |f(x) - f(y)|$$

le plus grand module de continuité possible. On a :

$$|B_n f(x) - f(x)|| \leqslant \mathbb{E}\left[\left|f\left(\frac{S_n}{n}\right) - x\right|\right] \leqslant \mathbb{E}\left[\omega_f\left(\left|\frac{S_n}{n} - x\right|\right)\right],$$

il faut donc majorer $\mathbb{E}[\omega_f(|Y|)]$ avec $Y = \frac{S_n}{n} - x$, sachant que l'on sait ce que valent $\mathbb{E}[|Y|]$ et $\mathbb{E}[Y^2]$:

$$\mathbb{E}[|Y|] \leqslant \sqrt{\mathbb{E}[Y^2]} = \sqrt{\frac{x(1-x)}{n}} \leqslant \frac{1}{2\sqrt{n}}.$$

On montre alors que ω_f croît au plus linéairement :

Lemme 2.2.33. *Si $g : [0,1] \to \mathbf{R}$ est continue alors pour tous $R, \delta > 0$ on a :*

$$\omega_g(R\delta) \leqslant (R+1)\omega_g(\delta).$$

Démonstration. On choisit $x < y \in [0,1]$ tels que $|x-y| \leqslant R\delta$, et le plus grand entier n tel que $x + n\delta \leqslant y$. Il suffit alors d'écrire :

$$g(y) - g(x) = g(y) - g(x+n\delta) + \sum_{k=0}^{n-1} g(x+(k+1)\delta) - g(x+k\delta)$$

pour obtenir le résultat. □

On en déduit qu'en choisissant $\delta > 0$ et en posant $R = |Y|/\delta$, on a :

$$\mathbb{E}[\omega_f(R\delta)] \leqslant \mathbb{E}[R+1]\omega_f(\delta) = \left(\frac{1}{\delta}\mathbb{E}[|Y|] + 1\right)\omega_f(\delta) \leqslant \left(\frac{1}{2\delta\sqrt{n}} + 1\right)\omega_f(\delta),$$

donc $\|B_n f - f\|_\infty \leqslant \left(\frac{1}{2\delta\sqrt{n}} + 1\right)\omega_f(\delta)$. On prend maintenant $\delta = 1/\sqrt{n}$ pour conclure. □

Remarques.

- Il faut faire un dessin pour expliquer le lemme, c'est même probablement suffisant pour remplacer la démonstration.
- L'inégalité sur la vitesse est optimale. On le voit en appliquant l'inégalité de Khintchine avec la fonction $x \mapsto \left|x - \frac{1}{2}\right|$.
- Si f est de classe C^2 alors la vitesse est encore plus rapide, avec Taylor-Lagrange.
- Bien sûr on vient de démontrer le théorème de densité des fonctions polynomiales dans les fonctions continues sur un compact. On peut par exemple en déduire que si $f : [0,1] \to \mathbf{R}$ est continue et vérifie $\int_0^1 f(x)x^n \, \mathrm{d}x = 0$ pour tout n alors f est nulle, ou que la limite uniforme d'une suite de polynômes est encore un polynôme.

Recasages.

- 201 : Le théorème de Stone-Weierstrass est assez incontournable dans cette leçon, que ce soit sous cette forme ou sous la forme améliorée par Stone (à propos des sous-algèbres de $C^0(X)$ qui séparent les points, avec X un espace compact quelconque).
- 203 : Le théorème de Stone-Weierstrass est une application importante du théorème de Heine. On l'utilise au début de manière directe, et après en introduisant le module de continuité, qui utilise parce que f est uniformément continue.

— 209 : Ici aussi le théorème de Stone-Weierstrass est important, même si le développement n'est pas très original.
— 241 : On fabrique une suite de fonctions $(B_n f)_{n\in\mathbf{N}}$ qui converge uniformément, c'est donc un bon exemple de convergence uniforme, qui en plus est utile hors de la leçon.
— 264 : Les variables de Bernoulli et binomiale que l'on introduit jouent un rôle important, même si on ne les voit pas dans l'énoncé : on obtient donc une jolie application des probabilités à un autre domaine des mathématiques.
— 266 : De même que pour 264, et l'indépendance des X_j est important dans la démonstration. Attention quand même, c'est assez secondaire.

2.2.16 Sunyer i Balaguer

Leçons 204, 205, 223, 228

Référence Gourdon Analyse 2e Ed

Prérequis. Théorème de Baire.

Théorème 2.2.34. *Soit $f : \mathbf{R} \to \mathbf{R}$ une application de classe C^∞ telle que pour tout $x \in \mathbf{R}$, il existe $n \in \mathbf{N}$ tel que $f^{(n)}(x) = 0$. Alors f est polynomiale.*

Démonstration. Posons $F_n = \{x \in \mathbf{R} \mid f^{(n)}(x) = 0\}$ pour $n \in \mathbf{N}$ et $\Omega = \bigcup_{n\in\mathbf{N}} \mathring{F}_n$. On veut montrer que f est polynomiale sur chaque composante connexe de l'ouvert Ω, puis que le complémentaire X de celui-ci est vide.

Soient ainsi $]a, b[$ une composante connexe de Ω, et $[c, d]$ un segment inclus dans $]a, b[$. En fixant $x_0 \in]c, d[$, comme il est dans un $\mathring{F}_{n_0}$, on peut trouver $\alpha > 0$ tel que $f^{(n)} = 0$ sur $]x_0 - \alpha, x_0 + \alpha[$. Sur cet intervalle f est donc égale à un polynôme P, on va montrer que c'est le cas sur $[c, d]$. Pour cela on pose :

$$\Gamma = \{t \in]x_0, d] \mid \forall x \in [x_0, t], f(x) = P(x)\}$$

qui n'est pas vide, et l'on remarque que $M = \sup \Gamma$ ne peut pas valoir autre chose que d. En effet si $M < d$ alors on peut trouver $\eta > 0$ tel que f coïncide avec un certain polynôme Q sur $]M - \eta, M + \eta[$, mais alors P et Q coïncident sur l'ensemble infini $]M - \eta, M[$ et cela contredit le fait que M est le supremum de Γ. Ainsi $f = P$ sur $[x_0, d]$, et le même raisonnement vers la gauche montre que $f = P$ sur $[c, d]$. Comme c'est valable pour tous les segments $[c, d]$, c'est valable sur toute la composante connexe $]a, b[$.

Reste à montrer que X est vide. On commence par montrer qu'il est parfait (il est fermé et n'a pas de point isolé). Si X admet un point isolé x_0, alors il existe $\varepsilon > 0$ tel que $]x_0 - \varepsilon, x_0 + \varepsilon[\cap X = \{x_0\}$. D'après ce qui précède, f coïncide avec un polynôme P sur $]x_0 - \varepsilon, x_0[$, et avec un polynôme Q sur $]x_0, x_0 + \varepsilon[$. En dérivant successivement en x_0, on obtient $P = Q$ d'où $f = P$ sur $]x_0 - \varepsilon, x_0 + \varepsilon[$. Ainsi cet intervalle est inclus dans un $\mathring{F}_{\deg(P)+1} \subset \Omega$ ce qui est absurde.

Supposons finalement $X \neq \emptyset$. Comme $\bigcup_{n\in\mathbf{N}} F_n = \mathbf{R}$ on a $X = \bigcup_{n\in\mathbf{N}} (X \cap F_n)$. Comme X est fermé dans $\mathbf{R}$ donc complet, le théorème de Baire montre que l'intérieur d'un certain $X \cap F_{n_0}$ (pour la topologie de X) doit être non vide. Il existe donc $a < b$ dans $\mathbf{R}$ tels que :

$$]a, b[\cap X \neq \emptyset \quad \text{et} \quad]a, b[\cap X \subset F_{n_0}.$$

On choisit alors $x \in]a, b[$. Si $x \in X$ alors $f^{(n_0)}(x) = 0$. Mieux, x n'est pas isolé donc on peut y trouver une suite injective (x_p) qui tend vers x, et par théorèmes de Rolle successifs et passage à la limite, on a $f^{(n)}(x) = 0$ pour tout $n \geqslant n_0$. Si $x \notin X$ alors $x \in \Omega$, et alors la composante connexe Ω_x de x dans Ω admet une extrémité x_0 dans $]a, b[$ (parce que $]a, b[$ intersecte X). D'après ce qui précède, f coïncide avec un polynôme P sur Ω_x.

Mais $x_0 \in X \cap]a,b[$ donc d'après ce qui précède, $f^{(n)}(x_0) = 0$ pour tout $n \geqslant n_0$, d'où $\deg(P) < n_0$. En particulier $f^{(n_0)}(x) = 0$.

Dans tous les cas on a $f^{(n_0)} = 0$ sur $]a,b[$, donc $]a,b[\subset \mathring{F}_{n_0}$ ce qui est absurde. Donc X est vide, donc $\Omega = \mathbf{R}$ qui est connexe et f y est polynomiale. □

Remarques.

- Le développement est assez compliqué, donc mieux vaut avoir les idées claires. Voici un résumé des idées de la preuve :
  - f est polynomiale sur chaque composante connexe de Ω parce que l'on peut y propager la propriété locale d'être polynomiale. Il suffit donc de montrer que $\Omega = \mathbf{R}$, c'est-à-dire que X est vide.
  - On montre que X n'a pas de point isolé, un tel point étant forcément le raccordement de deux polynômes donc en fait dans Ω.
  - On montre que X est vide en supposant qu'il ne le soit pas, et en trouvant grâce au théorème de Baire un intervalle $]a,b[$ sur lequel $f^{(n_0)}$ est identiquement nulle mais qui intersecte X. Pour montrer que $f^{(n_0)}$ y est identiquement nulle, on distingue les cas où le point est dans X, et celui où le point est dans Ω en se ramenant au premier cas.
- Il faut passer un peu vite sur certains détails pour pouvoir tout expliquer, en particulier sur l'utilisation du théorème de Rolle. On dispose d'une suite injective (x_p) qui converge vers x, et l'on a $f^{(n_0)}(x_p) = 0$ pour tout p. En passant à la limite en p on trouve $f^{(n_0)}(x) = 0$. Ensuite on peut trouver, en supposant sans perdre de généralité que (x_p) est monotone, un point x'_p situé entre x_p et x_{p+1} où $f^{(n_0+1)}$ s'annule (c'est le théorème de Rolle). On obtient une nouvelle suite (x'_p) qui annule $f^{(n_0+1)}$ et à la limite on obtient $f^{(n_0+1)}(x) = 0$. On recommence par récurrence pour obtenir toutes les dérivées supérieures à n_0 qui sont nulles en x.
- Il faut pouvoir énoncer précisément le théorème de Baire, savoir le démontrer, et savoir l'utiliser dans la situation du développement. Il énonce que dans tout espace métrique complet, toute réunion dénombrable de fermés d'intérieurs vides est encore d'intérieur vide (mais pas forcément fermée !). De même en passant au complémentaire, toute intersection dénombrable d'ouverts denses est encore dense (mais pas forcément ouverte !). On l'utilise dans le développement au sens suivant : si tous les $X \cap F_n$ étaient d'intérieur vide alors leur réunion le serait aussi, or X n'est pas d'intérieur vide (dans lui-même) puisque X est supposé ne pas être vide. Donc il existe un $X \cap F_{n_0}$ dont l'intérieur n'est pas vide. Pour démontrer le théorème de Baire, on choisit la version avec les ouverts denses, et l'on fixe un ouvert ω quelconque, dont on veut qu'il rencontre l'intersection $\bigcap O_n$. Pour cela, on construit par récurrence une suite de points (x_n) et une suite de rayons (η_n) qui décroît au moins géométriquement, tels que $B(x_0, \eta_0) \subset \omega$ et $\overline{B}(x_{n+1}, \eta_{n+1}) \subset B(x_n, \eta_n) \cap O_{n+1}$. La suite (x_n) est alors de Cauchy et la limite est à la fois dans tous les O_n et dans ω.
- On risque de demander une application de ce théorème. Il a été utilisé par Pinkus pour démontrer qu'un réseau de neurones à une couche cachée peut approximer

toute fonction continue (pour la convergence uniforme sur tout compact) si et seulement si la fonction d'activation n'est pas polynomiale. Sinon, on peut aussi dire que les maths peuvent être une fin en soi et demander d'arrêter de toujours demander des applications à tous les théorèmes.

