Développements agrégation externe de mathématiques

Année 2018-2019

Thomas Cavallazzi

Table des matières

1 Introduction

Ce document contient les développements que j'ai travaillés durant mon année de préparation à l'agrégation à Rennes. Je tiens à remercier l'ensemble des personnes qui m'ont permis d'écrire ce document. Merci aux anciens agrégatifs qui ont partagé leur travail sur leur page web dont je me suis largement inspiré. Les pages que j'ai le plus utilisées sont celles de Corentin Kilque, Adrien Laurent, Antoine Diez, Florian Lemonnier, Paul Alphonse et Adrien Fontaine donc un grand merci à eux. Merci également à mes amis de Rennes, l'année de préparation à l'agrégation a été l'occasion de beaucoup de partage et de collaboration et certains ont contribué à améliorer ces développements. Enfin merci aux professeurs de Rennes dont les cours m'ont également inspiré. Ce document est donc le fruit d'un travail de synthèse.

Le document comporte ma liste de couplages leçons/développements. Pour chaque leçon j'indique les développements que j'utilisais pour les oraux ainsi que les autres développements possibles. Cela aidera sûrement les agrégatifs qui tentent d'optimiser leur liste de développements.... Je n'ai rédigé que certains développements, probablement mes préférés et les plus originaux ou ceux auxquels j'ai ajouté une touche personnelle. Pour chaque développement j'ai mis les sources que j'ai utilisées ainsi qu'une version rédigée disponible sur internet.

Ce document comporte certainement des erreurs, il est donc à utiliser avec prudence. N'hésitez surtout pas à me contacter si vous avez une question ou remarque pour que je puisse corriger ce qui doit l'être. Ce sera bénéfique pour tout le monde !

Pour les agrégatifs, il y a sans doute quelques développements difficiles qu'il faut bien maîtriser si on veut les présenter le jour J. Il y a quelques développements qui pourront intéresser les plus probabilistes d'entre vous. Il faut donc bien s'approprier chaque développement. Le mieux est de les faire au tableau pour être dans les conditions. N'hésitez pas à faire ça en groupe, le regard critique (mais bienveillant !) des camarades est d'une redoutable efficacité. Et cela permet de travailler un peu sur sa pédagogie, malgré le format de l'épreuve qui s'y prête peu. Enfin je vous souhaite bon courage !

2 Couplages

2.1 Leçons d'Algèbre et de Géométrie

- **101 : Groupe opérant sur un ensemble. Exemples et applications.**
  - Réciprocité quadratique
  - Théorème de la base de Burnside
  - *Nombre de matrices diagonalisables sur les corps finis*
  - *Table des caractères de S_4*
  - *Sous-groupes compacts de $GL_n(\mathbb{R})$*

- **102 : Groupe des nombres complexes de module 1. Sous-groupes des racines de l'unité. Applications.**
  - Irréductibilité des polynômes cyclotomiques sur $\mathbb{Z}$
  - Polygones réguliers constructibles
  - *Théorème de structure de groupes abéliens finis*

- **103 : Exemples de sous-groupes distingués et de groupes quotients. Applications.**
  - Théorème de la base de Burnside
  - Sous-groupes distingués et noyaux des caractères
  - *Frobenius Zolotarev*
  - *Simplicité de $S0_3(\mathbb{R})$*

- **104 : Groupes finis. Exemples et applications.**
  - Théorème de la base de Burnside
  - Théorème de structure des groupes abéliens finis
  - *Réciprocité quadratique*
  - *Nombre de matrices diagonalisables sur les corps finis*
  - *Table des caractères de S_4*

- **105 : Groupe des permutations d'un ensemble fini. Applications**
  - Frobenius Zolotarev
  - Table des caractères de S_4

- **106 : Groupe linéaire d'un espace vectoriel de dimension finie E, sous-groupes de $GL(E)$. Applications.**
  - Frobenius Zolotarev
  - Sous-groupes compacts de $GL_n(\mathbb{R})$

- *Simplicité de $SO_3(\mathbb{R})$*

- **107 : Représentations et caractères d'un groupe fini sur un $\mathbb{C}$-espace vectoriel. Exemples.**
  - Théorème de structure des groupes abéliens finis
  - Sous-groupes distingués et noyaux de caractères
  - *Tables des caractères de S_4*

- **108 : Exemples de parties génératrices d'un groupe. Applications.**
  - Théorème de la base de Burnside
  - Simplicité de $SO_3(\mathbb{R})$
  - *Quaternions et rotations*

- **110 : Structure et dualité des groupes abéliens finis. Applications.**
  - Théorème de structure des groupes abéliens finis
  - Formule de Poisson discrète

- **120 : Anneaux $\mathbb{Z}/n\mathbb{Z}$. Applications.**
  - Théorème de structure des groupes abéliens finis
  - Théorème de Sophie Germain
  - *Réciprocité quadratique*

- **121 : Nombres premiers. Applications.**
  - Réciprocité quadratique
  - Polygones réguliers constructibles
  - *Théorème des deux carrés*
  - *Théorème de Sophie Germain*

- **122 : Anneaux principaux. Applications.**
  - Théorème des deux carrés
  - Autour des endomorphismes semi-simples
  - *Réduction de Frobenius*

- **123 : Corps finis. Applications.**
  - Réciprocité quadratique
  - Nombre de polynômes irréductibles sur les corps finis
  - *Frobenius Zolotarev*

- *Algorithme de Berlekamp*

- **125 : Extensions de corps. Exemples et applications.**
  - Polygones réguliers constructibles
  - Nombre de polynômes irréductibles sur les corps finis

- **126 : Exemples d'équations en arithmétique.**
  - Théorème des deux carrés
  - Réciprocité quadratique
  - *Théorème de Sophie Germain*

- **141 : Polynômes irréductibles à une indéterminée. Corps de rupture. Exemples et applications.**
  - Irréductibilité des polynômes cyclotomiques sur $\mathbb{Z}$
  - Nombres de polynômes irréductibles sur les corps finis
  - *Autour des endomorphismes semi-simples*
  - *Algorithme de Berlekamp*

- **142 : PGCD et PPCM, algorithmes de calcul. Applications**
  - Théorème de Sophie Germain
  - Algorithme de Berlekamp

- **144 : Racines d'un polynôme. Fonctions symétriques élémentaires. Exemples et applications.**
  - Structure de polybômes symétriques
  - Irréductibilité des polynômes cyclotomiques sur $\mathbb{Z}$

- **150 : Exemples d'actions de groupes sur les espaces de matrices.**
  - Sous-groupes compacts de $GL_n(\mathbb{R})$
  - Nombre de matrices diagonalisables sur les corps finis
  - *Réduction de Frobenius*
  - *Réciprocité quadratique*

- **151 : Dimension d'un espace vectoriel (on se limitera au cas de la dimension finie). Rang. Exemples et applications.**
  - Polygones réguliers constructibles
  - Théorème de la base de Burnside

- *Réduction de Frobenius*
- *Coniques passant par 5 points*
- *Algorithme de Berlekamp*

- **152 : Déterminant. Exemples et applications.**
  - Frobenius Zolotarev
  - Coniques passant par 5 points

- **153 : Polynômes d'endomorphisme en dimension finie. Réduction d'un endomorphisme en dimension finie. Applications.**
  - Image de l'exponentielle
  - Décomposition de Dunford effective
  - *Réduction de Frobenius*
  - *Autour des endomorphismes semi-simples*

- **154 : Sous-espaces stables par un endomorphisme ou une famille d'endomorphismes d'un espace vectoriel de dimension finie. Applications.**
  - Réduction de Frobenius
  - Autour des endomorphismes semi-simples

- **155 : Endomorphismes diagonalisables en dimension finie.**
  - Décomposition de Dunford effective
  - Nombre de matrices diagonalisables sur les corps finis
  - *Autour des endomorphismes semi-simples*

- **156 : Exponentielle de matrices. Applications.**
  - Image de l'exponentielle
  - Homéomorphisme de exp : $\mathcal{S}_n(\mathbb{R}) \to \mathcal{S}_n^{++}(\mathbb{R})$

- **157 : Endomorphismes trigonalisables. Endomorphismes nilpotents.**
  - Décomposition de Dunford effective
  - Convergence des méthodes itératives

- **158 : Matrices symétriques réelles, matrices hermitiennes.**
  - Homéomorphisme de exp : $\mathcal{S}_n(\mathbb{R}) \to \mathcal{S}_n^{++}(\mathbb{R})$
  - Lemme de Morse

- **159 : Formes linéaires et dualité en dimension finie. Exemples et applications.**
  - Réduction de Frobenius
  - Théorème des extrema liés

- **160 : Endomorphismes remarquables d'un espace vectoriel euclidien (de dimension finie).**
  - Simplicité de $SO_3(\mathbb{R})$
  - Homéomorphisme de exp : $\mathcal{S}_n(\mathbb{R}) \to \mathcal{S}_n^{++}(\mathbb{R})$
  - *Sous-groupes compacts de $GL_n(\mathbb{R})$*
  - *Table des caractères de S_4*

- **161 : Distances et isométries d'un espace affine euclidien.**
  - Simplicité de $SO_3(\mathbb{R})$
  - Table des caractères de S_4
  - *Quaternions et rotations*

- **162 : Systèmes d'équations linéaires ; opérations élémentaires, aspects algorithmiques et conséquences théoriques.**
  - Méthode du gradient à pas optimal
  - Coniques passant par 5 points
  - *Méthodes itératives pour les systèmes linéaires*

- **170 : Formes quadratiques sur un espace vectoriel de dimension finie. Orthogonalité, isotropie. Applications.**
  - Réciprocité quadratique
  - Lemme de Morse

- **171 : Formes quadratiques réelles. Coniques. Exemples et applications.**
  - Lemme de Morse
  - Coniques passant par 5 points

- **181 : Barycentres dans un espace affine réel de dimension finie, convexité. Applications.**
  - Sous-groupes compacts de $GL_n(\mathbb{R})$
  - Coniques passant par 5 points

- **182 : Applications des nombres complexes à la géométrie.**
  - Polygones réguliers constructibles

- Quaternions et rotation

- **183 : Utilisation des groupes en géométrie.**
  - Quaternions et rotations
  - Polygones réguliers constructibles

- **190 : Méthodes combinatoires, problèmes de dénombrement.**
  - Nombre de polynômes irréductibles sur les corps finis
  - Réciprocité quadratique
  - *Nombre de matrices diagonalisables sur les corps finis*

2.2 Leçons d'Analyse

- **201 : Espaces de fonctions. Exemples et applications.**
  - Représentation des fonctions lipschitziennes
  - Densité des fonctions continues nulle part dérivables
  - *Equation de Schrödinger sur* $\mathbb{R}$

- **202 : Exemples de parties denses et applications.**
  - Équirépartition presque sûre des orbites du doublement de l'angle
  - Théorème de Féjèr
  - *Densité des fonctions continues nulle part dérivables*

- **203 : Utilisation de la notion de compacité.**
  - Théorème de Hadamard-Lévy
  - Optimisation dans un Hilbert
  - *Théorèmes de Prokhorov et de Lévy*
  - *Homéomorphisme de exp* $: \mathcal{S}_n(\mathbb{R}) \to \mathcal{S}_n^{++}(\mathbb{R})$

- **204 : Connexité. Exemples et applications.**
  - Image de l'exponentielle
  - Théorème de Hadamard-Lévy
  - *Simplicité de* $SO_3(\mathbb{R})$

- **205 : Espaces complets. Exemples et applications.**
  - Optimisation dans un Hilbert
  - Densité des fonctions continues nulle part dérivables
  - *Lax-Milgram et application*

- **207 : Prolongement de fonctions. Exemples et applications.**
  - Représentation des fonctions lipschitziennes
  - Prolongement de la fonction Zêta
  - *Optimisation dans un Hilbert*

- **208 : Espaces vectoriels normés, applications linéaires continues. Exemples.**
  - Représentation des fonctions lipschitziennes
  - Optimisation dans un Hilbert
  - *Lax-Milgram et application*

- **209 : Approximation d'une fonction par des polynômes et des polynômes trigonométriques. Exemples et applications**
  - Théorème de Féjèr
  - Théorème taubérien fort

- **213 : Espaces de Hilbert. Bases hilbertiennes. Exemples et applications.**
  - Optimisation dans un Hilbert
  - Lax-Milgram et application

- **214 : Théorème d'inversion locale, théorème des fonctions implicites. Exemples et applications en analyse et en géométrie.**
  - Théorème de Hadamard-Lévy
  - Image de l'exponentielle
  - *Lemme de Morse*
  - *Théorème des extrema liés*

- **215 : Applications différentiables définies sur un ouvert de $\mathbb{R}^n$. Exemples et applications.**
  - Théorème de Hadamard-Lévy
  - Théorème des extrema liés
  - *Lemme de Morse*
  - *Méthode du gradient à pas optimal*

- **219 : Extremums : existence, caractérisation, recherche. Exemples et applications.**
  - Optimisation dans un Hilbert
  - Théorème des extrema liés
  - *Méthode du gradient à pas optimal*

- **220 : Équations différentielles $X' = f(t, X)$. Exemples d'étude des solutions en dimension 1 et 2.**
  - Théorème de Hadamard-Lévy
  - Théorème de stabilité de Liapunov
  - *Nombre de zéros d'une EDO*

- **221 : Équations différentielles linéaires. Systèmes d'équations différentielles linéaires.**
  - Nombre de zéros d'une EDO
  - Théorème de stabilité de Liapunov

- **222 : Exemples d'équations aux dérivées partielles linéaires.**
  - Equation de la chaleur sur le cercle
  - Lax-Milgram et application
  - *Equation de Schrödinger*

- **223 : Suites numériques. Convergence, valeurs d'adhérence. Exemples et applications.**
  - Équirépartition presque sûre des orbites du doublement de l'angle
  - Convergence presque sûre des sous-martingales bornées dans L^1 (oups j'aime les probas ...)