— On risque aussi de demander un contre-exemple si l'on ne suppose plus que l'espace de départ est complet. Comme on utilise plusieurs fois des propriétés de $\mathbf{R}$ il faut par exemple trouver une fonction lisse, admettant une partie dense dans $\mathbf{R}$ où toutes les dérivées sauf un nombre fini s'annulent. La fonction de Fabius est un tel contre-exemple (c'est d'ailleurs une fonction lisse mais nulle part analytique). Maintenant pour une généralisation plus large ça paraît compliqué.

Recasages.

— 204 : On utilise plein de fois la connexité, les composantes connexes, et le fait de pouvoir propager des informations locales. C'est vraiment bien, mais ce n'est pas une application d'un gros théorème de la leçon donc il faut arriver à bien le positionner. Par exemple au niveau de la définition des composantes connexes, ou après le calcul des parties connexes de $\mathbf{R}$.

— 205 : C'est vraiment bien, on utilise la complétude de $\mathbf{R}$ avec une jolie illustration du théorème de Baire. Et on utilise le fait que les fermés dans les espaces complets sont complets (c'est pas grand chose, mais ça fait un lien de plus).

— 223 : On utilise le théorème de Baire qui utilise fortement des suites (numériques ici puisqu'on est dans $\mathbf{R}$) et aussi le fait que X est parfait pour montrer qu'il est vide. C'est un peu léger comme lien, mais c'est probablement assez pour pouvoir mettre le développement dans cette leçon.

— 228 : Ici c'est vraiment bien, ça donne un résultat intuitif mais quand même fort et loin d'être trivial à propos des fonctions lisses sur $\mathbf{R}$. Donc en plein milieu du thème. Et on utilise le théorème de Rolle, et des passages à la limite pour dire que des polynômes sont égaux.

2.2.17 Théorème taubérien de Hardy-Littlewood

Leçons 209, 230, 235, 241, 243

Référence Gourdon Analyse 2e Ed

Prérequis. Théorème de Weierstrass

Théorème 2.2.35. *Soit $(a_n)_{n\in\mathbf{N}}$ une suite de nombres réels en $O(1/n)$. Si $\sum a_n z^n$ a un rayon $\geqslant 1$ et si sa somme F est de limite nulle en 1^-, alors $\sum a_n$ converge et sa somme est nulle.*

Démonstration. On note :

$$\Phi = \left\{ \varphi : [0,1] \to \mathbf{R} \mid \forall x \in [0,1[, \sum a_n \varphi(x^n) \text{ converge, et } \lim_{x\to 1^-} \sum_{n=0}^{+\infty} a_n \varphi(x^n) = 0 \right\}.$$

On veut montrer que l'indicatrice de $[1/2,1]$ est dans Φ. Pour cela, on commence par les fonctions polynomiales, et l'on appliquera le théorème de Weierstrass.

Il est clair que Φ contient toutes les fonctions polynomiales qui s'annulent en 0. Maintenant si q est une fonction polynomiale, on montre que :

$$\lim_{x\to 1^-} (1-x) \sum_{n=0}^{+\infty} x^n q(x^n) = \int_0^1 q.$$

La somme à gauche converge toujours absolument puisque q est bornée. Supposons que q soit monomiale, disons $q(x) = x^k$. Alors pout tout $x \in [0,1[$:

$$(1-x) \sum_{n=0}^{+\infty} x^n q(x^n) = (1-x) \sum_{n=0}^{+\infty} (x^{k+1})^n = \frac{1}{1+x+\cdots+x^k},$$

d'où le résultat en passant à la limite. Par linéarité, c'est donc vrai pour toutes les fonctions polynomiales.

On pose alors $g = \mathbf{1}_{[1/2,1]}$, dont on veut montrer qu'elle appartient à Φ. Pour cela, on fixe $\varepsilon > 0$ et l'on trouve deux polynômes p_1 et p_2 tels que :

1. $p_1(0) = p_2(0) = 0$ et $p_1(1) = p_2(1) = 1$;
2. $p_1 \leqslant g \leqslant p_2$ sur $[0,1]$;
3. $\int_0^1 q < \varepsilon$ avec $q(x) = \frac{p_2(x)-p_1(x)}{x(1-x)}$.

Pour les trouver, posons $h(x) = \frac{g(x)-x}{x(1-x)}$. Pour contrôler la discontinuité, on introduit s_1 et s_2 deux fonctions continues telles que $s_1 \leqslant h \leqslant s_2$, et $\int_0^1 (s_2 - s_1) < \varepsilon$. D'après le théorème de Weierstrass, on peut trouver t_1 et t_2 deux fonctions polynomiales telles que $|s_i - t_i| < \varepsilon$ sur $[0,1]$. On note alors $u_1 = t_1 - \varepsilon$ et $u_2 = t_2 + \varepsilon$ pour avoir :

$$u_1 < s_1 \leqslant h \leqslant s_2 < u_2, \quad \text{et} \quad u_2 - u_1 = t_2 - t_1 + 2\varepsilon \leqslant s_2 - s_1 + 4\varepsilon.$$

Ainsi :

$$\int_0^1 (u_2 - u_1) \leqslant \int_0^1 (s_2 - s_1 + 4\varepsilon) < 5\varepsilon.$$

Comme on vient d'approcher h vérifiant $g(x) = x + x(1-x)h(x)$ pour $x \in [0,1]$, on peut approcher g en posant :

$$p_i = x + x(1-x)u_i(x).$$

Reste à voir en quoi ces deux polynômes nous permettent de conclure que $g \in \Phi$. Pour $x \in [0,1[$ il est clair que $\sum a_n g(x^n)$ converge puisque la somme est finie, il faut donc démontrer l'autre point. Comme $a_n = O(1/n)$, on peut poser $M > 0$ tel que $|a_n| \leqslant M/n$ pour tout $n \geqslant 1$. Pour $x \in [0,1[$ on obtient alors :

$$\begin{aligned}\left|\sum_{n=0}^{+\infty} a_n g(x^n) - \sum_{n=0}^{+\infty} a_n p_1(x^n)\right| &\leqslant \sum_{n=1}^{+\infty} |a_n|(p_2 - p_1)(x^n) \\ &\leqslant M \sum_{n=1}^{+\infty} \frac{x^n(1-x)^n}{n} q(x^n) \\ &\leqslant M(1-x) \sum_{n=1}^{+\infty} x^n q(x^n).\end{aligned}$$

Ce dernier membre tend vers $M\int_0^1 q < M\varepsilon$ donc est plus petit que $M\varepsilon$ lorsque x est assez proche de 1^-, et la somme $\left|\sum_{n=0}^{+\infty} a_n p_1(x^n)\right|$ est plus petite que ε pour $x > \lambda$ car $p_1 \in \Phi$. Finalement pour $x > \lambda$:

$$\left|\sum_{n=0}^{+\infty} a_n g(x^n)\right| \leqslant \varepsilon + M\varepsilon.$$

On fait finalement tendre ε vers 0 pour obtenir $g \in \Phi$.

Et maintenant le théorème est démontré, puisque :

$$\sum_{n=0}^{+\infty} a_n g(x^n) = \sum_{n=0}^{\lfloor -\ln 2/\ln x \rfloor} a_n$$

pour tout $x \in [0,1[$. □

Remarques.

- Le théorème est assez technique à démontrer et difficile à retenir, mais il est assez visuel pour que ce ne soit pas insurmontable.
- Il faut absolument faire des dessins quand on énonce que l'on va définir p_1 et p_2, et faire un dessin pour la construction de h, s_i, t_i et u_i. Et il faut faire ce dessin en prenant de la place.
- Il faut bien annoncer les idées de la démonstration au début, et pendant : elle est technique et on peut vite se perdre dans le raisonnement, alors c'est utile à la fois pour se rappeler du développement et pour que le jury comprenne.

— Si le développement est trop difficile, on peut se contenter du théorème taubérien faible.
— Ce théorème est en général plutôt appelé *théorème taubérien de Littlewood* car c'est lui qui l'a démontré (puis c'est Karamata qui a trouvé la démonstration du développement). Le *théorème taubérien de Hardy-Littlewood* est un peu plus général, et énonce que si a_n est une suite de réels positifs tels que $\sum_{n=0}^{+\infty} a_n \mathrm{e}^{-ny} \sim 1/y$ lorsque $y \to 0+$, alors les sommes partielles des a_k sont équivalentes à n. Une application de ce théorème est le théorème des nombres premiers. On commence par établir que :

$$\sum_{n=2}^{+\infty} \Lambda(n)\mathrm{e}^{-ny} \sim \frac{1}{y},$$

puis l'on en déduit avec le théorème que $\sum_{n\leqslant x} \Lambda(n) \sim x$. Ainsi si le jury demande à quoi peut servir ce théorème, on peut expliquer que c'est le précurseur du théorème plus général, ou bien donner un exemple un peu plus bête, comme le calcul de la somme harmonique alternée qui vaut $-\ln 2$.

Recasages.

— 209 : C'est une très jolie et surprenante application du théorème de Weierstrass, et ce n'est pas la seule approximation que l'on fait dans le développement. On commence par approximer avec des fonctions continues, puis avec des fonctions polynomiales, donc on illustre un théorème de densité des fonctions continues.
— 230 : La leçon ne porte pas sur les séries entières donc il faut bien expliquer que le développement peut servir à calculer des sommes infinies. Par exemple la série harmonique alternée peut s'étudier avec ce théorème.
— 235 : Le théorème est justement un théorème d'interversion limite-somme, qui ne découle pas déjà des théorèmes d'interversion connus. On peut alors citer d'autres théorèmes taubériens et des théorèmes abéliens pour accompagner une partie d'approfondissement.
— 241 : Cette fois c'est les séries entières qui sont dans le thème, et on illustre en particulier leur utilité dans le calcul de sommes infinies. On fait aussi apparaître la série entière géométrique qui est probablement la plus utile de toutes.
— 243 : cf. 241.

2.2.18 Trois droites de Hadamard / Riesz-Thorin

Leçons 201, 206, 208, 219, 234, 245, 253, 267

Référence Zuily-Queffélec

Prérequis. Principe du maximum, inégalité de Hölder

Théorème 2.2.36 (des trois droites de Hadamard). *On note $\mathcal{B}$ la bande des complexes de partie réelle comprise (strictement) entre 0 et 1. Soit $f : \overline{\mathcal{B}} \to \mathbf{C}$ une fonction continue sur $\overline{\mathcal{B}}$, bornée sur $\partial\mathcal{B}$, et holomorphe sur $\mathcal{B}$. Alors en posant $M(\theta) = \sup_{\mathfrak{Re}(z)=\theta} |f(z)|$, on a l'inégalité de log-convexité :*

$$M(\theta) \leqslant M(0)^{1-\theta} M(1)^{\theta}$$

pour tout $\theta \in [0,1]$.

Démonstration. Soient $\varepsilon > 0$ et $\lambda \in \mathbf{R}$ quelconques, et :

$$f_{\varepsilon,\lambda}(z) = \mathrm{e}^{\varepsilon z^2 + \lambda z} f(z).$$

Choisissons un $R > 0$ assez grand pour que le raisonnement suivant soit valide. On applique le principe du maximum à $f_{\varepsilon,\lambda}$ qui est holomorphe sur le rectangle $\mathcal{R} = [0,1] + \mathrm{i}[-R,R]$. Si R a été choisi assez grand (et si f n'est pas identiquement nulle) alors on peut trouver z_0 à l'intérieur de $\mathcal{R}$ tel que $f_{\varepsilon,\lambda}(z_0) \neq 0$. Par le principe du maximum, $f_{\varepsilon,\lambda}$ atteint son maximum sur le bord de $\mathcal{R}$. Montrons que cela ne peut pas se faire sur les bords en haut et en bas.

Si c'est le cas, alors pour $z = x \pm \mathrm{i}R$ on aurait :

$$|f_{\varepsilon,\lambda}(z)| = \mathrm{e}^{\varepsilon(x^2-R^2)+\lambda x}|f(z)| \leqslant \mathrm{e}^{\varepsilon+|\lambda|-\varepsilon R^2}\|f\|_{\infty,\mathcal{R}} < |f_{\varepsilon,\lambda}(z_0)|$$

pour un R choisi assez grand, ce qui est absurde. Donc le maximum de $f_{\varepsilon,\lambda}$ doit être atteint sur un des murs. On obtient :

$$|f_{\varepsilon,\lambda}(z)| \leqslant \max\left(\sup_{|t|\leqslant R} \mathrm{e}^{-\varepsilon t^2}|f(\mathrm{i}t)|, \sup_{|t|\leqslant R} \mathrm{e}^{\varepsilon(1-t^2)+\lambda}|f(1+\mathrm{i}t)|\right).$$

En notant $z = \theta + \mathrm{i}y$ on a alors :

$$|f(z)| \leqslant \max\left(\sup_{|t|\leqslant R} \mathrm{e}^{-\varepsilon(\theta^2-y^2)-\lambda\theta-\varepsilon t^2}|f(\mathrm{i}t)|, \sup_{|t|\leqslant R} \mathrm{e}^{-\varepsilon(\theta^2-y^2)-\lambda\theta+\varepsilon(1-t^2)+\lambda}|f(1+\mathrm{i}t)|\right).$$

On fait finalement tendre R vers $+\infty$ et ε vers 0 pour obtenir :

$$|f(z)| \leqslant \max\left(\mathrm{e}^{-\lambda\theta}M(0), \mathrm{e}^{\lambda(1-\theta)}M(1)\right),$$

ce qui donne le résultat en posant $\mathrm{e}^{-\lambda} = M(1)/M(0)$. □

Théorème 2.2.37 (de Riesz-Thorin). *Soient $1 \leqslant p_0, p_1, q_0, q_1 \leqslant +\infty$ quatre exposants, $(E, \mathcal{A}_E, \mu)$ et $(F, \mathcal{A}_F, \nu)$ deux espaces mesurés σ-finis, et T un opérateur linéaire continu de $\mathrm{L}^{p_0}(E)$ dans $\mathrm{L}^{q_0}(F)$, qui est aussi un opérateur linéaire continu de $\mathrm{L}^{p_1}(E)$ dans $\mathrm{L}^{q_1}(F)$. Alors pour tout $\theta \in [0,1]$, en posant :*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad et \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1},$$

l'opérateur T est aussi linéaire continu de $\mathrm{L}^p(E)$ dans $\mathrm{L}^q(F)$. De plus, en notant $M(\theta) = \|T\|_{\mathrm{L}^p(E)\to\mathrm{L}^q(F)}$, on dispose de l'inégalité de log-convexité :

$$M(\theta) \leqslant M(0)^{1-\theta} M(1)^{\theta}.$$

Démonstration. On se fiche des cas limites qui sont automatiques, on suppose donc que $1 < p, q < +\infty$ et que $\theta \in]0,1[$. Remarquons que pour $f \in \mathrm{L}^q(F)$, on peut par dualité calculer sa norme de la manière suivante :

$$\|f\|_q = \sup_{\|g\|_{q'}=1} \int_F f\overline{g}\, \mathrm{d}\nu$$

où q' est l'exposant conjugué de q. On montre alors simplement que pour $\|f\|_{\mathrm{L}^p(E)} = 1$ et pour $\|g\|_{\mathrm{L}^{q'}(F)} = 1$, on a :

$$\int_F Tf \cdot \overline{g}\, \mathrm{d}\nu \leqslant M(0)^{1-\theta} M(1)^{\theta}.$$

Pour cela, on applique le théorème des trois droites à une fonction bien choisie. On pose pour $z \in \mathcal{B}$:

$$\frac{1}{p(z)} = \frac{1-z}{p_0} + \frac{z}{p_1}, \qquad \frac{1}{q(z)} = \frac{1-z}{q_0} + \frac{z}{q_1},$$
$$\frac{1}{p'(z)} = \frac{1-z}{p'_1} + \frac{z}{p'_2}, \qquad \frac{1}{q'(z)} = \frac{1-z}{q'_1} + \frac{z}{q'_2},$$

de sorte à avoir toujours $p(z)$ et $p'(z)$ (resp. $q(z)$ et $q'(z)$) qui sont des exposants conjugués, et à avoir $p(\theta) = p$ et $q(\theta) = q$. On pose alors :

$$F(z) : x \in E \mapsto |f(x)|^{p/p(z)} \frac{f(x)}{|f(x)|},$$
$$G(z) : y \in F \mapsto |g(y)|^{q'/q'(z)} \frac{g(y)}{|g(y)|},$$
$$\varphi(z) = \int_F TF(z) \cdot \overline{G(z)}\, \mathrm{d}\nu,$$

en posant $F(z)(x) = 0$ (resp. $G(z)(y) = 0$) lorsque $f(x)$ (resp. $g(y)$) est nul. On a alors $F(\theta) = f$ et $G(\theta) = g$, et il suffit donc de montrer que :

$$\varphi(\theta) \leqslant M(0)^{1-\theta} M(1)^{\theta}.$$

On applique le théorème des trois droites à φ. On obtient $\varphi(\theta) \leqslant N(0)^{1-\theta}N(1)^{\theta}$ avec $N(t) = \sup_{\mathfrak{Re}(z)=t} |\varphi(z)|$. On a juste à montrer que $N(0) \leqslant M(0)$ et que $N(1) \leqslant M(1)$. Montrons-le par exemple pour 0. Si $\mathfrak{Re}(z) = 0$, alors par Hölder :

$$|\varphi(z)| = \left| \int_F TF(z) \cdot \overline{G(z)} \, \mathrm{d}\nu \right| \leqslant \|TF(z)\|_{q_0} \|G(z)\|_{q_0'} \leqslant M(0)\|F(z)\|_{p_0}\|G(z)\|_{q_0'}.$$

Or, comme $\mathfrak{Re}(z) = 0$ on a $\mathfrak{Re}(p/p(z)) = p/p_0$ d'où :

$$\|F(z)\|_{p_0}^{p_0} = \int_E |F(z)|^{p_0} \, \mathrm{d}\mu = \int_E |f|^p \, \mathrm{d}\mu = 1$$

et de même $\|G(z)\|_{q_0'} = 1$. Cela conclut.

Pour pouvoir utiliser le théorème des trois droites, il faudrait vérifier que φ (qui est clairement continue sur $\overline{\mathcal{B}}$ et bornée sur $\partial\mathcal{B}$) est holomorphe sur $\mathcal{B}$. Mais en fait il n'y a pas besoin de le vérifier pour le cas général. Avec un passage à la limite, il suffit de le vérifier pour le cas où f et g sont étagées à supports de mesure finie, et dans ce cas $\varphi(z)$ est une combinaison linéaire d'intégrales du type $\int_F T\mathbf{1}_{A_i} \cdot \overline{\mathbf{1}_{B_j}} \, \mathrm{d}\nu$, qui sont de la forme $\alpha\beta^z$ donc holomorphes. □

Remarques.

— Le développement est clairement beaucoup trop long pour quinze minutes, il faut alors aller vite sur certains points. Par exemple on peut s'éviter d'écrire toutes les marjorations dans la démonstration des trois droites, en disant à l'oral ce que l'on fait. Les détails pourront être écrits lors des questions s'il le faut. De même pour l'application du théorème des trios droites, pour la dualité, et pour la conclusion.

— Il y a des applications directes de ce théorème même au niveau de l'agreg. Par exemple on peut prendre pour T la transformée de Fourier qui est à la fois définie $\mathrm{L}^1 \to \mathrm{L}^\infty$ et $\mathrm{L}^2 \to \mathrm{L}^2$. On obtient une transformée de Fourier sur L^p pour tout $p \in [1, 2]$, à valeurs dans L^q avec un certain $q \in [2, +\infty]$. On peut même majorer sa norme d'opérateur avec l'inégalité de log-convexité, mais elle n'est atteinte que pour les deux transformées de Fourier que l'on connaît déjà. On peut aussi l'appliquer à l'opérateur des coefficients de Fourier $\mathrm{L}^2 \to \ell^2$ et $\mathrm{L}^1 \to \ell^\infty$, pour obtenir l'*inégalité de Hausdorff-Young* :

$$\left(\sum_{n=-\infty}^{+\infty} |\widehat{f}(n)|^q \right)^{1/q} \leqslant \left(\frac{1}{2\pi} \int_0^{2\pi} |f(t)|^p \, \mathrm{d}t \right)^{1/p},$$

pour $1 \leqslant p \leqslant 2$ et q son exposant conjugué. Une dernière application facile est la convolution par une fonction L^1, qui définit un opérateur $\mathrm{L}^1 \to \mathrm{L}^1$ mais aussi $\mathrm{L}^\infty \to \mathrm{L}^\infty$.

— Pour pouvoir écrire $\int Tf \cdot \overline{g} \, \mathrm{d}\nu$, on suppose déjà que Tf est dans L^q, donc que T définit un opérateur linéaire de L^p dans L^q. En fait cette partie est facile, puisque une fois T défini sur L^{p_0} et sur L^{p_1}, on le définit sur les L^p par densité. Si on hésite, on peut se contenter de démontrer tout le résultat pour f et g continues

à support compact, avec les espaces mesurés qui sont des ouverts de $\mathbf{R}^n$ avec la mesure de Lebesgue.
— Attention au livre de Zuily et Queffélec qui fait des grands détours. Il vaut mieux énoncer le théorème dans le cas d'ouverts de $\mathbf{R}^n$ avec la mesure de Lebesgue et faire la démonstration simple plutôt que d'essayer d'adapter la preuve du livre.
— Le théorème des trois droites a un analogue circulaire qui en est aussi un corollaire (en considérant l'exponentielle de f) : on remplace la bande par la couronne $\{r < |z| < R\}$, et le supremum sur les droites verticales devient un supremum sur les cercles intermédiaires ; le reste de l'énoncé est le même.
— Attention le théorème est d'un haut niveau, et le jury risque de poser des questions sur l'interpolation des espaces de fonctions. Le théorème de Riesz-Thorin est historiquement le point de départ de l'interpolation des espaces de Banach, et énonce que L^p est un espace d'interpolation entre L^{p_0} et L^{p_1} pour tout $p \in [p_0, p_1]$. La théorie de l'interpolation a été appliquée aux espaces L^p, aux espaces de Sobolev, aux espaces de Besov, ...

Recasages.
— 201 : Le théorème est à la base de l'interpolation des espaces de Banach, qui sont presque toujours des espaces de fonctions. Donc c'est dans le thème, attention à ne pas faire une leçon trop hétérogène.
— 206 : J'avais besoin d'un deuxième développement dans cette leçon qui est mon impasse. Mon idée est que l'on utilise le théorème des trois droites, qui est un théorème en dimension un (ou deux), pour énoncer un théorème en dimension infinie, utile en analyse. L'analogie s'arrête ici, donc je ne conseille pas vraiment.
— 208 : C'est très bien, surtout si l'on fait une partie sur les espaces L^p, ce qui est incontournable. En plus on utilise des techniques de dualité pour calculer des normes, ce qui est apprécié.
— 219 : C'est surtout pour le théorème des trois droites, qui est un principe du maximum. C'est vraiment une bonne idée de faire une partie sur les principes du maximum en analyse complexe, et le théorème de Riesz-Thorin en est une application spectaculaire. Comme le développement est trop long, on pourrait presque se contenter du théorème des trois droites, et si jamais il reste du temps, rajouter celui des trois cercles ou expliquer rapidement la démonstration de Riesz-Thorin.
— 234 : C'est comme pour la leçon 208, ça rentre bien. En plus ici on peut l'appliquer à tous les opérateurs de Fourier que l'on connaît.
— 245 : Pour le théorème des trois droites, et les principes du maximum en général (comme pour la leçon 219). Le théorème de Riesz-Thorin est une application qui ne doit pas prendre trop de place dans le développement : on s'attarde surtout sur les questions d'analyse complexe du début.
— 253 : Le théorème en lui-même n'est pas une utilisation de la convexité à proprement parler, mais les corollaires en sont. On peut par exemple démontrer les trois droites et les trois cercles, passer rapidement sur Riesz-Thorin, et donner

des corollaires au théorème pour expliquer en quoi l'inégalité de log-convexité est intéressante. Mais ça reste un peu limite.