- **224 : Exemples de développements asymptotiques de suites et de fonctions.**
  - Nombre de zéros d'une EDO
  - Quelques ordres moyens

- **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.**
  - Équirépartition presque sûre des orbites du doublement de l'angle
  - Méthode du gradient à pas optimal
  - *Méthodes itératives pour les systèmes linéaires*

- **228 : Continuité et dérivabilité des fonctions réelles d'une variable réelle. Exemples.**
  - Représentation des fonctions lipschitziennes
  - Densité des fonctions continues nulle part dérivables

- **229 : Fonctions monotones. Fonctions convexes. Exemples et applications.**
  - Optimisation dans un Hilbert
  - Théorèmes de Prokhorov et de Lévy
  - *Méthode du gradient à pas optimal*

- **230 : Séries de nombres réels ou complexes. Comportement des restes ou des sommes partielles des séries numériques. Exemples.**
  - Théorème taubérien fort
  - Quelques ordres de moyens

- **233 : Analyse numérique matricielle : résolution approchée de systèmes linéaires, Recherche de vecteurs propres, exemples.**
  - Méthode du gradient à pas optimal

- Méthodes itératives pour les systèmes linéaires

- **234 : Fonctions et espaces de fonctions Lebesgue-intégrables.**
  - Théorème de Féjèr
  - Représentation des fonctions lipschitziennes
  - *Inversion de Fourier*

- **235 : Problèmes d'interversion de limites et d'intégrales.**
  - Inversion de Fourier
  - Théorème taubérien fort
  - *Formule des compléments*
  - *Équation de la chaleur sur le cercle*

- **236 : Illustrer par des exemples quelques méthodes de calcul d'intégrales de fonctions d'une ou plusieurs variables.**
  - Formule des compléments
  - Inversion de Fourier

- **239 : Fonctions définies par une intégrale dépendant d'un paramètre. Exemples et applications.**
  - Inversion de Fourier
  - Prolongement de la fonction Zêta
  - *Équation de Schrödinger sur* $\mathbb{R}$

- **241 : Suites et séries de fonctions. Exemples et contre-exemples.**
  - Equation de la chaleur sur le cercle
  - Théorème taubérien fort
  - *Théorème de Prokhorov et de Lévy*

- **243 : Convergence des séries entières, propriétés de la somme. Exemples et applications.**
  - Problème des moments
  - Théorème taubérien fort

- **245 : Fonctions holomorphes sur un ouvert de $\mathbb{C}$. Exemples et applications.**
  - Formule des compléments
  - Prolongement de la fonction Zêta

- **246 : Séries de Fourier. Exemples et applications.**
  - Équation de la chaleur sur le cercle
  - Théorème de Féjèr

- **250 : Transformation de Fourier. Applications.**
  - Équation de Schrödinger sur $\mathbb{R}$
  - Inversion de Fourier

- **253 : Utilisation de la notion de convexité en analyse.**
  - Optimisation dans un Hilbert
  - Méthode du gradient à pas optimal
  - *Sous-groupes compacts de* $GL_n(\mathbb{R})$

- **260 : Espérance, variance et moments d'une variable aléatoire.**
  - Marche aléatoire sur $\mathbb{Z}^d$
  - Problème des moments

- **261 : Loi d'une variable aléatoire : caractérisations, exemples, applications.**
  - Problème des moments
  - Équirépartition presque sûre des orbites du doublement de l'angle

- **262 : Convergences d'une suite de variables aléatoires. Théorèmes limites. Exemples et applications.**
  - Convergence presque sûre des sous-martingales bornées L^1
  - Théorème de Prokhorov et de Lévy
  - *Marche aléatoire sur* $\mathbb{Z}^d$
  - *Équirépartition presque sûre des orbites du doublement de l'angle*

- **264 : Variables aléatoires discrètes. Exemples et applications.**
  - Marche aléatoire sur $\mathbb{Z}^d$
  - Équirépartition presque sûre des orbites du doublement de l'angle

- **265 : Exemples d'études et d'applications de fonctions usuelles et spéciales.**
  - Formule des compléments
  - Prolongement de la fonction Zêta

3 Développements d'Analyse

3.1 Théorème de Hadamard-Lévy

Références : C. Zuily et H. Queffelec, *Elements d'analyse*, 2ème édition – M. Zavidovique, *Un max de maths.*

Leçons : 203, 204, 214, 215, 220.

Définition 3.1.1

Une application $f : \mathbb{R}^n \to \mathbb{R}^n$ est dite propre si et seulement si pour tout compact $K \subset \mathbb{R}^n$, $f^{-1}(K)$ est compact. Ceci équivaut à $\|f(x)\| \underset{\|x\|\to+\infty}{\to} +\infty$.

L'objectif du développement est de démontrer le théorème suivant.

Théorème 3.1.2 (*Théorème d'inversion globale de Hadamard-Lévy*)

Soit $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$. Il y a équivalence entre :

1. *f est un $\mathcal{C}^1$-difféomorphisme global de $\mathbb{R}^n$.*
2. *f est propre et pour tout $x \in \mathbb{R}^n$, $df_x \in GL(\mathbb{R}^n)$.*

Démonstration : ($\Rightarrow$) Supposons que f soit un $\mathcal{C}^1$-difféomorphisme global de $\mathbb{R}^n$. En différentiant la relation $f^{-1} \circ f = \mathrm{Id}$, on obtient par le théorème des fonctions composées que pour tout $x \in \mathbb{R}^n$, $d(f^{-1})_{f(x)} \circ df_x = \mathrm{Id}$, ce qui assure l'inversibilité de la différentielle en tout point. Pour le caractère propre, il suffit de remarquer que si K est un compact de $\mathbb{R}^n$, alors $f^{-1}(K)$ est un compact comme image du compact K par l'application continue f^{-1}.

($\Leftarrow$) **Étape 1 :** Supposons maintenant que f est une application propre et dont la différentielle est inversible en tout point. Cette dernière hypothèse permet d'appliquer le théorème d'inversion locale à f ce qui assure que f est localement un $\mathcal{C}^1$-difféomorphisme. Il s'agit donc de montrer que f est bijective, on pourra ainsi utiliser le théorème d'inversion globale. On se donne $y \in \mathbb{R}^n$ et on cherche à montrer que y admet un unique antécédent par f, c'est-à-dire que $\mathrm{Card}\ (f^{-1}(\{y\})) = 1$. Quitte à poser $g = f - y$, qui vérifie les mêmes hypothèses que f, on se ramène à étudier le nombre d'antécédents de 0 par f.

Étape 2 : L'idée est alors simple : il s'agit de trouver un flot le long duquel f décroît vers 0. Pour cela, on pose :

$$F : \left\{ \begin{array}{ll} \mathbb{R}^n & \to \mathbb{R}^n \\ x & \mapsto -(df_x)^{-1}.f(x) \end{array} \right. .$$

Il s'agit d'une application de classe $\mathcal{C}^1$ puisqu'on a supposé f de classe $\mathcal{C}^2$. Ainsi on considère le problème de Cauchy suivant :

$$\left\{ \begin{array}{ll} y' & = F(y) \\ y(0) & = q \in \mathbb{R}^n \end{array} \right. .$$

Le théorème de Cauchy-Lipschitz assure que ce problème admet une unique solution maximale qu'on note $\varphi(.,q)$ définie sur un intervalle maximal $[0, T^*[$. Voyons maintenant en quoi le choix de cette fonction F permet d'atteindre l'objectif qu'on s'est fixé. On considère l'application

$$t \in [0, T^*[\mapsto g(t) := f \circ \varphi(t,q).$$

Cette application est $\mathcal{C}^1$ et vérifie :

$$\begin{aligned}\forall t \in [0, T^*[, \quad g'(t) &= df_{\varphi(t,q)}.\partial_t \varphi(t,q) \\ &= df_{\varphi(t,q)}.(-(df_{\varphi(t,q)})^{-1}.f(\varphi(t,q))) \\ &= -g(t).\end{aligned}$$

Ainsi par unicité de la solution, on en déduit que :

$$\forall t \in [0, T^*[,\ f(\varphi(t,q)) = e^{-t} f(q).$$

En particulier, on en déduit que le flot $\varphi(.,q)$ est à valeurs dans $f^{-1}(\overline{B}(0, \|f(q)\|))$ qui est compact puisque f est propre. Le théorème de sortie de tout compact assure que la solution φ est globale, c'est-à-dire que $T^* = +\infty$. De plus comme le flot est à valeurs dans un compact, on en déduit qu'il existe $y \in \mathbb{R}^n$ et une sous-suite $(t_k)_k$ strictement croissante vers $+\infty$ telle que :

$$\varphi(t_k, q) \underset{k \to +\infty}{\to} y.$$

On déduit alors, par continuité de f, que $f(y) = 0$ puisque $g(t_k) \to 0$. On a donc l'existence d'un antécédent de 0 pour f (ce qui montre la surjectivité de f). Il s'agit maintenant de montrer que cet antécédent est unique.

Étape 3 : On remarque que les équilibres du système différentiel introduit sont exactement les zéros de f. Nous allons voir qu'ils sont asymptotiquement stables. Pour cela appliquons le théorème d'inversion locale en y. On dispose de U^y un voisinage de y dans $\mathbb{R}^n$ et de $B^y := B(0, \delta_y)$ tel que f soit un $\mathcal{C}^1$-difféomorphisme de U^y sur B^y. Supposons qu'il existe t_0 tel que $\varphi(t_0, q) \in U^y$. Montrons que l'on a :

$$\{t \in [t_0, +\infty[,\ \varphi(t,q) \in U^y\} = \{t \in [t_0, +\infty[,\ \varphi(t,q) = (f_{|U^y})^{-1}(e^{-t} f(q))\}.$$

L'inclusion directe vient du fait que $f(\varphi(t_0,q)) = e^{-t_0} f(q) \in B^y$ et comme la norme décroît quand t augmente, $f(\varphi(t,q)) = e^{-t} f(q) \in B^y$, pour tout $t \geq t_0$. En appliquant $(f_{|U^y})^{-1}$, on obtient l'inclusion directe. L'inclusion réciproque est claire.

Le premier ensemble est ouvert comme pré-image d'un ouvert par l'application continue $\varphi(.,q)$. Le second est fermé comme pré-image de 0 par l'application continue $\varphi(t,q) - f^{-1}_{|U^y}(e^{-t} f(q))$. Comme les ensembles sont égaux, il s'agit d'un ouvert fermé non vide, donc égal à $[t_0, +\infty[$, par connexité de ce dernier ensemble. En laissant tendre t vers $+\infty$ dans l'égalité : $\varphi(t,q) = f^{-1}_{|U^y}(e^{-t} f(q))$, on obtient que :

$$\varphi(t,q) \underset{t \to +\infty}{\to} y,$$

puisque y est le seul 0 de f dans U^y. Or dans l'étape 2, on a fait converger le flot vers y à extraction près, on en déduit qu'en fait il converge globalement vers y lorsque t tend vers $+\infty$.

Étape 4 : On utilise un argument de connexité pour conclure. Posons, pour $y \in f^{-1}(\{0\})$:

$$W^y := \left\{ q \in \mathbb{R}^n,\ \varphi(t,q) \underset{t \to +\infty}{\to} y \right\}.$$

Dans les étapes précédentes, on a vu que :

$$(*) \quad \mathbb{R}^n = \bigcup_{y \in f^{-1}(\{0\})} W^y.$$

En effet, la trajectoire issue de n'importe quelle donnée initiale converge vers un 0 de f.

Montrons que les W^y sont des ouverts non vides. Le caractère non vide provient du fait que $y \in W^y$, puisque y est un équilibre du système différentiel. Pour le caractère ouvert, on reprend le voisinage U^y de y obtenu à l'étape 2 par le théorème d'inversion locale. On dispose de $\eta_y > 0$, tel que $B(y, 2\eta_y) \subset U^y$. Soit $q \in W^y$, on dispose de $T > 0$ tel que $\varphi(T, q) \in B(y, \eta_y)$. De plus, la continuité du flot par rapport à la donnée initiale assure qu'il existe $\delta > 0$ tel que :

$$\|q - q'\| < \delta \Rightarrow \|\varphi(T, q) - \varphi(T, q')\| < \eta_y.$$

L'inégalité triangulaire assure alors que pour $\|q - q'\| < \delta$, on a : $\varphi(T, q') \in B(y, 2\eta_y) \subset U^y$. L'étape 3 assure alors que :

$$\varphi(t, q') \underset{t \to +\infty}{\to} y.$$

Ainsi $B(q, \delta)$ est inclus dans W^y, qui est donc ouvert. L'égalité $(*)$ ainsi que la connexité de $\mathbb{R}^n$ assurent que :

$$\text{Card } f^{-1}(\{0\}) = 1,$$

ce qui achève la preuve.

$\square$

Remarque 3.1.3.

1. *Le théorème reste vrai dans le cas où f est seulement supposée de classe $\mathcal{C}^1$ mais la preuve est plus compliquée (on ne peut pas directement utiliser le théorème de Cauchy-Lipschitz avec le lemme de sortie de tout compact).*
2. *On peut se demander d'où vient l'EDO qu'on introduit. En fait on peut l'interpréter comme une version continue de la méthode de Newton, en interprétant y' comme $u_{n+1} - u_n$ (dérivée discrète).*
3. *La continuité du flot par rapport à la donnée initiale à temps fixé résulte facilement du lemme de Gronwall.*

Annexe

Donnons une application de ce théorème. Je ne suis pas sûr mais ça doit se trouver sur le site de J. Lafontaine.