— 267 : Ici par contre le développement rentre très bien ! Les courbes dans le plan complexe sont un peu trop souvent reléguées au rang de chemins d'intégration, alors que l'on peut les utiliser pour démontrer des résultats beaucoup plus forts sur des espaces L^p. En plus, on *utilise* réellement ces droites.

2.3 Développements mixtes

2.3.1 Composantes connexes des formes quadratiques

Leçons 170, 171, 204

Référence Oraux Algèbre 3

Prérequis. Loi d'inertie de Sylvester.

Théorème 2.3.1. *Soit E un espace vectoriel réel de dimension finie $n \geqslant 1$. On note $Q(E)$ l'espace des formes quadratiques sur E et $\Omega(E) \subset Q(E)$ la partie formée des formes quadratiques non dégénérées. Alors les composantes connexes de $\Omega(E)$ sont les :*

$$\Omega_k(E) = \{q \in \Omega(E) \mid \sigma(q) = (k, n-k)\}.$$

Démonstration. L'énoncé de ce théorème dépend *a priori* de la topologie de E, mais on choisit bien sûr la topologie induite par n'importe quelle norme $\|-\|$, que l'on fixe pour la suite.

Soit $q \in \Omega(E)$, de signature (r, s). Dans une base orthogonale pour q, on lit que l'on peut trouver deux sous-espaces supplémentaires F et G de dimensions respectives r et s, tels que les restrictions de q à F et à G soient respectivement définie positive et définie négative. Sur F, la fonction $\sqrt{q}$ est une norme euclidienne, donc par équivalence des normes il existe k_1 tel que $q(x) \geqslant k_1\|x\|^2$ pour tout $x \in F$. De même, $\sqrt{-q}$ est une norme euclidienne sur G, et il existe k_2 tel que $-q(x) \geqslant k_2\|x\|^2$ pour tout $x \in G$. On remplace alors k_1 et k_2 par le maximum des deux, de sorte à disposer d'un réel $k > 0$ tel que :

$$\forall x \in F,\ q(x) \geqslant k\|x\|^2, \quad \text{et} \quad \forall x \in G,\ q(x) \leqslant -k\|x\|^2.$$

Modulo le choix d'une base, l'espace $Q(E)$ (resp. $\Omega(E)$) est en bijection avec l'espace des matrices symétriques réelles (resp. matrices symétriques réelles inversibles). La topologie sur cet espace est donc celle induite par n'importe quelle norme, et l'on choisit la suivante :

$$N(q') = \sup_{\|x\|=1} |q'(x)|.$$

On va montrer que pour cette norme, si q' est assez proche de q alors les deux formes ont la même signature.

Soit alors $q' \in Q(E)$ telle que $N(q'-q) < k$. Pour tout vecteur non nul $x \in E$, on a alors :

$$|q'(x) - q(x)| < k\|x\|^2.$$

Ainsi pour $x \in F$ non nul, l'inégalité triangulaire donne $q'(x) > q(x) - k\|x\|^2 \geqslant 0$ et pour $x \in G$ non nul, de même $q'(x) < q(x) + k\|x\|^2 \leqslant 0$. Donc q' est définie positive sur F et définie négative sur G, donc elle est de même signature que q.

On peut maintenant conclure. On vient de montrer que chaque $\Omega_k(E)$ est ouvert, mais ils partitionnent $\Omega(E)$ par Sylvester. Il reste donc simplement à montrer que chaque $\Omega_k(E)$ est connexe.

Soient A et B deux matrices symétriques de même signature $(k, n-k)$. En posant D la diagonale formée de k fois 1 puis $n-k$ fois -1, on dispose donc de P, Q orthogonales telles que $A = P^\top DP$ et $B = Q^\top DQ$. Quitte à remplacer la première ligne de P ou Q par son opposée, on peut supposer que ces deux matrices sont de déterminant > 0. On choisit alors un arc qui relie P à Q dans l'espace connexe par arcs des matrices réelles de déterminant > 0, ce qui fournit un arc reliant A à B dans l'espace des matrices symétriques réelles de signature $(k, n-k)$. □

Remarques.

— L'avantage est que la démonstration est très intuitive donc facile à retenir. On passe beaucoup de temps à montrer que les $\Omega_k(E)$ sont ouverts et très peu à démontrer qu'ils sont connexes, ce qui peut sembler dommage.

— Bien sûr il faut savoir expliquer pourquoi les matrices réelles de déterminant > 0 forment un espace connexe par arcs. Pour cela, il faut se rappeler que le groupe linéaire est engendré par les transvections et les dilatations.

Recasages.

— 170 : Bof, parce que c'est seulement le cas réel. Ça permet de remplir la leçon si vraiment on ne sait pas quoi y mettre, mais c'est un peu faiblard.

— 171 : C'est déjà mieux. Ce n'est pas mon développement préféré, mais j'avais besoin d'un deuxième développement dans cette leçon. Il illustre quand même un peu les différentes formes quadratiques sur un espace réel et la loi d'inertie de Sylvester, ce qui n'est pas si mal. Attention à ne pas oublier les coniques par contre !

— 204 : Le développement est assez bon pour cette leçon. On y utilise la connexité de $\mathrm{GL}_n^+(\mathbf{R})$ et on montre que l'on sait ce que sont des composantes connexes. En plus, c'est dans un espace assez abstrait et ça fait des liens avec l'algèbre. Le jury aime ce genre de liens avec les autres domaines.

2.3.2 Convergence d'une suite de polygones vers l'isobarycentre

Leçons 149, 152, 181, 226

Référence Gourdon algèbre

Prérequis. Aucun.

Théorème 2.3.2. *Soient $z_1^{(0)}, \dots, z_n^{(0)} \in \mathbf{C}$ des points du plan complexe définissant un polygone P_0. Pour tout $k \in \mathbf{N}$, on définit :*

$$z_i^{(k+1)} = \frac{z_i^{(k)} + z_{i+1}^{(k)}}{2}$$

en convenant que $z_{n+1} = z_1$. Chaque n-uplet $(z_1^{(k)}, \dots, z_n^{(k)})$ définit le polygone P_k. Alors la suite (P_k) converge vers le polygone dégénéré concentré en l'isobarycentre de P_0.

Démonstration. Notons $z^{(k)} = (z_1^{(k)}, \dots, z_n^{(k)})$ pour tout $k \in \mathbf{N}$. La définition de la suite $(z^{(k)})_{k\in\mathbf{N}}$ se réécrit :

$$z^{(k+1)} = Az^{(k)} = A^k z^{(0)}, \text{ avec } A = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & \dots & 0 \\ 0 & 1 & 1 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

On étudie alors les valeurs propres de la matrice A. On pose :

$$J = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}$$

qui est un bloc de Jordan auquel on a rajouté un 1 en bas à gauche. Pour tout polynôme $P = \sum_{k=0}^{n-1} a_k X^k$, on a :

$$P(J) = \begin{pmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \dots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \dots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \dots & a_0 \end{pmatrix}.$$

En particulier $J^n = I_n$ donc J se diagonalise en $D = \operatorname{diag}(\omega^j : 0 \leqslant j < n)$ avec $\omega = \exp(\frac{2\mathrm{i}\pi}{n})$. Ainsi, $P(J)$ se diagonalise en $P(D)$ et donc $\det P(J) = \prod_{j=0}^{n-1} P(\omega^j)$.

Ici, on a $A = \frac{1}{2}(I_n + J)$, et plus généralement $XI_n - A = (X - \frac{1}{2})I_n - \frac{1}{2}J$ d'où :

$$\chi_A(X) = \prod_{j=0}^{n-1} \left(X - \frac{1+\omega^j}{2} \right).$$

Ce polynôme est à racines simples donc A se diagonalise avec les valeurs propres écrites ci-dessus, et donc A^k converge (géométriquement, pour n'importe quelle norme d'algèbre donc pour n'importe quelle norme) vers une matrice B qui est conjuguée à $\mathrm{diag}(1, 0, \ldots, 0)$. Comme $X = (1, \ldots, 1)$ est point fixe de A et donc à la limite point fixe de B qui est de rang 1, on a $B = (b_1 X, \ldots, b_n X)$. Comme A préserve les isobarycentres, à la limite B aussi donc chaque b_i vaut $1/n$, ce qui conclut. □

Remarques.

— On peut, au lieu de définir chaque point comme étant le milieu d'un segment, le définir comme étant une combinaison convexe de ces deux points avec un coefficient $\lambda \in]0, 1[$. On a démontré le cas $\lambda = 1/2$, mais le résultat reste vrai pour tout λ.

— Il ne faut pas hésiter à commencer le développement en dessinant un polygone et en itérant deux ou trois fois la construction. Ça donne l'idée intuitive derrière la démonstration.

Recasages.

— 149 : On calcule de manière exacte des valeurs propres, et même toutes les valeurs propres de toutes les matrices circulantes. Et l'on voit aussi comment les valeurs propres influencent le comportement asymptotique de la suite (A^k), ce qui bien sûr est central dans cette leçon.

— 152 : On obtient un calcul de déterminant très élégant car il ne demande pas de gros calculs, juste un peu de réduction. Et l'on obtient une conséquence géométrique très intuitive, ce qui peut adoucir le côté trop abstrait de la leçon.

— 181 : Dans cette leçon le développement rentre parfaitement, encore mieux si l'on remplace le $1/2$ par un $\lambda \in]0, 1[$ quelconque fixé.

— 226 : La suite $(z^{(k)})$ est une suite vectorielle définie par une relation du type $z^{(k+1)} = f(z^{(k)})$, ici f est linéaire. On peut en particulier faire le lien (fort) entre ce développement discret et l'étude qualitative des équations différentielles linéaires, où le spectre de la matrice détermine le comportement asymptotique des solutions.

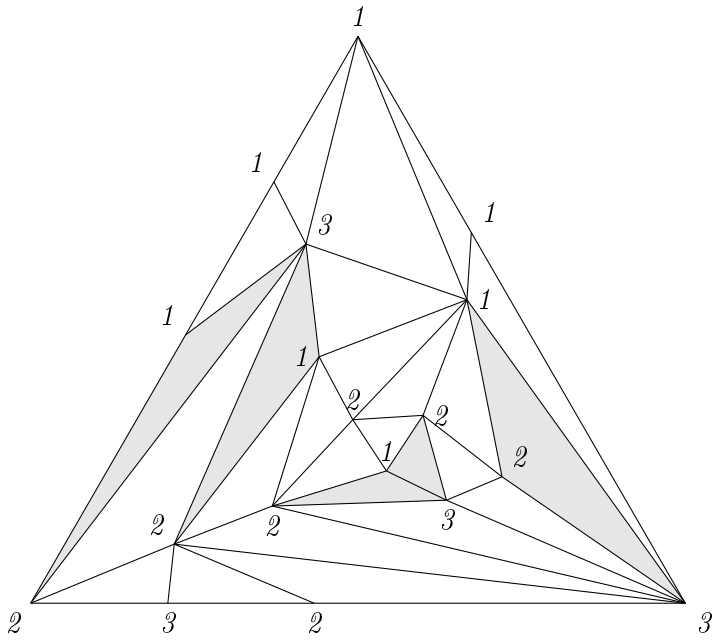
2.3.3 Lemme de Sperner + Théorème de Brouwer

Leçons 190, 203, 253

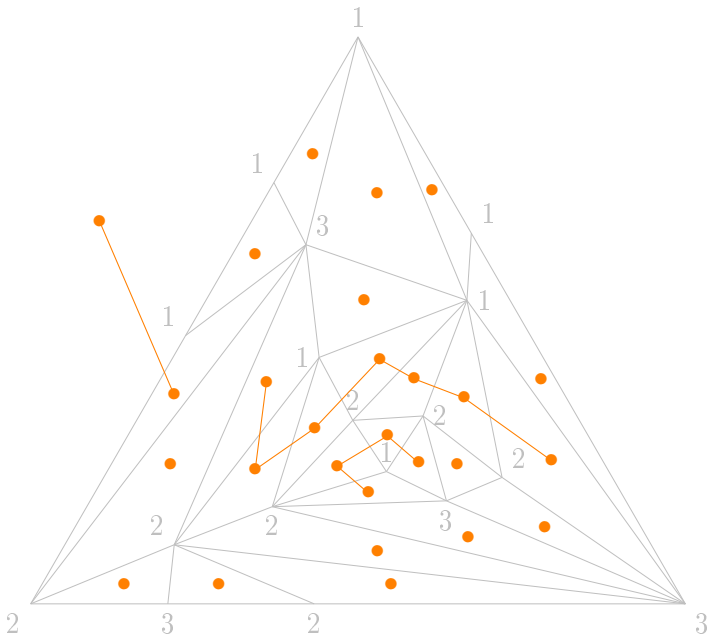
Référence Raisonnements divins

Prérequis. Aucun.