Application 3.1.4. *Il existe $\epsilon > 0$ tel que pour tout $a \in \mathbb{R}^n$, pour tout $L \in GL_n(\mathbb{R})$ si :*

$$\|a\| < \epsilon, \quad et \quad \|L - Id\| < \epsilon,$$

alors il existe un difféomorphisme f de $\mathbb{R}^n$ tel que :

- $\forall \|x\| < 1, \quad f(x) = Lx + a.$
- $\forall \|x\| > 2, \quad f(x) = x.$

Autrement dit f est un isomorphisme affine autour de 0 et f est l'identité assez loin de 0.

Démonstration : On note $\|.\|$ la norme euclidienne usuelle sur $\mathbb{R}^n$. On considère une fonction plateau $\chi \in \mathcal{D}(\mathbb{R})$ qui vaut 1 sur $[-1, 1]$ et 0 sur $[-2, 2]^C$. On définit, pour $x \in \mathbb{R}^n$:

$$f(x) = x + g(\|x\|^2)(Lx + a - x).$$

Cette application est clairement C^1 et propre car c'est l'identité en dehors d'une boule compacte. La différentielle de f est clairement inversible en x si $\|x\| < 1$ ou si $\|x\| < 2$. Soit maintenant x tel que $1 \leq \|x\| \leq 2$. Un calcul montre que pour $h \in \mathbb{R}^n$:

$$df_x.h = h + g(\|x\|^2)(Lh - h) + 2\langle x, h\rangle g'(\|x\|^2)(Lx + a - x).$$

Notons $M = \|g'\|_\infty.$, on a :

$$\|df_x - \mathrm{Id}\| \leq (1 + 4M)\|L - \mathrm{Id}\| + 4M\|a\|.$$

Ainsi si L est assez proche de Id et si a est assez proche de 0, on a :

$$\|df_x - \mathrm{Id}\| < 1,$$

ce qui entraîne classiquement que $df_x \in GL_n(\mathbb{R})$ (série de Neumann). Le théorème de Hadamard-Lévy permet de conclure. $\square$

3.2 Théorèmes de Prokhorov et Lévy

Références : O.Garet et A.Kurtzmann, *De l'intégration aux probabilités.*[1]

Leçons : 203, 229, 241, 262, *261, 250.*

Définition 3.2.1

Soient $(\mu_n)_n$ une suite de probabilités sur $\mathbb{R}^n$ et μ une probabilité sur $\mathbb{R}^n$. On dit que cette suite converge étroitement vers μ si :

$$\forall \phi \in \mathcal{C}_b^0(\mathbb{R}^n, \mathbb{R}) \int_{\mathbb{R}^n} \phi(x)\, d\mu_n(x) \to \int_{\mathbb{R}^n} \phi(x)\, d\mu(x).$$

Définition 3.2.2 (*Tension*)

Soit $(\mu_n)_n$ une suite de probabilités sur $\mathbb{R}^n$. On dit que cette suite est tendue si pour tout $\epsilon > 0$, il existe un compact K tel que :

$$\forall n \in \mathbb{N},\ \mu_n(K) \geq 1 - \epsilon.$$

Cela signifie que la masse de toutes les probabilités de la suite reste localisée dans un compact à ϵ près.

Le théorème suivant est un critère de compacité pour la convergence étroite de probabilités. Nous en aurons besoin pour démontrer le théorème de Lévy.

Théorème 3.2.3 (*Théorème de Prokhorov*)

Soit $(\mu_n)_n$ une suite tendue de probabilités sur $\mathbb{R}$. Alors il existe une extraction de cette suite qui converge étroitement.

Avant de démontrer ce théorème on aura besoin du lemme suviant.

Théorème 3.2.4 (*Théorème de Helly*)

Soit $(F_n)_n$ une suite de fonctions de répartition. Alors il existe une extraction $(\alpha_n)_n$ et une fonction $F : \mathbb{R} \to [0,1]$ croissante et continue à droite telle qu'en tout point de continuité x de F, on a :

$$F_{\alpha_n}(x) \underset{n\to\infty}{\to} F(x).$$

Démonstration : Soit $q \in \mathbb{Q}$, la suite $(F_n(q))_n$ est une suite de $[0,1]$, on peut donc extraire une sous-suite qui converge vers un réel qu'on note $\tilde{F}(q) \in [0,1]$. Par extraction diagonale, il existe une extraction $(\alpha_n)_n$ telle que :

$$\forall q \in \mathbb{Q},\ F_{\alpha_n}(q) \underset{n\to\infty}{\to} \tilde{F}(q).$$

On peut alors définir, si $x \in \mathbb{R}$:

$$F(x) = \inf_{q\in\mathbb{Q},\, x<q} \{\tilde{F}(q)\}.$$

1. Merci à Alain pour les nombreuses discussions sur ce développement !

F est croissante par construction. Montrons qu'elle est continue à droite. Soit $x \in \mathbb{R}$ et $\epsilon > 0$. Il existe $q > x$ tel que $\tilde{F}(q) \leq F(x) + \epsilon$. Ainsi, si $y \in [x, q]$, on a :

$$F(x) \leq F(y) \leq \tilde{F}(q) \leq F(x) + \epsilon.$$

Soit maintenant x un point de continuité de F et $\epsilon > 0$. Par continuité, il existe $y < x$ tel que :

$$F(x) - \epsilon \leq F(y).$$

Soient $r, s \in \mathbb{Q}$ tels que : $y < r < x < s$ et $\tilde{F}(s) \leq F(x) + \epsilon$. On a alors :

$$F(x) - \epsilon \leq F(y) \leq \tilde{F}(r) \leq \tilde{F}(s) \leq F(x) + \epsilon.$$

Or par définition, on a :

$$F(x) - \epsilon \leq \tilde{F}(r) = \lim_n F_{\alpha_n}(r) \leq \underline{\lim}\, F_{\alpha_n}(x) \leq \overline{\lim}\, F_{\alpha_n}(x) \leq \overline{\lim}\, F_{\alpha_n}(s) = \tilde{F}(s) \leq F(x) + \epsilon.$$

En laissant tendre ϵ vers 0, on obtient $F_{\alpha_n}(x) \underset{n \to \infty}{\to} F(x)$. $\square$

Démonstration : (Théorème de Prokhorov).

On note F_n la fonction de répartition de μ_n. Il suffit de montrer que la fonction F obtenue est une fonction de répartition. Par caractérisation de la convergence étroite avec les fonctions de répartition, on aura le résultat [2]. Si on montre que F tend vers 0 en $-\infty$ et vers 1 en $+\infty$, F sera une fonction de répartition grâce à l'inverse généralisé [3]. Fixons $\epsilon > 0$, Par hypothèse il existe $M > 0$, qu'on peut supposer point de continuité de F puisque les discontinuités d'une fonction monotone forment un ensemble au plus dénombrable, tel que :

$$\forall n,\ \mu_n([-M, M]) \geq 1 - \epsilon.$$

En passant à la limite suivant l'extraction du lemme de Helly, on obtient :

$$F(M) - F(-M) \geq 1 - \epsilon.$$

Compte tenu de la monotonie de F, cela assure que F est une fonction de répartition.

$\square$

Le corollaire suivant est un critère de convergence étroite (analogue au critère pour les suites réelles bornées).

Corollaire 3.2.5

Si $(\mu_n)_n$ est tendue et admet une unique valeur d'adhérence pour la convergence étroite, alors $(\mu_n)_n$ converge étroitement vers μ.

2. Voir la fin du développement pour la preuve du critère de convergence étroite avec les fonctions de répartition
3. Voir la fin pour la proposition concernant l'inverse généralisé

Démonstration : Soit f continue bornée. La suite $\left(\int f\, d\mu_n\right)_n$ converge vers $\int f\, d\mu$ puisque de toute extraction de cette suite, on peut extraire une sous-suite qui converge vers $\int f\, d\mu$ grâce au théorème de Prokhorov et car il y a une unique valeur d'adhérence. □

Théorème 3.2.6 (*Théorème de Lévy*)

Soit $(\mu_n)_n$ une suite de probabilités sur $\mathbb{R}$. On note ϕ_n la fonction caractéristique de μ_n. On suppose qu'il existe une fonction ϕ continue en 0 telle que $(\phi_n)_n$ converge simplement vers ϕ. Alors ϕ est la fonction caractéristique d'une probabilité μ et on a convergence étroite de $(\mu_n)_n$ vers μ.

Démonstration : Il s'agit de montrer que la suite $(\mu_n)_n$ est tendue. En effet, cette suite admet une unique valeur d'adhérence puisqu'une telle valeur d'adhérence admet nécessairement ϕ comme fonction caractéristique et que celle-ci caractérise la loi [4]. Le corollaire précédent permet de conclure. Commençons par un calcul. Soit $u > 0$, :

$$\begin{aligned}\frac{1}{2u}\int_{-u}^{u} 1-\phi_n(t)\,dt &= \int_{\mathbb{R}} 1-\frac{\sin(ux)}{ux}\,d\mu_n(x) \quad \text{(Fubini)}\\ &\geq \left(1-\frac{2}{\pi}\right)\mu_n([-\frac{\pi}{2u},\frac{\pi}{2u}]^{\complement}),\end{aligned}$$

en utilisant l'inégalité :

$$\forall |y| \geq \frac{\pi}{2}, \quad \sin(y) \leq \frac{2y}{\pi}.$$

Le théorème de convergence dominée et l'hypothèse de continuité en 0 assurent que :

$$\frac{1}{2u}\int_{-u}^{u} 1-\phi(t)\,dt \underset{u\to 0}{\to} 0.$$

De plus si on fixe u, on a :

$$\frac{1}{2u}\int_{-u}^{u} 1-\phi_n(t)\,dt \underset{u\to 0}{\to} \frac{1}{2u}\int_{-u}^{u} 1-\phi(t)\,dt.$$

On déduit que si $\epsilon > 0$, il existe $u > 0$ et N (qui dépend de u mais ce n'est pas important) tels que :

$$\forall n > N, \quad \mu_n([-\frac{\pi}{2u},\frac{\pi}{2u}]^{\complement}) \leq \epsilon.$$

Puisqu'un ensemble fini de probabilités est tendu, on a démontré la tension de la suite $(\mu_n)_n$ ce qui achève la preuve. □

Remarque 3.2.7. - *Pour que ça tienne en* 15 *minutes, on peut seulement démontrer le lemme de Helly, et le théorème de Lévy, si on a du temps on détaille plus ou moins le théorème de Prokhorov.*

4. Un argument rapide étant l'inversion de Fourier dans les distributions tempérées

- *Le critère de Prokhorov est vrai sur* $\mathbb{R}^n$ *voire dans un cadre plus général mais la preuve est plus compliquée puisqu'on ne dispose pas d'une fonction de répartition comme sur* $\mathbb{R}$.
- *La continuité en* 0 *est cruciale dans le théorème de Lévy. La tension correspond au fait que la masse de la famille de probabilités reste localisée dans un même compact à* ϵ *près, ce qui évite la perte de masse à l'infinie (on voit bien dans la preuve que c'est la tension de la suite de probabilités qui assure que la fonction* F *du lemme de Helly est bien une fonction de répartition, c'est-à-dire qu'il n'y a pas eu perte de masse en passant à la limite). Comme la masse totale d'une mesure finie est égale à l'évaluation en* 0 *de sa fonction caractéristique, il est naturel qu'il y ait une hypothèse sur ce point dans le théorème de Lévy.*
- *On peut utiliser le théorème de Lévy pour démontrer le théorème central limite sur* $\mathbb{R}$.
- *La convergence étroite est métrisable.*

Annexe

Je démontre ici le critère de convergence en loi avec les fonctions de répartition, ainsi que la propriété utilisée sur l'inverse généralisé.

Théorème 3.2.8

Soit $(X_n)_n$ une suite de variables aléatoires et X une variable aléatoire. On note F_n la fonction de répartition de X_n et F celle de X. Alors $(X_n)_n$ converge en loi vers X si et seulement si $(F_n)_n$ converge simplement vers F en chaque point de continuité de F.

Démonstration :

- ($\Rightarrow$) Supposons qu'on ait la convergence en loi. Soit x un point de continuité de F. Soit $\epsilon > 0$, on dispose, par continuité à droite de F en x de $x_1 > x$ tel que : $F(x) \leq F(x_1) \leq F(x) + \epsilon$. Soit f_1 la fonction continue affine par morceaux égale à 1 sur $]-\infty, x]$ et à 0 sur $[x_1, +\infty[$. On a ainsi en intégrant :

$$\forall n,\, F_n(x) \leq \mathbb{E}(f_1(X_n)) \leq F_n(x_1) \quad \text{et} \quad F(x) \leq \mathbb{E}(f_1(X)) \leq F(x_1).$$

  Puisque f_1 est continue bornée, on en déduit qu'il existe $n_1 > 0$ tel que :

$$\forall n \geq n_1,\, \mathbb{E}(f_1(X_n)) \leq \mathbb{E}(f_1(X)) + \epsilon.$$

  En combinant les deux inégalités précédentes, on obtient que pour tout $n \geq n_1$, on a :

$$F_n(x) \leq F(x) + 2\epsilon.$$

  Un raisonnement analogue utilisant la continuité à gauche de F montre qu'il $n_0 > 0$ tel que :

$$\forall n \geq n_0,\, F_n(x) \geq F(x) - 2\epsilon.$$

  Cela assure la convergence simple.

- ($\Leftarrow$) Supposons qu'on a le résultat de convergence simple. On veut montrer la convergence en loi. Pour cela on commence par montrer que

$$\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X)),$$

  pour une fonction C^1 à support compact.

  (a) Soit f une fonction de $C^1_c(\mathbb{R})$. D'après le théorème de Fubini :

$$\begin{aligned}\mathbb{E}(f(X_n)) &= \int_{\mathbb{R}} \left(\int_{-\infty}^{x} f'(y)\, dy \right) d\mathbb{P}_{X_n}(x) \\ &= \int_{\mathbb{R}} (1 - F_n(y)) f'(y)\, dy.\end{aligned}$$

  La fonction F ayant une ensemble de points de discontinuités au plus dénombrable, on peut appliquer le théorème de convergence dominée, ce qui donne que : $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$.

(b) Soit maintenant f une fonction de $C_c^0(\mathbb{R})$. Soit $\epsilon > 0$, on dispose de $g \in C_c^1(\mathbb{R})$ telle que :

$$\|f - g\|_\infty \leq \epsilon.$$

D'après la première étape, on dispose de $n_0 > 0$ tel que :

$$|\mathbb{E}(g(X_n)) - \mathbb{E}(g(X))| \leq \epsilon.$$

On a alors :

$$\begin{aligned}
&|\mathbb{E}(f(X_n)) - \mathbb{E}(f(X))| \\
&\leq |\mathbb{E}(f(X_n)) - \mathbb{E}(g(X_n))| + |\mathbb{E}(g(X_n)) - \mathbb{E}(g(X))| + |\mathbb{E}(g(X)) - \mathbb{E}(f(X))| \\
&\leq 3\epsilon.
\end{aligned}$$

(c) Pour finir, soit $f \in C_b^0$. Soit $\epsilon > 0$, on dispose d'une fonction $h \in C_c^0(\mathbb{R})$ à valeurs dans $[0, 1]$ telle que :

$$\mathbb{E}(h(X)) \geq 1 - \epsilon,$$

par exemple en prenant une fonction plateau continue à support compact valant 1 sur un intervalle du type $[-M, M]$ et en utilisant le théorème de convergence monotone. Par convergence en loi, on dispose de $n_0 > 0$ tel que :

$$\forall n \geq n_0,\ \mathbb{E}(h(X_n)) \geq 1 - 2\epsilon.$$

La fonction fh étant continue à support compact, on a $n_1 > 0$ tel que :

$$\forall n \geq n_1,\ |\mathbb{E}(fh(X_n)) - \mathbb{E}(fh(X))| \leq \epsilon.$$

Ainsi pour tout $n \geq \max(n_0, n_1)$, on a :

$$\begin{aligned}
&|\mathbb{E}(f(X_n)) - \mathbb{E}(f(X))| \\
&\leq |\mathbb{E}(fh(X_n)) - \mathbb{E}(fh(X))| + |\mathbb{E}(f(X_n)(1 - h(X_n)))| + |\mathbb{E}(f(X)(1 - h(X)))| \\
&\leq \epsilon + \|f\|_\infty(\mathbb{E}((1 - h(X_n))) + \mathbb{E}((1 - h(X)))) \\
&\leq \epsilon + 3\epsilon\|f\|_\infty.
\end{aligned}$$

Cela montre bien la convergence en loi.

□

Proposition 3.2.9

Soit F une fonction croissante, continue à droite, qui tend vers 0 en $-\infty$ et vers 1 en $+\infty$. Alors F est une fonction de répartition. Plus précisément : on définit pour tout $x \in]0, 1[$:

$$F^*(x) = \inf\{s \in \mathbb{R},\ F(s) \geq x\},$$

alors si U suit une loi uniforme sur $]0, 1[$, $F^*(U)$ a pour fonction de répartition F.

Démonstration : Remarquons d'abord que l'hypothèse sur les limites de F assure que F^* est bien définie et à valeurs réelles. La proposition découle de l'équivalence suivante :

$$x \leq F(s) \Leftrightarrow F^*(x) \leq s.$$

Le sens direct est immédiat par définition et pour le sens réciproque : si $F^*(x) \leq s$, on prend une suite $(s_n)_n$ de réels strictement plus grands que s qui tend vers s, par définition de F^*, on a :

$$\forall n, \, F(s_n) \geq x.$$

En laissant tendre n vers $+\infty$ et en utilisant la continuité à droite, on a le résultat. □

3.3 Le problème des moments

Références : P. Billingsley, *Probability and Measure* – R. Durrett, Probability : *Theory and Examples*[5]

Leçons : 243, 260, 261.

Définition 3.3.1

Soit μ une probabilité sur $\mathbb{R}$. Soit $n \in \mathbb{N}$, on définit les moments absolus et les moments (lorsque cela est possible) de μ respectivement par :

$$\mu_n = \int_{\mathbb{R}} |x|^n d\mu(x).$$

et

$$m_n = \int_{\mathbb{R}} x^n d\mu(x),$$

L'objet du développement est le théorème suivant.

Théorème 3.3.2

Soit μ une probabilité sur $\mathbb{R}$ admettant des moments (finis) à tout ordre. On suppose que :

$$\limsup_n \frac{(\mu_n)^{1/n}}{n} < +\infty.$$

Alors μ est caractérisée par ses moments.

Démonstration : Le but est de montrer l'analyticité de la fonction caractéristique ce qui suffira à conclure. On note ϕ la fonction caractéristique de la loi μ. Le théorème de régularité sous l'intégrale assure que ϕ est de classe C^∞ sur $\mathbb{R}$ (ce qui n'assure pas l'analyticité ...), de plus on a :

$$\forall n \in \mathbb{N}, \ \frac{\phi^{(n)}(0)}{n!} = i^n m_n.$$

- **Étape 1 :** On se donne $x, t, h \in \mathbb{R}$ et on applique la formule de Taylor à la fonction $s \mapsto e^{isx}$, qui est de classe C^∞, entre 0 et h :

$$e^{ixh} = \sum_{m=0}^{n} \frac{(ix)^m}{m!} h^m + \int_0^h \frac{(h-s)^n}{n!} (ix)^{n+1} e^{isx}\, ds.$$

On déduit donc que :

$$e^{ix(t+h)} = \sum_{m=0}^{n} \frac{(ix)^m}{m!} e^{itx} h^m + \int_0^h \frac{(h-s)^n}{n!} (ix)^{n+1} e^{i(s+t)x}\, ds.$$

On intègre ensuite sur $\mathbb{R}$ pour la mesure μ :

$$\phi(t+h) - \sum_{m=0}^{n} \frac{h^m}{m!} \int_{\mathbb{R}} (ix)^m e^{itx}\, d\mu(x) = \int_{\mathbb{R}} \int_0^h \frac{(h-s)^n}{n!} (ix)^{n+1} e^{i(s+t)x}\, ds\, d\mu(x).$$

5. Également inspiré de la version d'Antoine Diez

On suppose $h \geq 0$, les calculs étant analogues dans l'autre cas. On a par inégalité triangulaire et d'après le théorème de Fubini :

$$\left|\phi(t+h) - \sum_{m=0}^{n} \frac{\phi^{(m)}(t)}{m!} h^m\right| \leq \int_{\mathbb{R}} \int_0^h \frac{|h-s|^n}{n!} |x|^{n+1} \, ds \, d\mu(x)$$
$$= \frac{|h|^{n+1}}{(n+1)!} \mu_{n+1}.$$

- **Etape 2 :** On note $R = \limsup_n \frac{(\mu_n)^{1/n}}{n} < +\infty$. On dispose donc de $N \in \mathbb{N}$ tel que :

$$\forall n \geq N, \; \frac{(\mu_n)^{1/n}}{n} \leq R+1.$$

On a donc :

$$\forall n \geq N, \; \mu_n \leq n^n (R+1)^n.$$

Ainsi si on suppose que $|h| \leq \frac{1}{2e(R+1)}$, on a :

$$\forall n \geq N, \; \frac{|h|^n}{n!} \mu_n \leq \frac{n^n}{n!} |h|^n (R+1)^n.$$

Le développement en série de e^n montre que : $\frac{n^n}{n!} \leq e^n$. Cela assure que :

$$\forall n \geq N, \; \frac{|h|^n}{n!} \mu_n \leq \left(\frac{1}{2}\right)^n.$$

Ainsi :

$$\forall |h| \leq \frac{1}{2e(R+1)}, \quad \phi(t+h) = \sum_{m=0}^{\infty} \frac{\phi^{(m)}(t)}{m!} h^m,$$

ce qui est exactement le caractère analytique de ϕ.

- **Étape 3** On conclut : la fonction ϕ est analytique sur $\mathbb{R}$ sous les hypothèses précédentes. Puisque les dérivées de ϕ en 0 dépendent uniquement des moments de μ on déduit μ est caractérisée par ses moments. En effet si ν est une autre probabilité vérifiant les mêmes hypothèses, alors les deux fonctions caractéristiques coïncident autour de 0 et sont analytiques donc égales. Comme la fonction caractéristique caractérise la loi, on peut conclure que $\mu = \nu$.

□

Corollaire 3.3.3

Une probabilité à support compact est caractérisée par ses moments, c'est en particulier le cas des lois à support fini.

Démonstration : Soit $M > 0$ tel que $\mu([-M, M]) = 1$. Alors $\mu_n \leq M^n$ dont on déduit :

$$\limsup_n \frac{(\mu_n)^{1/n}}{n} = 0.$$

□

Donnons maintenant un exemple et un contre-exemple.

Proposition 3.3.4

La loi $\mathcal{N}(0,1)$ est caractérisée par ses moments.

Démonstration : On a pour tout n par intégration par partie :

$$\begin{aligned}\mu_n &= 2\int_0^\infty x^n e^{-\frac{x^2}{2}}\,\frac{dx}{\sqrt{2\pi}}\\ &= 2\int_0^\infty (n-1)x^{n-2}e^{-\frac{x^2}{2}}\,\frac{dx}{\sqrt{2\pi}}\\ &= (n-1)\mu_{n-2}.\end{aligned}$$

On en déduit qu'il existe une constante C (en réalité on a une expression exacte) telle que : $\forall n \in \mathbb{N}$, $\mu_n \leq C(n!) \leq Cn^n$. Cela assure que : $\limsup_n \frac{(\mu_n)^{1/n}}{n} \leq 1$.

□

Proposition 3.3.5

Soit N une variable aléatoire de loi $\mathcal{N}(0,1)$. Alors e^N est à densité $x \mapsto f(x) = \frac{1}{\sqrt{2\pi}x}e^{-\frac{ln(x)^2}{2}}\mathbf{1}_{x>0}$ par rapport à la mesure de Lebesgue. La loi à densité $x \mapsto g(x) = f(x)(1+\sin(2\pi ln(x)))$ a les mêmes moments que f, mais n'est pas égale à f. Cette loi n'est donc pas caractérisée par ses moments.

Démonstration : Soit $k \in \mathbb{N}$. On a, en posant $y = ln(x)$:

$$\begin{aligned}\int_0^\infty x^k f(x)\sin(2\pi ln(x))\,dx &= \int_{\mathbb{R}} e^{ky}e^{-\frac{y^2}{2}}\sin(2\pi y)\frac{dy}{\sqrt{2\pi}}\\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{y^2-2yk}{2}}\sin(2\pi y)\,dy\\ &= \frac{e^{\frac{k^2}{2}}}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{(y-k)^2}{2}}\sin(2\pi y)\,dy\\ &= \frac{e^{\frac{k^2}{2}}}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{u^2}{2}}\sin(2\pi u)\,du\\ &= 0\end{aligned}$$

f et g ont les mêmes moments mais ne sont pas égales. □

Remarque 3.3.6.
- *Pour que le développement tienne en 15 minutes, on peut se contenter de démontrer le théorème en énonçant le corollaire et selon le temps qu'il reste traiter l'exemple de la gaussienne voire également le contre exemple si on est rapide !*
- *Le fait que la fonction caractéristique caractérise la loi peut se justifier rapidement par inversion de Fourier sur les distributions tempérées.*
- *En utilisant l'inégalité de Cauchy-Schwarz, on obtient :*

$$\mu_{2n+1} = \int_{\mathbb{R}} |x|^{n+1}|x|^n\,d\mu(x) \leq \sqrt{\mu_{2n}\mu_{2n+2}}.$$

Cela permet de montrer que la condition suivante suffit à assurer la caractérisation de la loi par les moments :

$$\limsup_n \frac{(\mu_{2n})^{1/2n}}{2n} < +\infty.$$

Annexe 1

La suite est plutôt " culturelle " bien que très intéressante notamment pour la leçon 262 ! Le théorème nécessite le critère de convergence étroite d'une suite tendue qui a été démontré dans le développement sur les théorèmes de Prokhorov et de Lévy mais que je remets en Annexe 2. On commence d'abord par des résultats généraux sur l'uniforme intégrabilité qu'on pourra trouver dans le livre de O. Garet et A. Kurtzmann : *De l'intégration aux probabilités*. Ces résultats trouvent tout à fait leur place dans la leçon 234 !