Lemme 2.3.3 (de Sperner). *Soient $P_1P_2P_3$ un triangle et $\mathcal{T}$ une triangulation de celui-ci. Pour tout coloriage $c : \mathcal{T}_{sommets} \to \{1,2,3\}$ des sommets de $\mathcal{T}$ en trois couleurs avec $c(P_i) = i$ et $c(A) \in \{i,j\}$ pour tout sommet $A \in [P_iP_j]$, il existe un petit triangle de $\mathcal{T}$ qui est tricolore.*



Démonstration. On considère le graphe dual de la triangulation, c'est-à-dire le graphe avec un sommet pour chaque petit triangle, un sommet disposé hors du triangle (représentant la face extérieure), et une arête entre deux sommets lorsque les faces correspondantes partagent une arête commune dans la triangulation originale. De ce graphe dual, on ne garde que les arêtes qui coupent des anciennes arêtes coloriées $1-2$ ou $2-1$:



On examine alors les degrés des sommets. Un sommet intérieur de ce graphe a un degré valant :

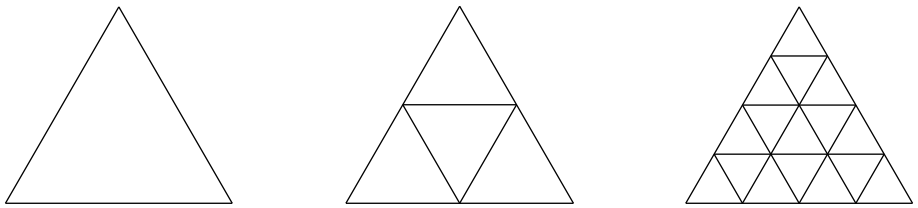
— 0 lorsque le triangle correspondant n'affiche pas les deux couleurs 1 et 2 ;
— 1 lorsque le triangle correspondant est tricolore ;
— 2 lorsque le triangle correspondant n'affiche pas la couleur 3 ;
— 3 dans aucun cas.

Le degré du sommet extérieur est toujours impair car il y a un nombre impair de changements entre 1 et 2 le long du segment $[P_1P_2]$. Or, la somme de tous les degrés des sommets d'un graphe vaut deux fois le nombre d'arêtes, donc est paire. Ainsi, la somme des degrés des sommets intérieurs est impaire, donc il y a un nombre impair de triangles tricolores (qui sont les seuls avec un degré impair). □

On se sert du lemme purement combinatoire de Sperner pour obtenir une démonstration du théorème de Brouwer en dimension deux.

Théorème 2.3.4 (de Brouwer)**.** *Toute application continue* $\mathbf{B}^2 \to \mathbf{B}^2$ *admet un point fixe.*

Démonstration. Soient (e_1, e_2, e_3) la base canonique de $\mathbf{R}^3$ et $\Delta = \mathrm{Conv}(e_1, e_2, e_3)$ le triangle standard. Comme Δ et $\mathbf{B}^2$ sont homéomorphes, il suffit de démontrer le résultat pour une application continue $f : \Delta \to \Delta$, dont on suppose désormais qu'elle n'admet pas de point fixe. Étant donnée une triangulation $\mathcal{T}$ de Δ, on note $\delta(\mathcal{T})$ la longueur maximale de ses arêtes. On peut construire une suite $(\mathcal{T}_n)_{n\in\mathbf{N}}$ de triangulations de Δ telle que $\delta(\mathcal{T}_n) \to 0$, comme suit :



Étant fixé un entier $n \in \mathbf{N}$, on colorie les sommets de la triangulation $\mathcal{T}_n$ de la manière suivante. Pour tout sommet v, la différence $f(v) - v$ n'est pas nulle et appartient à l'hyperplan $\{x + y + z = 0\}$ donc l'une de ses coordonnées v_i est strictement négative ; on colorie alors v avec la plus petite telle couleur i. Notons que si $v \in [e_1e_2]$ alors $v_3 = 0$ donc $(f(v) - v)_3 \geqslant 0$ et v n'est pas colorié avec la couleur 3. De même pour $[e_1e_3]$ et $[e_2e_3]$. En particulier, e_i est bien colorié avec la couleur i, et la coloration de $\mathcal{T}_n$ ainsi construite vérifie les hypothèses du lemme de Sperner. Il existe alors un petit triangle tricolore $(v^{n,1}, v^{n,2}, v^{n,3})$, le sommet $v^{n,i}$ étant de couleur i. Comme Δ est compact, à extraction près la suite $(v^{n,1})_{n\in\mathbf{N}}$ converge vers un $v \in \Delta$. Comme $\|v^{n,2} - v\| \leqslant \|v^{n,2} - v^{n,1}\| + \|v^{n,1} - v\| \leqslant \delta(\mathcal{T}_n) + \|v^{n,1} - v\| \to 0$, la suite $(v^{n,2})_{n\in\mathbf{N}}$ et de même la suite $(v^{n,3})_{n\in\mathbf{N}}$ convergent aussi vers v. Puisque f est continue, on a $(f(v) - v)_i \leqslant 0$ pour tout i. Cela est impossible puisque $f(v) - v$ est dans l'hyperplan $\{x + y + z = 0\}$ et est supposé non nul. □

Remarques.

— Il faut savoir démontrer le fait que la somme de tous les degrés d'un graphe vaut deux fois le nombre d'arêtes. Si V est l'ensemble des sommets et E l'ensemble des arêtes, on peut par exemple dire que d'une part :

$$\begin{aligned}|\{(v,e) \in V \times E \mid v \text{ est un sommet de } e\}| &= \left|\coprod_{e\in E} \{v \in V \mid v \text{ est un sommet de } e\}\right| \\ &= \sum_{e\in E} 2 = 2|E|,\end{aligned}$$

mais que d'autre part :

$$\begin{aligned}|\{(v,e) \in V \times E \mid v \text{ est un sommet de } e\}| &= \left|\coprod_{v\in V} \{e \in E \mid v \text{ est un sommet de } e\}\right| \\ &= \sum_{v\in V} \deg e.\end{aligned}$$

— On justifie toute la fin du raisonnement à extraction près. On pourrait croire que c'est une extraction diagonale où l'on compose trois extractions, mais on n'a besoin d'extraire que pour $(v^{n,1})$, la même extraction fait directement converger $(v^{n,2})$ et $(v^{n,3})$.

— La méthode se généralise parfaitement à toutes les dimensions. On peut définir une triangulation de Sperner d'un tétraèdre par exemple, en imposant à la coloration sur chaque face d'être une coloration de Sperner (de dimension 2). Et ainsi de suite. Et le théorème de Brouwer s'en déduit tout pareillement !

— Le théorème de Brouwer peut aussi se déduire d'un autre théorème combinatoire : il n'y a jamais de match nul au jeu de Hex.
— Le lemme de Sperner donne un algorithme assez efficace en pratique pour calculer des approximations de points fixes. Il suffit d'évaluer la fonction sur une triangulation, en déduire la coloration de Sperner, trouver un triangle tricolore, et recommencer jusqu'à avoir un triangle assez petit à notre goût.
— On déduit aussi du lemme de Sperner le théorème de Monsky : on ne peut pas partitionner un carré en un nombre impair de triangles de même aire. Les idées de la démonstration sont sur la page Wikipédia du théorème.
— Le théorème de Brouwer a énormément d'applications. Par exemple, l'existence d'équilibres de Nash en théorie des jeux, ou le théorème de Jordan.

Recasages.
— 190 : C'est la leçon idéale, une méthode combinatoire pour démontrer un théorème utile en analyse, c'est super cool. On peut faire une sous-partie de théorie des graphes pour mettre le lemme de Sperner à l'aise, mais c'est pas obligé.
— 203 : On utilise vraiment la compacité du triangle (ou en dimension supérieure, du simplexe standard) pour déduire le théorème de Brouwer. Donc ça rentre tout pile.
— 253 : On utilise la convexité pour définir le triangle (ou en dimension supérieure, le simplexe standard). En quelque sorte on utilise le fait que l'hyperplan $\{x+y+z = 1\}$ est convexe, mais c'est un peu tiré par les cheveux.

2.3.4 Loi des cycles d'une permutation aléatoire

Leçons 104, 105, 261, 262, 264

Référence

Prérequis. Décomposition en produits de cycles à supports disjoints, moments factoriels d'une loi de Poisson.

Théorème 2.3.5. *Pour $n \geqslant 1$, on considère π_n une variable aléatoire de loi uniforme sur le groupe symétrique $\mathfrak{S}_n$. Pour $1 \leqslant j \leqslant n$, on note $c_j(\pi_n)$ le nombre de cycles de longueur j qui apparaissent dans la décomposition de π_n en produit de cycles à supports disjoints. Alors lorsque n tend vers l'infini, $c_j(\pi_n)$ converge en loi vers une loi de Poisson de paramètre $1/j$.*

Démonstration. On commence par déterminer la loi de $c(\pi_n)$.

Lemme 2.3.6. *Pour $k \in \mathbf{N}^n$, on a :*

$$\mathbb{P}(c(\pi_n) = k) = \left(\prod_{j=1}^{n} \left(\frac{1}{j}\right)^{k_j} \frac{1}{k_j!} \right) \mathbf{1}_{\sum_{j=1}^n jk_j = n}.$$

Démonstration. On remarque déjà que k est la structure des cycles d'une permutation de $\mathfrak{S}_n$ si et seulement si $\sum_{j=1}^n jk_j = n$. On suppose cette condition vérifiée et l'on va compter le nombre N_k de permutations qui ont cette structure, c'est-à-dire le cardinal de la classe de conjugaison C_σ d'un élément σ qui a cette structure. Soit Z_σ le centralisateur de σ, on a :

$$N_k |Z_\sigma| = |\mathfrak{S}_n| = n!.$$

Pour qu'un élément commute avec σ, il faut et il suffit :

- soit qu'il décale les éléments des cycles, il y a ℓ manières de faire cela pour chaque cycle de longueur ℓ, donc un total de $\prod_{j=1}^n j^{k_j}$ choix ;
- soit qu'il permute les cycles de longueurs identiques entre eux, soit un total de $\prod_{j=1}^n k_j!$ choix.