Définition 3.3.7

Une famille de variables aléatoires réelles $(X_i)_i$ *est dite uniformément intégrable si :*

$$\lim_{M \to +\infty} \sup_i \mathbb{E}(|X_i| \mathbf{1}_{|X_i| \geq M}) = 0.$$

Lemme 3.3.8

Soit $(X_n)_n$ *une suite de variables aléatoires qui converge en loi vers* X*. Alors :*

$$\mathbb{E}(|X|) \leq \liminf_n \mathbb{E}(|X_n|).$$

Démonstration : D'après le théorème de Fubini-Tonelli, on a :

$$\mathbb{E}(|X_n|) = \int_{\mathbb{R}^+} \mathbb{P}(|X_n| > t)\, dt.$$

Or en tout point t de continuité de la fonction de répartition de X, on a [6] :

$$\mathbb{P}(|X_n| > t) \to \mathbb{P}(|X| > t).$$

Comme l'ensemble des discontinuités d'une fonction monotone est dénombrable, on peut appliquer le lemme de Fatou :

$$\mathbb{E}(|X|) = \int_{\mathbb{R}^+} \liminf_n \mathbb{P}(|X_n| > t)\, dt \leq \liminf_n \mathbb{E}(|X_n|).$$

□

Théorème 3.3.9

Soit $(X_n)_n$ *une suite de variables aléatoires uniformément intégrable qui converge en loi vers* X*. Alors* $X \in L^1$ *et on a :*

$$\mathbb{E}(X_n) \to \mathbb{E}(X).$$

6. Pour la preuve on peut consulter le développement sur les théorèmes de Prokhorov et de Lévy

Démonstration : Puisqu'une famille uniformément intégrable est bornée dans L^1, on sait déjà que X est intégrable par le lemme précédent. Soit $\epsilon > 0$. On dispose de $M > 0$ tel que :

$$\sup_n \mathbb{E}(|X_n|\mathbf{1}_{|X_n|\geq M}) \leq \epsilon \quad \text{et} \quad \mathbb{E}(|X|\mathbf{1}_{|X|\geq M}) \leq \epsilon.$$

On a alors successivement, en notant $\wedge$ le minimum :

$$\begin{aligned}
|\mathbb{E}(X_n) - \mathbb{E}(X)| &= |\mathbb{E}(X_n \wedge M) - \mathbb{E}(X \wedge M) + \mathbb{E}((X_n - M)\mathbf{1}_{X_n \geq M}) - \mathbb{E}((X - M)\mathbf{1}_{X \geq M})\\
&\leq |\mathbb{E}(X_n \wedge M) - \mathbb{E}(X \wedge M)| + \mathbb{E}(|X_n - M|\mathbf{1}_{X_n \geq M}) + \mathbb{E}(|X - M|\mathbf{1}_{X \geq M})\\
&\leq |\mathbb{E}(X_n \wedge M) - \mathbb{E}(X \wedge M)| + \mathbb{E}(|X_n|\mathbf{1}_{X_n \geq M}) + \mathbb{E}(|X|\mathbf{1}_{X \geq M})\\
&\leq |\mathbb{E}(X_n \wedge M) - \mathbb{E}(X \wedge M)| + 2\epsilon.
\end{aligned}$$

La fonction $x \mapsto x \wedge M$ étant continue bornée, on a : $\mathbb{E}(X_n \wedge M) \to \mathbb{E}(X \wedge M)$. Ainsi :

$$\limsup_n |\mathbb{E}(X_n) - \mathbb{E}(X)| \leq 2\epsilon,$$

D'où le résultat. □

J'énonce ici un corollaire que je n'utilise pas dans la suite mais qui est un résultat " classique ".

Théorème 3.3.10

Soit $(X_n)_n$ une suite de variables aléatoires uniformément intégrable qui converge en probabilité vers X. Alors $X \in L^1$ et on a :

$$X_n \underset{L^1}{\to} X.$$

Démonstration : Par hypothèse la suite $(|X_n - X|)_n$ converge en probabilité vers 0 donc également en loi. Il suffit d'appliquer le théorème précédent pour conclure. □

Passons maintenant au résultat de convergence grâce aux moments.

Théorème 3.3.11 (*Méthode des moments*)

Soit $(X_n)_n$ une suite de variables aléatoires admettant des moments à tout ordre. Soit X une variable aléatoire admettant également des moments à tout ordre et telle que la loi μ de X est caractérisée par ses moments. On suppose que :

$$\forall r \in \mathbb{N},\ \mathbb{E}(X_n^r) \to \mathbb{E}(X^r) := m_r.$$

Alors $(X_n)_n$ converge en loi vers X.

Démonstration : On note μ_n la loi de X_n et μ celle de X. D'après le corollaire de l'annexe, pour prouver le théorème, il suffit de montrer que la suite $(\mu_n)_n$ est tendue et admet μ comme unique valeur d'adhérence (pour la convergence en loi).

- **Tension :** Il suffit de voir que la suite $(X_n)_n$ est bornée dans L^1 par hypothèse. On note $M = \sup_n \mathbb{E}(|X_n|)$. L'inégalité de Markov assure que si $x > 0$:

$$\forall n \in \mathbb{N},\ \mathbb{P}(|X_n| \geq x) \leq \frac{M}{x},$$

ce qui implique directement la tension de la suite $(\mu_n)_n$.

- μ **est la seule valeur d'adhérence :** Soit ν une autre valeur d'adhérence de la suite $(\mu_n)_n$ et Y une variable aléatoire de loi ν. On dispose d'une extraction $(n_k)_k$ telle que $(X_{n_k})_k$ converge en loi vers Y. Si on montre l'uniforme intégrabilité de la suite $(|X_n|^r)_n$ pour tout $r \in \mathbb{N}$, on aura, d'après un résultat précédent :

$$\mathbb{E}(X_{n_k}^r) \underset{k \to +\infty}{\to} \mathbb{E}(Y^r)$$

puisque $(X_n^r)_n$ converge en loi vers X^r. Par unicité de la limite, on a alors :

$$\forall r \in \mathbb{N},\ \mathbb{E}(Y^r) = m_r.$$

Ainsi puisque X est caractérisée par ses moments, on obtient que $\nu = \mu$, ce qui achèvera la preuve. On se donne donc $r \in \mathbb{N}$. On note $M_{2r} = \sup_n \mathbb{E}(|X_n|^{2r})$. On a alors pour tout $C > 0$:

$$\mathbb{E}(|X_n|^r \mathbf{1}_{|X_n|^r \geq C}) \leq \frac{1}{C^r}\mathbb{E}(|X_n|^{2r}) \leq \frac{M_{2r}}{C^r}.$$

Cette majoration assure l'uniforme intégrabilité de $(X_n^r)_n$ pour tout $r \in \mathbb{N}$.

$\square$

Annexe 2

Je remets ici la preuve du critère de convergence étroite d'une suite tendue utilisé, on pourra trouver l'intégralité dans le développement sur les théorèmes de Prokhorov et de Lévy ou dans *De l'intégration aux probabilités.*

Définition 3.3.12

Soient $(\mu_n)_n$ une suite de probabilités sur $\mathbb{R}^n$ et μ une probabilité sur $\mathbb{R}^n$. On dit que cette suite converge étroitement vers μ si :

$$\forall \phi \in \mathcal{C}_b^0(\mathbb{R}^n, \mathbb{R}) \int_{\mathbb{R}^n} \phi(x)\, d\mu_n(x) \to \int_{\mathbb{R}^n} \phi(x)\, d\mu(x).$$

Définition 3.3.13 (*Tension*)

Soit $(\mu_n)_n$ une suite de probabilités sur $\mathbb{R}^n$. On dit que cette suite est tendue si pour tout $\epsilon > 0$, il existe un compact K tel que :

$$\forall n \in \mathbb{N},\ \mu_n(K) \geq 1 - \epsilon.$$

Le théorème suivant est un critère de compacité pour la convergence étroite de probabilités. Nous en aurons besoin pour démontrer le théorème de Lévy.

Théorème 3.3.14 (*Théorème de Prokhorov*)

Soit $(\mu_n)_n$ une suite tendue de probabilités sur $\mathbb{R}$. Alors il existe une extraction de cette suite qui converge étroitement.

Avant de démontrer ce théorème on aura besoin du lemme suviant.

Théorème 3.3.15 (*Théorème de Helly*)

Soit $(F_n)_n$ une suite de fonctions de répartition. Alors il existe une extraction $(\alpha_n)_n$ et une fonction $F : \mathbb{R} \to [0,1]$ croissante et continue à droite telle qu'en tout point de continuité x de F, on a :

$$F_{\alpha_n}(x) \underset{n\to\infty}{\to} F(x).$$

Démonstration : Soit $q \in \mathbb{Q}$, la suite $(F_n(q))_n$ est une suite de $[0,1]$, on peut donc extraire une sous-suite qui converge vers un réel qu'on note $\tilde{F}(q) \in [0,1]$. Par extraction diagonale, il existe une extraction $(\alpha_n)_n$ telle que :

$$\forall q \in \mathbb{Q},\ F_{\alpha_n}(q) \underset{n\to\infty}{\to} \tilde{F}(q).$$

On peut alors définir, si $x \in \mathbb{R}$:

$$F(x) = \inf_{q\in\mathbb{Q},\, x<q} \{\tilde{F}(q)\}.$$

F est croissante par construction. Montrons qu'elle est continue à droite. Soit $x \in \mathbb{R}$ et $\epsilon > 0$. Il existe $q > x$ tel que $\tilde{F}(q) \leq F(x) + \epsilon$. Ainsi, si $y \in [x, q]$, on a :

$$F(x) \leq F(y) \leq \tilde{F}(q) \leq F(x) + \epsilon.$$

Soit maintenant x un point de continuité de F et $\epsilon > 0$. Par continuité, il existe $y < x$ tel que :

$$F(x) - \epsilon \leq F(y).$$

Soient $r, s \in \mathbb{Q}$ tels que : $y < r < x < s$ et $\tilde{F}(s) \leq F(x) + \epsilon$. On a alors :

$$F(x) - \epsilon \leq F(y) \leq \tilde{F}(r) \leq \tilde{F}(s) \leq F(x) + \epsilon.$$

Or par définition, on a :

$$F(x) - \epsilon \leq \tilde{F}(r) = \lim_n F_{\alpha_n}(r) \leq \underline{\lim}\, F_{\alpha_n}(x) \leq \overline{\lim}\, F_{\alpha_n}(x) \leq \overline{\lim}\, F_{\alpha_n}(s) = \tilde{F}(s) \leq F(x) + \epsilon.$$

En laissant tendre ϵ vers 0, on obtient $F_{\alpha_n}(x) \underset{n\to\infty}{\to} F(x)$. □

Démonstration : (Théorème de Prokhorov).

On note F_n la fonction de répartition de μ_n. Il suffit de montrer que la fonction F obtenue est une fonction de répartition. Par caractérisation de la convergence étroite avec les fonctions de répartition, on aura le résultat. Si on montre que F tend vers 0 en $-\infty$ et vers 1 en $+\infty$, F sera une fonction de répartition grâce à l'inverse généralisé. Fixons $\epsilon > 0$, Par hypothèse il existe $M > 0$, qu'on peut supposé point de continuité de F puisque les discontinuités d'une fonction monotone forment un ensemble au plus dénombrable, tel que :

$$\forall n,\ \mu_n([-M, M]) \geq 1 - \epsilon.$$

En passant à la limite suivant l'extraction du lemme de Helly, on obtient :

$$F(M) - F(-M) \geq 1 - \epsilon.$$

Compte tenu de la monotonie de F, cela assure que F est une fonction de répartition.

□

Le corollaire suivant est un critère de convergence étroite (analogue au critère pour les suites réelles bornés, d'ailleurs la preuve en découle).

Corollaire 3.3.16

Si $(\mu_n)_n$ est tendue et admet une unique valeur d'adhérence pour la convergence étroite, alors $(\mu_n)_n$ convergence étroitement vers μ.

Démonstration : Soit f continue bornée. La suite $\left(\int f\,d\mu_n\right)_n$ converge vers $\int f\,d\mu$ puisque de toute extraction de cette suite, on peut extraire une sous-suite qui converge vers $\int f\,d\mu$ grâce au théorème de Prokhorov et car il y a une unique valeur d'adhérence. □

3.4 Optimisation dans un Hilbert

Référence : : P.G. Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*[7]

Leçons : 203, 205, 208, 213, 219, 229, 253.

Définition 3.4.1

Soit H un espace de Hilbert et $(x_n)_n$ une suite de H. On dit que $(x_n)_n$ converge faiblement vers $x \in H$ si :

$$\forall y \in H, \, \langle x_n, y \rangle \to \langle x, y \rangle.$$

On note alors : $x_n \rightharpoonup x$.

On commence par démontrer le résultat de compacité faible suivant.

Théorème 3.4.2 (*Théorème de Banach-Alaoglu*)

Soit H un espace de Hilbert séparable et $(x_n)_n$ une suite bornée de H. Alors il existe une extraction $(n_k)_k$ et $x \in H$ tels que : $x_{n_k} \rightharpoonup x$.

Démonstration : On considère une suite dense qu'on note $(h_n)_n$ et on note $M = \sup_n \|x_n\|$. Pour tout k, la suite $(\langle x_n, h_k \rangle)_n$ est bornée. On peut donc en extraire une sous-suite convergeant vers un réel noté $\phi(h_k)$. Par extraction diagonale, on dispose d'une extraction $(n_j)_j$ telle que :

$$\forall k \in \mathbb{N}, \, \langle x_{n_j}, h_k \rangle \underset{j \to +\infty}{\to} \phi(h_k) \quad (*).$$

Montrons maintenant que si $y \in H$, alors la suite $(\langle x_{n_j}, y \rangle)_j$ converge. Par complétude, il suffit de vérifier le critère de Cauchy. Soit $\epsilon > 0$. Par densité, on dispose de $k \in \mathbb{N}$ tel que : $\|y - h_k\| \leq \epsilon$. De plus, d'après la convergence $(*)$, on a $N \in \mathbb{N}$ tel que :

$$\forall p, q \geq N, \, |\langle x_{n_p}, h_k \rangle - \langle x_{n_q}, h_k \rangle| \leq \epsilon.$$

Ainsi on a pour tout $p, q \geq N$:

$$\begin{aligned} |\langle x_{n_p}, y \rangle - \langle x_{n_q}, y \rangle| &\leq |\langle x_{n_p}, y \rangle - \langle x_{n_p}, h_k \rangle| + |\langle x_{n_p}, h_k \rangle - \langle x_{n_q}, h_k \rangle| + |\langle x_{n_q}, h_k \rangle - \langle x_{n_q}, y \rangle| \\ &= |\langle x_{n_p}, y - h_k \rangle| + |\langle x_{n_p}, h_k \rangle - \langle x_{n_q}, h_k \rangle| + |\langle x_{n_q}, h_k - y \rangle \\ &\leq \epsilon + 2M\epsilon, \end{aligned}$$

d'après l'inégalité de Cauchy-Schwarz. Cela montre la convergence annoncée. On a donc une application $\phi : H \to \mathbb{R}$ telle que :

$$\forall y \in H, \, \langle x_{n_j}, y \rangle \underset{j \to +\infty}{\to} \phi(y) \quad (**).$$

ϕ est linéaire comme limite simple de formes linéaires. Elle est également continue. En effet, pour tout $y \in H$, on a d'après l'inégalité de Cauchy-Schwarz :

$$|\phi(y)| = |\lim_{j \to +\infty} \langle x_{n_j}, y \rangle| \leq M\|y\|.$$

7. Le développement n'y est pas rédigé, comme ça tout du moins, je me suis servi de la version rédigée par Corentin Kilque.

Le théorème de Riesz assure qu'il existe $x \in H$ tel que :

$$\forall y \in H,\ \phi(y) = \langle x, y \rangle.$$

La convergence $(**)$ assure la convergence faible de l'extraction vers x.

$\square$

Remarque 3.4.3. *On peut se passer de l'hypothèse de séparabilité dans le cadre hilbertien quitte à travailler dans le Hilbert séparable $\tilde{H} := \overline{Vect\{x_n,\, n \in \mathbb{N}\}}$. On effet soit G le supplémentaire orthogonal de ce sous-espace vectoriel fermé. D'après ce qui précède, il existe $x \in \tilde{H}$ et une extraction $(n_j)_j$ tels que :*

$$\forall y \in \tilde{H},\ \langle x_{n_j}, y \rangle \underset{j \to +\infty}{\to} \langle x, y \rangle.$$

Soit maintenant $y = \tilde{y} + g$ la décomposition fournie par : $H = \tilde{H} \oplus G$. On a alors par orthogonalité :

$$\begin{aligned}\langle x_{n_j}, y \rangle &= \langle x_{n_j}, \tilde{y} + g \rangle \\ &= \langle x_{n_j}, \tilde{y} \rangle \\ &\to \langle x, \tilde{y} \rangle = \langle x, \tilde{y} + g \rangle = \langle x, y \rangle.\end{aligned}$$

D'où le résultat.

Théorème 3.4.4

Soit H un espace de Hilbert et $J : H \to \mathbb{R}$ une fonction convexe continue et coercive, i.e telle que :

$$J(x) \underset{\|x\| \to +\infty}{\to} +\infty.$$

Alors J atteint son minimum en un point x^*.

Démonstration : On considère une suite minimisante $(x_n)_n$, i.e $J(x_n) \to \inf_H J$. Cette suite est bornée. En effet dans le cas contraire, on a une extraction $(n_k)_k$ telle que :

$$\|x_{n_k}\| \to +\infty.$$

Par coercivité, on aurait : $\inf_H J = +\infty$, ce qui est absurde. D'après le théorème de Banach-Alaoglu, on dispose de $x_* \in H$ et d'une extraction $(n_k))_k$ tels que :

$$x_{n_k} \rightharpoonup x_*.$$

Montrons maintenant que :

$$J(x_*) = \inf_H J.$$

Soit $\alpha > \inf_H J$. On pose :

$$C_\alpha = J^{-1}(]-\infty, \alpha]),$$

qui est non vide par choix de α. La continuité de J assure que C_α est fermé et la convexité de J assure que C_α est convexe. On peut donc bien définir P_α la projection sur C_α. Puisque $J(x_{n_k}) \to \inf_H J$, on dispose de $N > 0$ tel que : $\forall k \geq N,\ x_{n_k} \in C_\alpha$. La propriété des angles obtus assure que :

$$\forall k \geq N,\ \langle x_* - P_\alpha(x_*), x_{n_k} - P_\alpha(x_*) \rangle \leq 0.$$

En laissant tendre k vers $+\infty$, on déduit que :

$$\|x_* - P_\alpha(x_*)\|^2 \leq 0,$$

ce qui montre que :

$$x_* \in C_\alpha, \quad \forall \alpha > \inf_H J.$$

D'où $J(x_*) \leq \inf\limits_H J$ et l'égalité est démontrée.

$\square$

Annexe

L'application suivante permet de résoudre un problème aux limites non linéaire au sens faible grâce à la formulation variationnelle associée, dans un cas où le théorème de Lax-Milgram ne s'applique pas. Je n'ai pas trouvé de référence pour cette partie, l'énoncé est issu du cours d'un cours de Karine Beauchard.

Proposition 3.4.5

Soit $f \in L^2(0,1)$ et $p > 0$, alors il existe une unique $u \in H_0^1(0,1)$ tel que :

$$-u" + u|u|^{p-1} = f,$$

au sens des distributions.

Démonstration :

- **Existence :** Commençons par trouver la formulation variationnelle. Pour cela, on pose :

$$J : \left\{ \begin{array}{ll} H_0^1(0,1) & \to \mathbb{R} \\ u & \mapsto \displaystyle\int_0^1 \frac{|u'|^2}{2} + \frac{|u|^{p+1}}{p+1} - fu \; dx \end{array} \right. .$$

(a) J est différentiable et on a :

$$\forall u, v \in H, \; dJ_u.v = \int_0^1 u'v' + |u|^{p-1}uv - fv \; dx.$$

Le membre de droite est une forme linéaire continue sur $H_0^1(0,1)$ en utilisant l'inégalité de Cauchy-Schwarz et le fait que :

$$H_0^1(0,1) \subset \mathcal{C}([0,1]).$$

Montrons qu'il s'agit bien de la différentielle de J. Un simple calcul montre que :

$$J(u+v) = J(u) + \int_0^1 u'v' + \frac{|u+v|^{p+1}}{p+1} - \frac{|u|^{p+1}}{p+1} - fv \; dx + \frac{1}{2}\|v'\|_{L^2}^2.$$

Il suffit donc de montrer que :

$$\int_0^1 \frac{|u+v|^{p+1}}{p+1} - \frac{|u|^{p+1}}{p+1} - |u|^{p-1}uv \; dx = o(\|v\|_{H_0^1}).$$

Fixons $u, v \in \mathbb{R}$ et posons $\phi : t \in [0,1] \mapsto \frac{|u+tv|^{p+1}}{p+1} - \frac{|u|^{p+1}}{p+1} - t|u|^{p-1}uv$. L'objectif est de contrôler $|\phi(1)| = |\phi(1) - \phi(0)|$. Remarquons qu'on a :

$$\frac{d}{dt}|u+tv|^{p+1} = v(p+1)\text{sgn}(u+tv)|u+tv|^p,$$

en notant sgn la fonction signe et en convenant que $\text{sgn}(0) = 0$. ϕ est donc dérivable et on a :

$$\forall t \in [0,1], \; \phi'(t) = v(p+1)\text{sgn}(u+tv)|u+tv|^p - u|u|^{p-1}v.$$

L'inégalité des accroissements finis assure alors que :

$$\left|\frac{|u+v|^{p+1}}{p+1}-\frac{|u|^{p+1}}{p+1}-|u|^{p-1}uv\right| \leq |v| \sup_{|t|\leq|v|} \{|\mathrm{sgn}(u+t)|u+t|^p - \mathrm{sgn}(u)|u|^p|\}\,.$$

On a alors d'après ce qui précède et l'inégalité de Cauchy-Schwarz :

$$\begin{aligned}
&\left|\int_0^1 \frac{|u+v|^{p+1}}{p+1}-\frac{|u|^{p+1}}{p+1}-|u|^{p-1}uv\;dx\right| \\
&\leq \int_0^1 |v| \sup_{|t|\leq\|v\|_\infty} \{|\mathrm{sgn}(u+t)|u+t|^p - \mathrm{sgn}(u)|u|^p|\}\;dx \\
&\leq \|v\|_{L^2}\left(\int_0^1\left(\sup_{|t|\leq\|v\|_\infty} \{|\mathrm{sgn}(u+t)|u+t|^p - \mathrm{sgn}(u)|u|^p|\}\right)^2 dx\right)^{1/2} \\
&\leq \|v\|_{H_0^1}\left(\int_0^1\left(\sup_{|t|\leq\|v\|_\infty} \{|\mathrm{sgn}(u+t)|u+t|^p - \mathrm{sgn}(u)|u|^p|\}\right)^2 dx\right)^{1/2}
\end{aligned}$$

Pour montrer le résultat, il suffit de voir que :

$$\int_0^1\left(\sup_{|t|\leq\|v\|_\infty} \{|\mathrm{sgn}(u+t)|u+t|^p - \mathrm{sgn}(u)|u|^p|\}\right)^2 dx \underset{\|v\|_{H_0^1}\to 0}{\to} 0.$$

Puisqu'on a : $\|v\|_\infty \leq \|v\|_{H_0^1}$, on a pour tout $x\in[0,1]$:

$$\sup_{|t|\leq\|v\|_\infty} \{|\mathrm{sgn}(u(x)+t)|u(x)+t|^p - \mathrm{sgn}(u(x))|u(x)|^p|\} \underset{\|v\|_{H_0^1}\to 0}{\to} 0,$$

par continuité de la fonction $x\mapsto \mathrm{sgn}(x)|x|^p$. On ne perd pas de généralité à supposer $\|v\|_{H_0^1}\leq 1$, on peut ainsi dominer l'intégrande par :

$$((|u|+1)^p+|u|^p)^2 \in L^1(0,1),$$

puisque u est continue. Le théorème de convergence dominée assure que :

$$\int_0^1\left(\sup_{|t|\leq\|v\|_\infty} \{|\mathrm{sgn}(u+t)|u+t|^p - \mathrm{sgn}(u)|u|^p|\}\right)^2 dx \underset{\|v\|_{H_0^1}\to 0}{\to} 0,$$

ce qui est le résultat voulu.

(b) J est strictement convexe. La convexité est claire compte tenu de la convexité de l'intégrande. Soit $\lambda\in]0,1[$ et $u,v\in H_0^1(0,1)$. Supposons que :

$$J(\lambda u+(1-\lambda)v) = \lambda J(u)+(1-\lambda)J(v).$$

Alors on a :

$$\int_0^1 |\lambda u+(1-\lambda)v|^{p+1}\,dx = \int_0^1 \lambda|u|^{p+1}+(1-\lambda)|v|^{p+1}\,dx.$$

D'où l'on déduit que :

$$|\lambda u+(1-\lambda)v|^{p+1} = \lambda|u|^{p+1}+(1-\lambda)|v|^{p+1},$$

presque-partout (et même partout car $H_0^1\subset C^0$) par convexité de $t\mapsto|t|^{p+1}$ et positivité de l'intégrale. La stricte convexité de $t\mapsto|t|^{p+1}$ montre que $u=v$.

(c) J est coercive. En effet, l'inégalité de Poincaré assure qu'il existe $C > 0$ telle que :

$$\forall u \in H_0^1(0,1),\ 2C\|u\|_{H_0^1}^2 \leq \|u'\|_{L^2}^2.$$

Ainsi on a :

$$\begin{aligned} J(u) &\geq \int_0^1 \frac{1}{2}|u'|^2\,dx - \|u\|_{L^2}\|f\|_{L^2} \\ &\geq C\|u\|_{H_0^1}^2 - \|u\|_{L^2}\|f\|_{L^2} \underset{\|u\|_{H_0^1}\to+\infty}{\longrightarrow} +\infty. \end{aligned}$$

D'après le théorème d'optimisation, il existe $u \in H_0^1(0,1)$ qui minimise J. On a donc la condition nécessaire d'optimalité suivante :

$$\forall v \in H_0^1(0,1),\ \int_0^1 u'v' + |u|^{p-1}uv - fv\ dx = 0,$$

donc on a bien : $-u" + u|u|^{p-1} = f$ au sens des distributions.

- **Unicité :** Soit $\tilde{u}$ une autre solution au sens dans distributions dans $H_0^1(0,1)$. Alors, on a par définition :

$$\forall v \in \mathcal{D}(0,1),\ dJ_{\tilde{u}}.v = 0.$$

Par densité de $\mathcal{D}(0,1)$ dans $H_0^1(0,1)$ pour la norme $\|.\|_{H_0^1}$ et par continuité de la forme linéaire $dJ_{\tilde{u}}$ sur $H_0^1(0,1)$, on déduit que $\tilde{u}$ est un point critique de J qui est unique par stricte convexité de J. D'où $u = \tilde{u}$.

$\square$

Remarque 3.4.6. - *L'inégalité de Poincaré sur $H_0^1(0,1)$ résulte simplement de l'identité :*

$$u(x) = \int_0^x u'(t)\,dt,$$

ce qui implique que : $\|u\|_{L^2} \leq \|u'\|_{L^2}$, par Cauchy-Schwarz. On peut en trouver une constante optimale grâce aux séries de Fourier.

- *La solution à cette EDO aux conditions de bords nulles est l'unique solution d'un problème de minimisation de l'énergie (la fonctionnelle J).*
- *Si f est continue, alors la solution u est en fait une solution au sens classique de classe C^2.*

3.5 Équation de la chaleur sur le cercle

Références : H. Dym et H.P. McKean, *Fourier Series and Integrals*[8]

Leçons : 222, 235, 241, 246, *209*

Notons $\mathcal{C}^0_{2\pi,x}(\mathbb{R}^+ \times \mathbb{R}, \mathbb{C})$ l'espace des fonctions continues sur $\mathbb{R}^+_t \times \mathbb{R}_x$ et 2π-périodiques par rapport à la seconde variable x.