Le cardinal du centralisateur est alors $|Z_\sigma| = \prod_{j=1}^n j^{k_j} k_j!$, d'où :

$$\mathbb{P}(c(\pi_n) = k) = \frac{N_k}{n!} = \prod_{j=1}^{n} \frac{1}{j^{k_j} k_j!}.$$

□

Pour $x \in \mathbf{R}$ et r un entier strictement positif, on pose $x^{[r]} = x(x-1)\cdot\ldots\cdot(x-r+1)$. On peut alors calculer les moments factoriels :

Lemme 2.3.7. *Pour tout $m \in \mathbf{N}^n$, on a :*

$$\mathbb{E}\left[\prod_{j=1}^{n} c_j(\pi_n)^{[m_j]}\right] = \left(\prod_{j=1}^{n} \left(\frac{1}{j}\right)^{m_j}\right) \mathbf{1}_{\sum_{j=1}^n jm_j \leqslant n}.$$

Démonstration. Si $\sum_{j=1}^n jm_j > n$, alors quelle que soit la valeur de π_n, il existe j_0 tel que $m_{j_0} > c_{j_0}(\pi_n)$ et alors $c_{j_0}(\pi_n)^{[m_{j_0}]} = 0$ donc le produit est nul. Supposons que ce ne soit pas le cas. On calcule alors :

$$\begin{aligned}
\mathbb{E}\left[\prod_{j=1}^{n} c_j(\pi_n)^{[m_j]}\right] &= \sum_{\pi\in\mathfrak{S}_n} \mathbb{P}(\pi_n = \pi) \prod_{j=1}^{n} c_j(\pi)^{[m_n]} \\
&= \sum_{k\in\mathbf{N}^n} \frac{1}{n!} \sum_{c(\pi)=k} \prod_{j=1}^{n} k_j^{[m_j]} \\
&= \sum_{k\in\mathbf{N}^n} \prod_{j=1}^{n} \frac{k_j^{[m_j]}}{j^{k_j} k_j!} \\
&= \prod_{j=1}^{n} \frac{1}{j^{m_j}} \left(\sum_{k\in\mathbf{N}^n} \prod_{j=1}^{n} \frac{1}{j^{k_j-m_j}(k_j-m_j)!}\right) \\
&= \prod_{j=1}^{n} \frac{1}{j^{m_j}} \left(\sum_{k\in\mathbf{N}^n} \mathbb{P}(c(\pi_n) = (k_1-m_1, \ldots, k_n-m_n))\right) \\
&= \prod_{j=1}^{n} \frac{1}{j^{m_j}}.
\end{aligned}$$

□

On applique maintenant ce lemme à $(0, \ldots, 0, m, 0, \ldots, 0)$ pour obtenir que pour tout $1 \leqslant j \leqslant n$ et tout $m \geqslant 0$, on a :

$$\mathbb{E}[c_j(\pi_n)^{[m]}] = \left(\frac{1}{j}\right)^m \mathbf{1}_{jm\leqslant n}.$$

En faisant tendre n vers l'infini, on observe alors que le moment factoriel d'ordre m de la loi limite de $c_j(\pi_n)$ est $\left(\frac{1}{j}\right)^m$, qui est exactement le moment factoriel d'ordre m d'une loi de Poisson de paramètre $1/j$.

C'est terminé, car les variables aléatoires en jeu sont à valeurs entières. Les lois et la convergence en loi se lisent sur les fonctions génératrices, dont les coefficients (quand on les développe autour de 1) donnent justement les moments factoriels. Une loi de Poisson de paramètre λ a pour fonction génératrice $\exp(\lambda(t-1))$, donc son moment factoriel d'ordre m est λ^m. □

Remarques.

— Il est difficile de trouver une source accessible pour ce raisonnement, donc il faut l'apprendre par cœur. Il est quand même fait page 11 de *Logarithmic Combinatorial Structures : a Probabilistic Approach* de Arratia, Barbour et Tavaré. Le plus dur est presque de retenir les énoncés des deux lemmes, la démonstration du premier est assez simple (et si on ne la retrouve pas, on peut s'en sortir à la main en comptant tous les choix de cycles). C'est la démonstration du second qui peut être technique à retenir.

— On peut expliquer, s'il reste du temps à la fin, que le résultat est vrai pour la loi jointe, et pas seulement pour les marginales. L'argument est le même, on peut calculer la fonction génératrice de la loi jointe et trouver $\bigotimes_j \mathcal{P}\left(\frac{1}{j}\right)$.

— Une conséquence immédiate de ce théorème est un autre développement plus facile : le nombre de points fixes d'une permutation aléatoire converge en loi vers une loi de Poisson de paramètre 1.

Recasages.

— 104 : C'est un peu limite, mais si l'on fait une grande partie sur l'étude des groupes symétriques, ça peut bien rentrer. On utilise quand même la relation orbite-stabilisateur, donc c'est dans le thème.

— 105 : De même que pour la 104, mais cette fois c'est déjà beaucoup plus dans le thème. On pourrait faire une grosse partie sur les permutations aléatoires.

— 261 : C'est parfait, et on utilise la caractérisation de la loi d'une variable discrète par ses moments factoriels (enfin, par sa fonction génératrice).

— 262 : C'est aussi parfait, on utilise la caractérisation de la convergence en loi aussi.

— 264 : Les permutations aléatoires sont un gros exemple de variables aléatoires discrètes, et qui change un peu des habituelles variables à valeurs entières. De quoi donner un élan de fraîcheur à cette leçon.

2.3.5 Probabilité que deux nombres soient premiers entre eux

Leçons 121, 190, 230

Référence Oraux Algèbre 1

Prérequis. Formule du crible de Poincaré, équivalent de la série harmonique, formule d'inversion de Möbius et/ou valeur de $\sum_{n=1}^{+\infty} \frac{\mu(n)}{n^2}$ (avec multiplicativité de la somme pour $*$)

Théorème 2.3.8. *Pour $n \geqslant 1$, soit r_n la probabilité que deux entiers choisis uniformément dans $\{1, \dots, n\}$ soient premiers entre eux. Alors la suite $(r_n)_{n\in\mathbf{N}}$ converge vers $6/\pi^2$.*

Démonstration. Soit $n \geqslant 1$. On fait la liste $p_1, \dots, p_k$ des nombres premiers inférieurs à n, et pour $i \leqslant k$ on pose :

$$U_i = \{(a,b) \in \{1,\dots,n\}^2 \mid p_i \text{ divise } a \text{ et } b\}.$$

Ainsi, si A est l'ensemble des couples (a,b) d'entiers premiers entre eux dans $\{1,\dots,n\}$, alors le complémentaire de A est la réunion des U_i. De plus, chaque intersection $U_{i_1} \cap \dots \cap U_{i_m}$ est formée des couples de multiples de $p_{i_1} \cdot \ldots \cdot p_{i_m}$, donc est de cardinal $\lfloor n/p_{i_1} \cdot \ldots \cdot p_{i_m} \rfloor^2$. D'après la formule du crible de Poincaré, on a alors :

$$\begin{aligned}
n^2 r_n = |A| &= n^2 - \left| \bigcup_{i=1}^{k} U_i \right| \\
&= n^2 - \sum_{m=1}^{k} (-1)^{m+1} \sum_{1\leqslant i_1 < \dots < i_m \leqslant k} |U_{i_1} \cap \dots \cap U_{i_m}| \\
&= n^2 - \sum_{m=1}^{k} (-1)^{m+1} \sum_{1\leqslant i_1 < \dots < i_m \leqslant k} \left\lfloor \frac{n}{p_{i_1} \cdot \ldots \cdot p_{i_m}} \right\rfloor^2 \\
&= \sum_{m=0}^{k} (-1)^{m} \sum_{1\leqslant i_1 < \dots < i_m \leqslant k} \left\lfloor \frac{n}{p_{i_1} \cdot \ldots \cdot p_{i_m}} \right\rfloor^2 \\
&= \sum_{d=1}^{n} \mu(d) \left\lfloor \frac{n}{d} \right\rfloor^2.
\end{aligned}$$

On peut alors estimer asymptotiquement r_n par :

$$\left| r_n - \sum_{d=1}^{n} \frac{\mu(d)}{d^2} \right| = \left| \sum_{d=1}^{n} \mu(d) \left(\frac{1}{n^2} \left\lfloor \frac{n}{d} \right\rfloor^2 - \frac{1}{d^2} \right) \right|.$$

De $\lfloor n/d \rfloor > n/d - 1$, on déduit en élevant au carré que :

$$\frac{1}{n^2} \left\lfloor \frac{n}{d} \right\rfloor^2 > \frac{1}{d^2} - \frac{2}{dn} + \frac{1}{n^2}.$$

L'inégalité triangulaire fournit alors :

$$\left| r_n - \sum_{d=1}^{n} \frac{\mu(d)}{d^2} \right| \leqslant \sum_{d=1}^{n} \left(\frac{2}{dn} + \frac{1}{n^2} \right) = O\left(\frac{\ln n}{n} \right),$$

d'où finalement :

$$\lim_{n \to +\infty} r_n = \sum_{d=1}^{+\infty} \frac{\mu(d)}{d^2} = \frac{6}{\pi^2}.$$

□

Remarques.

— Le développement est très joli et est une illustration très forte de la formule du crible de Poincaré. Attention par contre, il faut plusieurs résultats pas forcément triviaux pour que ça marche.

— La formule du crible de Poincaré se démontre en développant le produit suivant :

$$(1 - \mathbf{1}_{U_1})(1 - \mathbf{1}_{U_2}) \cdot \ldots \cdot (1 - \mathbf{1}_{U_k}),$$

qui est égal à l'indicatrice du complémentaire de la réunion des U_i.

— L'égalité suivante :

$$\sum_{m=0}^{k} (-1)^m \sum_{1 \leqslant i_1 < \cdots < i_m \leqslant k} \left\lfloor \frac{n}{p_{i_1} \cdot \ldots \cdot p_{i_m}} \right\rfloor^2 = \sum_{d=1}^{n} \mu(d) \left\lfloor \frac{n}{d} \right\rfloor^2$$

se justifie en regroupant les deux sommes par $p_{i_1} \cdot \ldots \cdot p_{i_m} = d$ constant, le coefficient qui apparaît est soit 0 si d a un facteur carré (parce qu'il n'apparaît pas dans la somme), soit -1 si m est impair, soit $+1$ si m est pair.

— L'estimation par $\frac{\ln n}{n}$ provient de l'estimation des sommes partielles de la série harmonique qui sont en $\ln n$, on le démontre par comparaison série-intégrale.

— La toute dernière égalité peut se justifier de deux manières dans le plan. Soit on écrit d'abord le théorème qui dit que si f et g sont deux fonctions multiplicatives dont les séries de Dirichlet convergent absolument en s alors c'est aussi le cas de $f * g$ et l'on a :

$$\sum_{n=1}^{+\infty} \frac{f(n)}{n^s} \sum_{n=1}^{+\infty} \frac{g(n)}{n^s} = \sum_{n=1}^{+\infty} \frac{(f * g)(n)}{n^s},$$

en l'appliquant alors à $f = 1$ et $g = \mu$ avec le théorème d'inversion de Möbius, soit on explique *à la main* que la somme est l'inverse de $\zeta(2)$. Cela revient au même mais le calcul peut être plus court : on a juste à expliquer que les familles

en jeu sont sommables et donc qu'on peut écrire :

$$\begin{aligned}\sum_{n=1}^{+\infty}\frac{\mu(n)}{n^2}\sum_{n=1}^{+\infty}\frac{1}{n^2} &= \sum_{d,n\geqslant 1}\frac{\mu(d)}{(dn)^2}\\ &= \sum_{m\geqslant 1}\sum_{d|m}\frac{\mu(d)}{m^2}\\ &= \sum_{m\geqslant 1}\frac{1}{m^2}\underbrace{\sum_{d|m}\mu(d)}_{=\delta_{p,1}} = 1.\end{aligned}$$

Bien sûr cela revient au même puisque c'est comme ça qu'on démontre le théorème en question, et que l'on a utilisé le théorème d'inversion de Möbius pour la toute dernière égalité.