Théorème 3.5.1

Soit $f \in \mathcal{C}^0_{2\pi}(\mathbb{R}, \mathbb{C})$ une fonction continue, 2π-périodique. Alors il existe une unique solution u du problème suivant :

- $u \in \mathcal{C}^0_{2\pi,x}(\mathbb{R}^+ \times \mathbb{R}, \mathbb{C})$
- $u \in \mathcal{C}^2_{2\pi,x}(\mathbb{R}^+_* \times \mathbb{R}, \mathbb{C})$
- $\forall (t,x) \in \mathbb{R}^+_* \times \mathbb{R}$, $\partial_t u(t,x) = \partial_x^2 u(t,x)$
- $u(0,.) = f$

De plus, la solution u est en fait dans $\mathcal{C}^\infty_{2\pi,x}(\mathbb{R}^+_* \times \mathbb{R}, \mathbb{C})$

Démonstration : On raisonne par analyse-synthèse.

- **Étape 1 : Analyse et unicité.**

Supposons qu'une telle solution u existe. Puisque pour tout $t > 0$, $u(t,.)$ est de classe $\mathcal{C}^1$, cette fonction est donc somme de sa série de Fourier. On note pour $t > 0$ et $n \in \mathbb{Z}$:

$$c_n(t) := \frac{1}{2\pi} \int_{-\pi}^{\pi} u(t,x) e^{-inx}\, dx.$$

On a alors :

$$\forall t > 0,\, \forall x \in \mathbb{R},\, u(t,x) = \sum_{n \in \mathbb{Z}} c_n(t) e^{inx}.$$

Or les fonctions c_n sont C^1 sur $\mathbb{R}^+_*$. En effet, le théorème de régularité des intégrales à paramètres s'applique car :

(a) $\forall x \in [-\pi, \pi]$, $u(.,x)e^{-inx}$ est C^1.

(b) Si $0 < a < b$, alors :

$$\forall t \in [a,b],\, \forall x \in \mathbb{R},\, |u(.,x)e^{-inx}| \leq \sup_{t \in [a,b],\, x \in \mathbb{R}} |\partial_x^2 u(t,x)|,$$

puisque u vérifie l'équation de la chaleur. Cette constante est finie car u est supposée $\mathcal{C}^2$ sur $\mathbb{R}^+_* \times \mathbb{R}$ et périodique en x.

8. Également inspiré de la version de Benjamin Havret.

Ainsi si on fixe $t > 0$, $n \in \mathbb{Z}$:

$$\begin{aligned} c_n'(t) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \partial_t u(t,x)e^{-inx}\,dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \partial_x^2 u(t,x)e^{-inx}\,dx \\ &= -n^2\frac{1}{2\pi}\int_{-\pi}^{\pi} u(t,x)e^{-inx}\,dx \quad \text{après deux IPPs} \\ &= -n^2 c_n(t). \end{aligned}$$

Il existe alors $C_n \in \mathbb{C}$ tels que :

$$c_n(t) = C_n e^{-n^2 t}.$$

Le théorème de continuité sous l'intégrale ainsi que les hypothèses faites montrent que les fonctions c_n sont continues sur $\mathbb{R}^+$ si bien que :

$$c_n(t) = c_n(f)e^{-n^2 t}.$$

On a bien montré l'unicité car la solution u est nécessairement :

$$u(t,x) = \begin{cases} f(x) & \text{si } t = 0 \\ \displaystyle\sum_{n\in\mathbb{Z}} c_n(f)e^{inx-n^2t} & \text{sinon.} \end{cases} \qquad (*)$$

- **Étape 2 : Existence sans la continuité en $t = 0$**

On définit u par $(*)$. La suite des coefficients de Fourier étant bornée, on montre que les séries des dérivées partielles de $(t,x) \mapsto c_n(f)e^{inx-n^2t}$ convergent normalement sur tout compact de $\mathbb{R}_*^+ \times \mathbb{R}$ (décroissance exponentielle). Cela montre que u est $\mathcal{C}^\infty$ sur $\mathbb{R}_*^+ \times \mathbb{R}$. Elle est évidemment 2π-périodique et vérifie l'équation de la chaleur en dérivant sous le signe somme. Il reste donc à vérifier la continuité en 0 pour achever la preuve.

- **Étape 3 : Continuité en 0 pour f de classe $\mathcal{C}^1$**

Si on suppose f de classe $\mathcal{C}^1$, alors on a :

$$\sum_{n\in\mathbb{Z}} |c_n(f)| < +\infty.$$

La série définissant u converge donc normalement sur $\mathbb{R}^+ \times \mathbb{R}$ et la somme vaut f en 0, ce qui montre la continuité de u.

- **Étape 4 : Principe du maximum**

Si f est C^1 et positive, alors la solution u est également positive. Le fait qu'elle soit réelle se vérifie facilement en conjuguant la somme et en utilisant que : $\overline{c_n(f)} = c_{-n}(f)$ pour f réelle. On raisonne par l'absurde : s'il existe $(t_0,x_0) \in \mathbb{R}_*^+ \times \mathbb{R}$ tel que $u(t_0,x_0) < 0$. Soit $\beta < 0$, l'application $v : (t,x) \in [0,t_0] \times \mathbb{R} \mapsto u(t,x)e^{\beta t}$ admet un minimum global strictement négatif atteint en (t_1, x_1), où $t_1 > 0$ par hypothèse. Comme il s'agit d'un minimum et que l'intervalle est fermé à droite, on a :

$$\partial_t v(t_1, x_1) \leq 0, \quad \text{et} \quad \partial_x^2 v(t_1, x_1) \geq 0.$$

Or on a :

$$\left\{ \begin{array}{ll} \partial_t v(t_1, x_1) & = \partial_t u(t_1, x_1) + \beta u(t_1, x_1) \leq 0 \\ \partial_x^2 v(t_1, x_1) & = \partial_x^2 u(t_1, x_1) e^{\beta t_1} \geq 0 \end{array} \right. \quad (*) .$$

Puisque : $\beta u(t_1, x_1) > 0$, la première inégalité assure que : $\partial_t u(t_1, x_1) < 0$. Or la seconde inégalité entraine que : $\partial_x^2 u(t_1, x_1) = \partial_t u(t_1, x_1) \geq 0$. Ceci est une contradiction.

- **Étape 5 : Étude du Noyau de la chaleur**

Pour $(t, x) \in \mathbb{R}_*^+ \times \mathbb{R}$, on pose :

$$p_t(x) = \sum_{n \in \mathbb{Z}} e^{inx - n^2 t}.$$

Pour tout $t > 0$, p_t est continue et 2π-périodique et en échangeant le signe somme et intégrale par un argument de convergence normale :

$$u(t, .) = p_t * f,$$

où $*$ désigne la convolution sur le cercle. Notons $(K_n)_n$ le noyau de Féjèr. Puisque p_t est continue, le théorème de Féjèr assure que :

$$\|p_t * K_N - p_t\|_\infty \underset{N \to +\infty}{\to} 0.$$

Comme K_N est positif et $\mathcal{C}^1$, le principe du maximum assure que pour tout N, $p_t * K_N \geq 0$. On déduit des deux faits précédents que :

$$\forall t > 0, \, p_t \geq 0.$$

On a donc pour $t > 0$:

$$\|p_t\|_{L^1} = c_0(p_t) = 1.$$

- **Étape 6 : Continuité en 0 pour $f \in \mathcal{C}_{2\pi}^0$**

Si $t > 0$, on a :

$$|u(t, x)| = |p_t * f(x)| \leq \|p_t\|_{L^1} \|f\|_\infty \leq \|f\|_\infty,$$

inégalité valable également pour $t = 0$. La solution u de notre problème est donc bornée. Cela permet de définir l'opérateur solution :

$$\Delta = \left\{ \begin{array}{ll} (\mathcal{C}_{2\pi}^0(\mathbb{R}, \mathbb{C}), \|.\|_\infty) & \to L_{2\pi,x}^\infty(\mathbb{R}_t^+ \times \mathbb{R}_x, \mathbb{C}) \\ f & \mapsto u \end{array} \right. ,$$

qui est linéaire et continu d'après ce qui précède. Or on a montré que :

$$\Delta(\mathcal{C}^1(\mathbb{R})) \subset \mathcal{C}_{2\pi,x}^0(\mathbb{R}^+ \times \mathbb{R}, \mathbb{C}),$$

qui est un sous-espace vectoriel fermé de $L_{2\pi,x}^\infty(\mathbb{R}_t^+ \times \mathbb{R}_x, \mathbb{C})$. Puisque Δ est continu et que $\mathcal{C}_{2\pi}^1$ est dense dans $\mathcal{C}_{2\pi}^0$ pour la norme uniforme, on a :

$$\Delta(\mathcal{C}_{2\pi}^0) = \Delta(\overline{\mathcal{C}_{2\pi}^1}) \subset \overline{\Delta(\mathcal{C}_{2\pi}^1)} \subset \mathcal{C}_{2\pi,x}^0(\mathbb{R}^+ \times \mathbb{R}, \mathbb{C}).$$

Cela achève la preuve de la continuité en 0 de la solution

$\square$

Remarque 3.5.2. - *Le développement est un peu trop long, on peut aller vite sur l'analyse, notamment en passant très vite sur les théorèmes de régularité sous l'intégrale et on peut admettre le principe du maximum.*

- *C'est une belle application du théorème de Féjèr et de la convolution !!!*

3.6 Équipartition presque sûre des orbites du doublement de l'angle

Références : P. Billingsley, *Probability and measure*[9]

Leçons : 202, 223, 226, 261, 262, 264.

On se place sur $E = [0,1[$ qu'on munit de la mesure de Lebesgue. On définit l'application f par $f(x) = 2x - \lfloor 2x \rfloor := \{2x\}$, pour tout x dans E. Cette application correspond au doublement de l'angle sur le cercle.

Définition 3.6.1

Soit $(u_n)_n$ une suite de E. On dit que $(u_n)_n$ est équirépartie si pour tout $a < b$ dans E on a :

$$\frac{1}{n} Card\{0 \leq k < n, u_k \in [a,b[\} \to b - a.$$

Théorème 3.6.2

Pour presque-tout x dans E, la suite des itérées $(f^k(x))_k$ est équirépartie.

On pose, pour tout $x \in E$, $X_0(x) = \lfloor 2x \rfloor = \mathbf{1}_{x \in [1/2,1[}$ ainsi que pour tout $n \in \mathbb{N}$, $X_n(x) = X_0(f^n(x))$. On commence par démontrer un lemme.

Lemme 3.6.3

On a pour tout $x \in E$ et pour tout $n \geq 1$:

$$x \in I_n(x) = \left[x_n, x_n + \frac{1}{2^n} \right[, \quad \text{où} \quad x_n = \sum_{k=0}^{n-1} \frac{X_k(x)}{2^{k+1}}.$$

Démonstration : On raisonne par récurrence.

- Pour $n = 1$, on a par définition : $X_0(x) \leq 2x < X_0(x) + 1$ ce qui est exactement le résultat voulu.
- Supposons le résultat établi au rang n et montrons le au rang suivant. On peut montrer par récurrence que :

$$\forall x \in E,\ f^n(x) = \{2^n x\}.$$

Ainsi, puisque par hypothèse de récurrence, $2^n(x - x_n) \in [0,1[$ et puisque $2^n x_n \in \mathbb{N}$, on a :

$$\begin{aligned} X_n(x) &= \lfloor 2 f^n(x) \rfloor \\ &= \lfloor 2\{2^n x\} \rfloor \\ &= \lfloor 2\{2^n(x - x_n)\} \rfloor \\ &= \lfloor 2^{n+1}(x - x_n) \rfloor \end{aligned}$$

9. Merci à Alain pour les discussions sur ce développement qui ont permis de l'améliorer.

On a donc :

$$X_n(x) \leq 2^{n+1}(x - x_n) < X_n(x) + 1,$$

ce qui est exactement le résultat voulu.

□

Le raisonnement précédent permet de montrer qu'en fait les X_k sont uniques au sens où si on se donne $\epsilon_0, \ldots, \epsilon_{n-1} \in \{0, 1\}$ et si on pose $x_n = \sum\limits_{k=0}^{n-1} \frac{\epsilon_k}{2^{k+1}}$, alors :

$$x \in \left[x_n, x_n + \frac{1}{2^n}\right[\Leftrightarrow X_0(x) = \epsilon_0, \ldots, X_{n-1}(x) = \epsilon_{n-1}.$$

On déduit immédiatement de cette remarque la proposition suivante.

Proposition 3.6.4

La suite $(X_k)_k$ est une suite de variables aléatoires sur E muni de la mesure de Lebesgue indépendantes, identiquement distribuées de loi Bernoulli de paramètre $1/2$.

Démonstration : La mesurabilité est claire. Ce sont donc bien des variables aléatoires. On se donne $n \geq 1$ et $\epsilon_0, \ldots, \epsilon_{n-1} \in \{0, 1\}$ et on a par définition :

$$\mathbb{P}(X_0 = \epsilon_0, \ldots, X_{n-1} = \epsilon_{n-1}) = \lambda\left(\left[x_n, x_n + \frac{1}{2^n}\right[\right) = \frac{1}{2^n},$$

en gardant les mêmes notations que précédemment. On reconnait la loi produit voulue ce qui achève la preuve. □

On peut maintenant passer à la preuve du théorème.