— Justement, la formule d'inversion de Möbius est équivalente à dire que la somme $\sum_{d|m}\mu(d)$ vaut 1 si $m=1$ et 0 sinon. Pour montrer cela, on dit que c'est évident pour $m=1$, et sinon on note P l'ensemble des facteurs premiers de m pour avoir :

$$\begin{aligned}\sum_{d|m}\mu(d) &= \sum_{D\subset P}\mu\left(\prod_{p\in D}p\right) = \sum_{D\subset P}(-1)^{|D|}\\ &= \sum_{t=0}^{|P|}(-1)^t|\{D\subset P\mid |D|=t\}| = \sum_{t=0}^{|P|}(-1)^t\binom{|P|}{t}\\ &= (-1+1)^t = 0.\end{aligned}$$

Recasages.

— 121 : C'est très bien pour illustrer l'utilité de la fonction de Möbius si on l'introduit, et c'est un des rares résultats assez forts sur la répartition des nombres premiers qui est accessible en développement.

— 190 : On illustre magistralement la formule du crible de Poincaré (il y a d'autres applications, comme compter le nombre de surjections entre deux ensembles finis fixés) et par la même occasion la formule d'inversion de Möbius qui sert dans d'autres problèmes de combinatoire (comme compter le nombres de polynômes irréductibles d'un certain degré sur un certain corps fini).

— 230 : Le développement a deux liens majeurs avec cette leçon. Le premier est l'estimation de la somme harmonique, qui permet d'établir que r_n vaut bien $\sum_{n=1}^{+\infty}\frac{\mu(n)}{n^2}$. L'autre est le calcul de cette somme, qui se fait avec des théorèmes de familles sommables, donc en plein dans la leçon aussi.

2.3.6 Simplicité de $\mathrm{SO}_3(\mathbf{R})$

Leçons 103, 106, 108, 160, 204

Référence H2G2 tome 1

Prérequis. Connexité par arcs, centre, générateurs et structure de $\mathrm{SO}(3,\mathbf{R})$

Théorème 2.3.9. *Le groupe* $\mathrm{SO}(3,\mathbf{R})$ *est simple.*

Démonstration. On rappelle qu'il est connexe par arcs, de centre trivial et engendré par les retournements, qui y sont tous conjugués.

Soit H un sous-groupe distingué non trivial de $\mathrm{SO}(3,\mathbf{R})$. On montre que H contient un retournement.

Pour $1 \neq h \in H$, on pose :

$$\varphi_h : \begin{array}{c} \mathrm{SO}(3,\mathbf{R}) \longrightarrow \mathbf{R} \\ g \longmapsto \mathrm{tr}[g,h] \end{array}$$

qui est continue. Comme la trace d'un élément de $\mathrm{SO}(3,\mathbf{R})$ est de la forme $1+2\cos\theta$, l'image de ϕ est contenue dans $[-1,3]$. Par connexité et compacité de $\mathrm{SO}(3,\mathbf{R})$, l'image est connexe et compacte, et contient $\varphi_h(1)=3$. Donc l'image est de la forme $[a,3]$ avec $a \leqslant 3$.

Si $a=3$, alors $[g,h]=1$ pour tout $g \in \mathrm{SO}(3,\mathbf{R})$ donc h est central donc trivial, ce que l'on a exclu.

Donc $a<3$. Pour n assez grand on a alors :

$$a < 1+2\cos\left(\frac{\pi}{n}\right) < 3$$

et l'on peut alors trouver un $g_n \in \mathrm{SO}(3,\mathbf{R})$ tel que $\varphi_h(g_n)$ soit ce nombre au milieu. Ainsi, $h_n=[g_n,h]$ est une rotation d'angle $\pm\pi/n$ qui appartient à H, et alors h_n^n est un retournement dans H.

Ainsi H contient un retournement, donc (comme ils sont tous conjugués dans $\mathrm{SO}(3,\mathbf{R})$) il les contient tous, donc il contient le sous-groupe engendré par les retournements, qui est $\mathrm{SO}(3,\mathbf{R})$ tout entier. Donc $\mathrm{SO}(3,\mathbf{R})$ est un groupe simple. □

Remarques.

— On passe plein de petits résultats sous le tapis, qu'il faut savoir démontrer :

— $\mathrm{SO}(3,\mathbf{R})$ est connexe par arcs :
On passe d'une matrice de rotation R_θ^ξ autour d'un axe $\xi \in \mathbf{R}^3\backslash\{0\}$ à l'identité par le chemin $t \mapsto R_{t\theta}^\xi$ qui reste dans $\mathrm{SO}(3,\mathbf{R})$.

— Son centre est trivial :
Si u est dans le centre, alors il commute avec tous les retournements donc il stabilise tous les plans. Il stabilise donc aussi les intersections de plans, c'est-à-dire les droites : c'est une homothétie. En dimension impaire, il n'y a que l'identité qui reste dans le groupe spécial orthogonal.

— Il est engendré par les retournements :
Comme le groupe orthogonal est engendré par les réflexions, un élément non trivial u est toujours produit d'un nombre pair de réflexions. En dimension 3, deux réflexions τ_1 et τ_2 suffisent. Mais alors $-\tau_i$ est un retournement et $u = (-\tau_1)(-\tau_2)$. En dimension quelconque, il faut d'abord expliquer pourquoi si τ_1 et τ_2 sont des réflexions, il existe toujours deux retournements σ_1 et σ_2 tels que $\tau_1\tau_2 = \sigma_1\sigma_2$. Si l'on demande d'expliquer pourquoi le groupe orthogonal est engendré par les réflexions : soit u dans le groupe orthogonal, on raisonne par récurrence sur la codimension p du sous-espace des points fixes de u pour montrer qu'il est produit d'au plus p réflexions. On prend un vecteur x orthogonal aux points fixes, et y son image par u qui reste orthogonale aux points fixes. Ils sont de même norme, et alors $x+y$ et $x-y$ sont orthogonaux : on compose par la réflexion dirigée par $x-y$ pour faire diminuer strictement p. Ce théorème reste vrai si la forme quadratique est quelconque, mais c'est plus technique.

— Les retournements sont tous conjugués dans $\mathrm{SO}(3,\mathbf{R})$:
Si v est un retournement de plan V et w un retournement de plan W, on complète des bases orthonormées à tout l'espace et l'on peut toujours trouver une transformation orthogonale qui envoie V sur W ; quitte à changer un signe cette transformation est de déterminant 1, et elle conjugue v en w.

Recasages.

— 103 : Un résultat de simplicité qui utilise le fait que les retournements soient conjugués dans $\mathrm{SO}(3,\mathbf{R})$. Ce développement est proposé dans le rapport de jury, donc pourquoi pas.

— 106 : Très bien si l'on fait une partie sur les sous-groupes importants du groupe linéaire et/ou sur l'aspect géométrique. Les raisonnements avec les retournements, les réflexions etc sont souvent appréciés car rarement bien maîtrisés.

— 108 : Il faut alors mettre l'accent très fort sur le fait que les retournements engendrent $\mathrm{SO}(3,\mathbf{R})$, voire même le démontrer au début du développement (qui est assez court).

— 160 : Les retournements sont des endomorphismes remarquables, et le titre de cette leçon est assez vague donc ça doit pouvoir rentrer sans problème.

— 204 : On utilise de manière assez forte la connexité du groupe pour établir un résultat qui est purement algébrique. C'est en cela que ce théorème est plus intéressant que la simplicité du groupe alterné : il fait un pont avec l'analyse.

2.3.7 Théorème de stabilité de Lyapounov

Leçons 156, 206, 215, 220, 221

Référence Petit guide de calcul différentiel

Prérequis. Lemme des noyaux

Théorème 2.3.10. *Soient $f : \mathbf{R}^n \to \mathbf{R}^n$ de classe C^1, u un zéro de f et A la différentielle de f en u. On s'intéresse au système différentiel :*

$$\begin{cases} y' = f(y), \\ y(0) = y_0, \end{cases}$$

qui se linéarise en :

$$\begin{cases} z' = Az, \\ z(0) = y_0 - u. \end{cases}$$

Si les valeurs propres de A sont toutes de partie réelle strictement négative alors :

1. *0 est un point attractif du système linéarisé ;*
2. *u est un point attractif du système non-linéaire ;*
3. *Si y_0 est assez proche de u alors $y(t) \to u$ avec vitesse exponentielle.*

Démonstration. Le système linéarisé se résout en $z(t) = \mathrm{e}^{tA} z(0)$, pour démontrer le point 1 il suffit donc de montrer le lemme suivant.

Lemme 2.3.11. *Soit $a \in \mathcal{L}(E)$ un endomorphisme d'un $\mathbf{R}$-espace vectoriel normé E. Si $(\lambda_1, \ldots, \lambda_r)$ est la liste des valeurs propres distinctes de a alors il existe un polynôme $P \in \mathbf{R}[X]$ tel que :*

$$\|\mathrm{e}^{ta}\| \leqslant P(|t|) \sum_{i=1}^{r} \mathrm{e}^{t\mathfrak{Re}(\lambda_i)}.$$

Démonstration. Le lemme des noyaux donne une décomposition :

$$E = \bigoplus_{i=1}^{r} E_i$$

et donc des projections $p_i : E \to E_i$ et des inclusions $q_i : E_i \to E$, qui vérifient les propriétés suivantes :

— $p_i q_i = \mathrm{id}_{E_i}$ pour tout i et $p_j q_i = 0$ si $i \neq j$;
— $\sum_{i=1}^{r} q_i p_i = \mathrm{id}_E$.

On note alors $a_i = p_i a q_i$ pour tout i de sorte à avoir $a = \sum_i q_i a_i p_i$, et de manière générale $a^n = \sum_i q_i a_i^n p_i$ d'où :

$$\mathrm{e}^{ta} = \sum_{i=1}^{r} q_i \mathrm{e}^{ta_i} p_i = \sum_{i=1}^{r} q_i \mathrm{e}^{t\lambda_i} \mathrm{e}^{t(a_i - \lambda_i \mathrm{id}_{E_i})} p_i.$$

Comme $a_i - \lambda_i \mathrm{id}_{E_i}$ est nilpotent d'indice la multiplicité m_i, on a :

$$\|\mathrm{e}^{ta_i}\| \leqslant |\mathrm{e}^{t\lambda_i}| \sum_{k=0}^{m_i} \frac{|t|^k}{k!} \|a_i - \lambda_i \mathrm{id}_{E_i}\|^k$$

et l'on peut alors conclure :

$$\|\mathrm{e}^{ta}\| \leqslant \left(\sum_{i=1}^{r} \sum_{k=0}^{m_i} \frac{|t|^k}{k!} \|a_i - \lambda_i \mathrm{id}_{E_i}\|^k \right) \sum_{i=1}^{r} \mathrm{e}^{t\mathfrak{Re}(\lambda_i)}.$$

□

On obtient directement le point 1, que l'on veut généraliser au système non-linéaire. Pour cela, on introduit une forme bilinéaire symétrique :

$$b(x, x') = \int_0^{+\infty} \langle \mathrm{e}^{tA}x, \mathrm{e}^{tA}x' \rangle \, \mathrm{d}t$$

qui est en fait définie positive, c'est un produit scalaire. Notons q la forme quadratique associée.

Soit alors y une solution du système non-linéaire. On va montrer que y s'approche de u au sens suivant :

Proposition 2.3.12. *Il existe $\alpha, \beta > 0$ tels que si $q(y(t) - u) \leqslant \alpha$ alors :*

$$\frac{\mathrm{d}}{\mathrm{d}t} q(y(t) - u) \leqslant -\beta q(y(t) - u) \leqslant 0.$$

Démonstration. Calculons premièrement :

$$\mathrm{d}q(x)(Ax) = 2b(x, Ax) = \int_0^{+\infty} 2\langle \mathrm{e}^{tA}x, \mathrm{e}^{tA}Ax \rangle \, \mathrm{d}t = \int_0^{+\infty} \frac{\mathrm{d}}{\mathrm{d}t} \langle \mathrm{e}^{tA}x, \mathrm{e}^{tA}x \rangle \, \mathrm{d}x = -\|x\|^2.$$

Cela démontre que Ax est orienté vers l'intérieur de l'ellipsoïde $\{x' \mid q(x') = q(x)\}$, puisque son produit scalaire avec le vecteur normal $\nabla q(x)$ est négatif.

On pose $r(y) = f(y) - A(y - u)$, et on dérive :

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} q(y - u) &= \mathrm{d}q(y - u) f(y) \\ &= \mathrm{d}q(y - u)(A(y - u)) + \mathrm{d}q(y - u)(r(y)) \\ &= -\|y - u\|^2 + 2b(y - u, r(y)). \end{aligned}$$

On va maintenant utiliser l'équivalence des normes pour contrôler les deux termes qui sont apparus. Pour le premier, l'équivalence des normes donne directement un $\beta > 0$ tel que :

$$-\|y - u\|^2 \leqslant -2\beta q(y - u).$$

Pour le second, on sait que f est différentiable pour n'importe quelle norme sur $\mathbf{R}^n$, en particulier pour $\sqrt{q}$: il existe $\varepsilon : \mathbf{R}^n \to \mathbf{R}^n$ de limite nulle en 0 telle que :

$$\begin{aligned} f(y) &= f(u) + \mathrm{d}f(u)(y-u) + \sqrt{q(y-u)}\varepsilon(y-u) \\ &= A(y-u) + \sqrt{q(y-u)}\varepsilon(y-u) \end{aligned}$$

d'où $r(y) = \sqrt{q(y-u)}\varepsilon(y-u)$. Par Cauchy-Schwarz on obtient :

$$b(y-u, r(y)) \leqslant \sqrt{q(y-u)}\sqrt{q(y-u)}\sqrt{q(\varepsilon(y-u))} = q(y-u)\sqrt{q(\varepsilon(y-u))}.$$

Pour y assez proche de u on a $\sqrt{q(\varepsilon(y-u))} \leqslant \beta/2$ ce qui donne le résultat. □

On peut maintenant conclure avec le lemme de Gronwall :

$$q(y-u) \leqslant \alpha \mathrm{e}^{-\beta t}.$$

□

Remarques.

- On peut passer plus rapidement sur le lemme car ce n'est pas le point 1 le plus important. D'ailleurs les systèmes linéaires sont leur propre linéarisé donc il n'y aurait même pas besoin d'expliquer ce point.
- Faire un dessin pour expliquer le sens de $\mathrm{d}q(x)(Ax) = -\|x\|^2$ peut être une très bonne idée. De même, c'est bien de motiver l'existence de la proposition.
- La forme quadratique q est une *fonction de Lyapounov*, car elle s'annule en 0, est strictement positive dans un voisinage de 0 (sauf en 0) et sa dérivée le long du champ de vecteurs A est négative dans un voisinage de l'origine. Les fonctions de Lyapounov servent plus généralement à étudier la stabilité des systèmes différentiels. L'idée est celle donnée dans le développement : comme la dérivée le long du champ de vecteurs est négative, la solution ne peut pas s'échapper d'un ellipsoïde qui devient de plus en plus petit si la dérivée est strictement négative. On obtient alors un point d'équilibre asymptotiquement stable.
- Comme on choisit f de classe C^1, on passe sous le tapis l'utilisation du théorème de Cauchy-Lipschitz. Il n'est pas crucial car ce serait valable même si aucune solution n'existait.

Recasages.

- 206 : Une très bonne alternative au théorème de Cauchy-Lipschitz. On utilise à plusieurs reprises l'équivalence des normes et donc le fait que l'on est en dimension finie. On pourrait chercher un contre-exemple en dimension infinie mais je n'en connais pas.
- 215 : On illustre le principe intuitif derrière le gradient, et l'on différentie plusieurs fois donc c'est dans le thème. Attention quand même à bien placer soigneusement l'énoncé dans le plan pour ne pas que ça fasse tache.

— 220 : Il n'y a rien à dire, les théorèmes de stabilité peuvent même faire une partie toute entière.
— 221 : Pareil que pour la 220, on illustre parfaitement le principe de linéarisation, qui apparaît partout en mathématiques. On ne comprend pas bien les systèmes non-linéaires, alors on se ramène au cas linéaire. Et le théorème de Lyapounov en est un exemple frappant et simple à comprendre.

2.3.8 Théorème des extrema liés

Leçons 159, 206, 214, 215, 219

Référence Objectif agrégation

Prérequis. Théorème des fonctions implicites, théorème des fonctions composées

Proposition 2.3.13. *Soient U un ouvert de $\mathbf{R}^n$, $F : U \to \mathbf{R}^{n-r}$ une application de classe C^1 définissant $M = \{x \in U \mid F(x) = 0\}$. Si la différentielle de F en tout point de M est surjective alors M est une sous-variété de $\mathbf{R}^n$ de classe C^1, ce que l'on suppose. Alors l'ensemble des vecteurs tangents à un point $a \in M$:*

$$T_aM = \{\gamma'(0) \colon \gamma : [-1,1] \to M \text{ de classe } C^1 \text{ tel que } \gamma(0) = a\}$$

est égal au sous-espace $\operatorname{Ker} \mathrm{d}F(a)$.

Démonstration. On pose $E_1 = \operatorname{Ker} \mathrm{d}F(a)$ et E_2 un supplémentaire de E_1 dans $\mathbf{R}^n$. On applique alors le théorème des fonctions implicites à F définie sur $E_1 \oplus E_2 = E_1 \times E_2$, pour obtenir que M est localement le graphe d'une fonction u de classe C^1 :

$$x_1 \in V_1, x_2 \in V_2, F(x_1 + x_2) = 0 \iff x_1 \in V_1, x_2 = u(x_1).$$

On a alors $T_aM = \{h + \mathrm{d}u(a_1)(h) : h \in E_1\}$, mais le théorème des fonctions implicites donne :

$$\mathrm{d}u(a_1) = -\left(\mathrm{d}F(a) \circ \pi_2\right)^{-1} \circ \left(\mathrm{d}F(a) \circ \pi_1\right)$$

avec π_i la projection $\mathbf{R}^n \to E_i$ parallèlement à l'autre E_j. Sauf que $E_1 = \operatorname{Ker} \mathrm{d}F(a)$ donc $\mathrm{d}u(a_1)(h) = 0$ pour tout $h \in E_1$, ce qui montre que $\operatorname{Ker} \mathrm{d}F(a) = E_1 = T_aM$. □

Théorème 2.3.14. *Soient U un ouvert de $\mathbf{R}^n$, $f, g_1, \ldots, g_r : U \to \mathbf{R}$ des fonctions de classe C^1 et :*

$$\Gamma = \{x \in U \mid g_1(x) = \cdots = g_r(x) = 0\}.$$

Si $f|_\Gamma$ admet un extremum local en $a \in \Gamma$ et si les $\mathrm{d}g_i(a)$ sont linéairement indépendants, alors il existe des réels $\lambda_1, \ldots, \lambda_r$ tels que :

$$\mathrm{d}f(a) = \sum_{i=1}^{r} \mathrm{d}g_i(a).$$

Démonstration. Comme les g_i sont de classe C^1 et les $\mathrm{d}g_i$ libres, la différentielle de $(g_1, \ldots, g_r)$ est surjective, et donc Γ est une sous-variété C^1 de $\mathbf{R}^n$. Soient $v \in T_a\Gamma$, et ainsi $\gamma : [-1,1] \to \Gamma$ valant a en 0 avec $\gamma'(0) = v$. Il suffit de dériver $f \circ \gamma$ en 0 grâce au théorème des fonctions composées :

$$(f \circ \gamma)'(0) = \mathrm{d}f(\gamma(0))(\gamma'(0)) = \mathrm{d}f(a)(v).$$

Comme $f|_\Gamma$ admet un extremum local en a, cette dérivée doit s'annuler donc $\mathrm{d}f(a)(v) = 0$. C'est valable pour tout v, donc $\mathrm{d}f(a)$ est nulle sur $T_a\Gamma$. Il suffit alors d'appliquer le lemme suivant à $T_a\Gamma = \bigcap_{i=1}^{r} \operatorname{Ker} \mathrm{d}g_i(a)$.

Lemme 2.3.15. *Soient $\phi, \psi_1, \dots, \psi_r$ des formes linéaires sur un espace vectoriel de dimension finie. Alors ϕ est combinaison linéaires des ψ_i si et seulement si :*

$$\operatorname{Ker}\phi \supset \bigcap_{i=1}^{r} \operatorname{Ker}\psi_i.$$

Démonstration. Le sens direct est évident. Pour le sens réciproque, il suffit de passer à l'orthogonal (dans le dual) :

$$(\operatorname{Ker}\phi)^\perp \subset \sum_{i=1}^{n} (\operatorname{Ker}\psi_i)^\perp$$

et remarquer que $(\operatorname{Ker}\phi)^\perp$ est la droite engendrée par ϕ. □

□

Corollaire 2.3.16. *Pour tous $v_1, \dots, v_n \in \mathbf{R}^n$, on a :*

$$|\det(v_1, \dots, v_n)| \leqslant \prod_{i=1}^{n} \|v_i\|_2.$$

Démonstration. On maximise le déterminant sur le tore :

$$\mathbf{S} = \{(v_1, \dots, v_n) \in (\mathbf{R}^n)^2 \mid \|v_1\| = \cdots = \|v_n\| = 1\}.$$

Le théorème des extrema liés montre que si det atteint son maximum en un $(v_1, \dots, v_n)$ alors c'est une base orthonormée donc de déterminant ± 1. Le résultat est obtenu par homogénéité. □

Remarques.

- Il faut faire attention car cette version est d'un assez haut niveau. On utilise le théorème des fonctions implicites pour montrer qu'une sous-variété définie localement avec une équation peut être définie localement avec un graphe. Ce théorème nous donne aussi le fait que l'espace tangent en un point est le noyau de la différentielle de l'équation en ce point, ce que l'on utilise dans la démonstration du théorème des extrema liés.
- Justement à ce propos, le jury déteste quand le contenu d'un développement est caché dans un lemme : ça repousse les mathématiques à ailleurs et le développement devient vide de raisonnement. Il faut donc absolument expliquer la proposition, même si les détails peuvent être un peu gommés. Se contenter du théorème et du lemme ne suffit pas, puisque tout réside dans l'égalité $T_a\Gamma = \bigcap \operatorname{Ker} \mathrm{d}g_i(a)$.
- C'est quand même probablement beaucoup mieux à l'oral de commencer par démontrer le théorème des extrema liés, puis de démontrer la proposition à la fin. Ainsi plus on prend du temps au début, plus on pourra gommer les détails de la proposition pour s'assurer que le développement soit clair, sans perdre en rigueur.

— Après ce développement, mieux vaut maîtriser les exercices simples qui peuvent suivre. Notamment l'utilisation du théorème des fonctions implicites, l'utilisation des extrema liés dans un cas simple ou énoncer d'autres applications. Une jolie application est la majoration $V^{2/3} \leqslant S/6$ où V et S sont respectivement le volume et l'aire d'un parallélépipède rectangle dans $\mathbf{R}^3$, avec égalité quand c'est un cube.

Recasages.

— 159 : Le lemme est un résultat facile de dualité en dimension finie, qui pourtant donne un résultat d'analyse très fort. Attention, le développement n'est pas très original dans cette leçon j'ai l'impression.

— 206 : On n'*utilise* pas la dimension finie à proprement parler, mais comme tout se déroule en dimension finie ça peut passer. Mais je trouve ça un peu limite.

— 214 : C'est une très bonne application du théorème des fonctions implicites, qui complète la suite logique inversion locale $\to$ fonctions implicites $\to$ extrema liés. C'est une bonne idée de démontrer un gros théorème du plan et de ne pas avoir que des développements anecdotiques.

— 215 : On utilise des différentielles partout et le résultat lui-même porte sur des applications de classe C^1, donc c'est en plein dans le thème. En plus dans les applications (comme l'inégalité de Hadamard), on calcule réellement des différentielles (ici celle du déterminant).

— 219 : Le nom du développement coïncide presque avec le titre de la leçon.