Démonstration : Soit $a < b \in E$ et soit $x \in E$, on note :

$$S_n^x(a, b) = \frac{1}{n}\mathrm{Card}\{0 \leq k < n, f^k(x) \in [a, b[\},$$

et on note $S_n(a, b)$ la variable aléatoire associée.

On commence par démontrer le résultat dans le cas dyadique et on procédera par densité. On fixe $l \geq 1$ et on se donne toujours $\epsilon_0, \ldots, \epsilon_{l-1} \in \{0, 1\}$. On note $a = \sum\limits_{k=0}^{l-1} \epsilon_k 2^{l-k-1} = 2^l \sum\limits_{k=0}^{l-1} \frac{\epsilon_k}{2^{k+1}}$. On a successivement :

$$\begin{aligned}
S_n^x\left(\frac{a}{2^l}, \frac{a+1}{2^l}\right) &= \frac{1}{n}\mathrm{Card}\left\{0 \leq k < n,\ f^k(x) \in \left[\frac{a}{2^l}, \frac{a+1}{2^l}\right[\right\}. \\
&= \frac{1}{n}\mathrm{Card}\left\{0 \leq k < n,\ X_0(f^k(x)) = \epsilon_0, \ldots, X_{l-1}(f^k(x)) = \epsilon_{l-1}\right\} \\
&= \frac{1}{n}\mathrm{Card}\left\{0 \leq k < n,\ X_k(x) = \epsilon_0, \ldots, X_{k+l-1}(x) = \epsilon_{l-1}\right\}
\end{aligned}$$

On écrit la division euclidienne de n et k par l : $n = q_n l + r_n$ et $k = ql + r$ si bien que :

$$\begin{aligned}
S_n^x\left(\frac{a}{2^l}, \frac{a+1}{2^l}\right) &= \frac{1}{n}\sum_{r=0}^{l-1}\sum_{q=0}^{q_n-1} \mathbf{1}_{X_{ql+r}(x)=\epsilon_0}\dots\mathbf{1}_{X_{ql+r+l-1}(x)=\epsilon_{l-1}} + o(1/n)\\
&= \frac{q_n}{n}\sum_{r=0}^{l-1}\frac{1}{q_n}\sum_{q=0}^{q_n-1} \mathbf{1}_{X_{ql+r}(x)=\epsilon_0}\dots\mathbf{1}_{X_{ql+r+l-1}(x)=\epsilon_{l-1}} + o(1/n)
\end{aligned}$$

De la proposition précédente on déduit qu'à r fixé, la famille de variables aléatoires $(Y_q^r := \mathbf{1}_{X_{ql+r}(.)=\epsilon_0}\dots\mathbf{1}_{X_{ql+r+l-1}(.)=\epsilon_{l-1}})_q$ est iid de loi Bernoulli de paramètre $\frac{1}{2^l}$. On peut appliquer la loi forte des grands nombres ce qui permet de déduire que presque sûrement :

$$\frac{1}{q_n}\sum_{q=0}^{q_n-1} Y_q^r \underset{n\to+\infty}{\to} \frac{1}{2^l},$$

on a donc toujours presque sûrement :

$$S_n\left(\frac{a}{2^l}, \frac{a+1}{2^l}\right) \underset{n\to+\infty}{\to} \frac{1}{2^l}.$$

Cela permet de déduire qu'on a presque sûrement :

$$\forall l \geq 1,\, \forall a \in \{0,\dots,2^l-1\},\, \forall b \in \{a+1,\dots,2^l\},\, S_n\left(\frac{a}{2^l}, \frac{b}{2^l}\right) \underset{n\to+\infty}{\to} \frac{b-a}{2^l}.$$

On se place sur cet événement presque sûr et on se donne $a < b \in E$ on dispose (d'après le premier lemme) de deux suites d'entiers $(a_l)_l$ et $(b_l)_l$ telles que :

$$\frac{a_l}{2^l} \leq a < \frac{a_l+1}{2^l} \quad \text{et} \quad \frac{b_l}{2^l} \leq b < \frac{b_l+1}{2^l}.$$

Les suites $\left(\frac{a_l}{2^l}\right)_l$ et$\left(\frac{b_l}{2^l}\right)_l$ convergent respectivement vers a et b. On a :

$$S_n\left(\frac{a_l+1}{2^l}, \frac{b_l}{2^l}\right) \leq S_n(a,b) \leq S_n\left(\frac{a_l}{2^l}, \frac{b_l+1}{2^l}\right),$$

d'où l'on déduit que :

$$\frac{b_l - a_l - 1}{2^l} = \liminf_n S_n\left(\frac{a_l+1}{2^l}\frac{b_l}{2^l}\right) \leq \liminf_n S_n(a,b)$$

et

$$\limsup_n S_n(a,b) \leq \limsup_n S_n\left(\frac{a_l}{2^l}, \frac{b_l+1}{2^l}\right) = \frac{b_l+1-a_l}{2^l}.$$

En laissant tendre l vers $+\infty$, on obtient que presque sûrement :

$$\forall a < b \in E, S_n(a,b) \to b-a,$$

ce qui achève la preuve. $\square$

Remarque 3.6.5. - *Le développement est un peu long tel qu'il est écrit ici. On peut juste énoncer la proposition puisqu'elle découle immédiatement du lemme et aller vite sur l'utilisation de la loi forte des grands nombres ou sur la fin au choix.*

- *Il est bon d'avoir le critère d'équirépartition de Weyl en tête.*
- *Une question facile : quels sont les points périodiques de f ? Et une question à laquelle je ne sais pas répondre : peut-on exhiber un $x \in [0,1[$ dont l'orbite est équirépartie ?*
- *On a montré que presque sûrement, un nombre $x \in [0,1[$ admet dans son développement dyadique la même proportion de 0 et de 1. On peut montrer en copiant la preuve que ce résultat se généralise au développement en base r quelconque.*
- *Une suite iid de variables aléatoires de loi uniforme sur $[0,1[$ est équirépartie presque sûrement (LGN).*

Annexe

Démontrons la loi forte des grands nombres qu'on a utilisée. Il s'agit de la version L^4 dont la preuve est bien plus simple que la version L^1. Cela ne fait pas partie du développement bien entendu !

Théorème 3.6.6

Soit $(X_n)_n$ une suite de variables aléatoires indépendantes, de même moyenne m et bornée dans L^4, c'est-à-dire il existe $C > 0$ telle que :

$$\forall n \in \mathbb{N},\ \mathbb{E}(X_n^4) \leq C.$$

Alors si on note $S_n = \sum_{k=1}^{n} X_k$, on a :

$$\frac{S_n}{n} \underset{n \to +\infty}{\to} m \quad ps.$$

Démonstration : Quitte à soustraire la moyenne, on peut supposer $m = 0$. Soit $n \in \mathbb{N}$, on a :

$$\mathbb{E}(S_n^4) = \sum_{i,j,k,l \in \{1,\dots,n\}} \mathbb{E}(X_i X_j X_k X_l).$$

Les termes du type $\mathbb{E}(X_i X_j X_k X_l)$ où l'une des variables apparaît seule sont nuls puisqu'on peut scinder l'espérance en 2 parties par indépendance dont l'une est l'espérance de la variable aléatoire seule, qui est nulle par hypothèse. Au total on a :

$$\begin{aligned}
\mathbb{E}(S_n^4) &= \sum_{k=1}^{n} \mathbb{E}(X_k^4) + \binom{4}{2} \sum_{i<j} \mathbb{E}(X_i^2 X_j^2) \\
&\leq nC + 6 \sum_{i<j} \sqrt{\mathbb{E}(X_i^4)\mathbb{E}(X_j)^4} \quad (\text{ Cauchy-Schwarz}) \\
&\leq nC + 3n(n-1)C.
\end{aligned}$$

Cela assure que :

$$\mathbb{E}\left(\sum_{n=1}^{\infty} \left(\frac{S_n}{n}\right)^4\right) \underset{\text{Fubini-Tonelli}}{=} \sum_{n=1}^{\infty} \mathbb{E}\left(\frac{S_n}{n}\right)^4 < +\infty.$$

Ainsi la série converge presque sûrement puisque d'espérance finie, d'où :

$$\frac{S_n}{n} \underset{ps}{\to} 0.$$

□

3.7 Représentation des fonctions lipschitziennes

Référence : `http://dyna.maths.free.fr/docs/lecons/developpement_analyse_554.pdf`

Leçons : 201, 207, 208, 228, 234.

Théorème 3.7.1

Une application $f : \mathbb{R} \to \mathbb{R}$ est lipschitzienne si et seulement si il existe une fonction $g \in L^\infty(\mathbb{R})$ telle que :

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

Démonstration : Le sens réciproque est immédiat : montrons le sens direct. On note $\mathcal{D}(\mathbb{R}) = \mathcal{C}_c^\infty(\mathbb{R})$ et on considère la dérivée distributionnelle T de f :

$$T : \begin{cases} \mathcal{D}(\mathbb{R}) & \to \mathbb{R} \\ \phi & \mapsto \langle f', \phi \rangle := -\displaystyle\int_{\mathbb{R}} f(x)\phi'(x)\, dx \end{cases}.$$

- **Étape 1 : prolongement de T à L^1.** On commence par montrer que T est une forme linéaire continue sur $(\mathcal{D}(\mathbb{R}), \|.\|_{L^1})$. Pour cela montrons que pour $\phi \in \mathcal{D}(\mathbb{R})$:

$$\langle T, \phi \rangle = \lim_{h \to 0} \int_{\mathbb{R}} f(x) \frac{\phi(x+h) - \phi(x)}{h}\, dx.$$

Soit $M > 0$ tel que $|x| > M \Rightarrow \phi(x) = 0$. On a :

(a) $\forall x \in \mathbb{R}$, $f(x)\frac{\phi(x+h)-\phi(x)}{h} \underset{h \to 0}{\to} f(x)\phi'(x)$.

(b) Par l'inégalité des accroissements finis, on a :

$$\forall x \in \mathbb{R}, \forall |h| \leq 1, \left| f(x) \frac{\phi(x+h) - \phi(x)}{h} \right| \leq |f(x)| \|\phi'\|_\infty \mathbf{1}_{[-M-1, M+1]}(x) \in L^1(\mathbb{R}).$$

On conclut par le théorème de convergence dominée. Par linéarité de l'intégrale et changement de variables, on déduit que :

$$\forall \phi \in \mathcal{D}(\mathbb{R}), \langle T, \phi \rangle = \int_{\mathbb{R}} \frac{f(x-h) - f(x)}{h} \phi(x)\, dx.$$

Ainsi si on note L tel que f soit L-lipschitzienne, on a :

$$\forall \phi \in \mathcal{D}(\mathbb{R}), |\langle T, \phi \rangle| \leq L \|\phi\|_{L^1}.$$

Par le théorème de prolongement des applications uniformément continues sur une partie dense, T admet un unique prolongement toujours noté T, en une forme linéaire continue sur $(L^1(\mathbb{R}), \|.\|_{L^1(\mathbb{R})})$ i.e en un élément de $L^1(\mathbb{R})'$. De plus : $\|T\|_{L^1(\mathbb{R})'} \leq L$.

- **Etape 2 : on construit g grâce théorème de Riesz**
  On se donne $n \geq 1$. On utilise les injections suivantes qui sont continues :

$$L^2(-n,n) \hookrightarrow L^1(-n,n) \hookrightarrow L^1(\mathbb{R}).$$

La première injection est continue par Cauchy-Schwarz et la seconde injection associe à $\phi \in L^1(-n,n)$ sont prolongement à $L^1(\mathbb{R})$ en prolongeant ϕ par 0 en dehors de $[-n,n]$: on note $\tilde{\phi}$ ce prolongement. On peut ainsi définir la forme linéaire continue suivante :

$$T_n : \left\{ \begin{array}{cl} L^2(-n,n) & \to \mathbb{R} \\ \phi & \mapsto \langle T, \tilde{\phi} \rangle \end{array} \right. .$$

D'après le théorème de Riesz, il existe une unique $g_n \in L^2(-n,n)$ telle que :

$$\forall \phi \in L^2(-n,n),\ \langle T, \tilde{\phi} \rangle = \int_{\mathbb{R}} g_n(x)\phi(x)\,dx.$$

De plus si $k \geq n$, l'unicité du théorème de Riesz assure que $g_k = g_n$ presque partout sur $[-n,n]$. On peut donc définir $g = \liminf_n g_n$ qui est mesurable. De plus pour tout n, $g_{|[-n,n]} = g_n$.

Montrons que g appartient à $L^\infty(\mathbb{R})$. Par l'absurde, supposons que $A = \{x \in \mathbb{R}, |g(x)| > L\}$ n'est pas négligeable. Puisqu'on a l'union croissante suivante :

$$A = \bigcup_n A_n,$$

où $A_n = A \cap [-n,n]$, il existe N tel que $\lambda(A_N) > 0$. Posons :

$$u = \mathbf{1}_{\mathbb{R}^+}(g) - \mathbf{1}_{\mathbb{R}_*^-}(g),$$

si bien que $|g| = ug$. On considère alors $\phi = u\mathbf{1}_{A_N} \in L^2(-n,n)$. On a alors :

$$\langle T, \phi \rangle = \int_{-N}^{N} g(x)u(x)\mathbf{1}_{A_N}(x)\,dx = \int_{A_N} |g(x)|\,dx > L\lambda(A_N) = L\|\phi\|_{L^1}.$$

Cela contredit $\|T\|_{L^1(\mathbb{R})'} \leq L$. On a donc montré que $T = f' = g$ est dans L^∞.

- **Étape 3 : conclusion** On définit :

$$G : \left\{ \begin{array}{cl} \mathbb{R} & \to \mathbb{R} \\ x & \mapsto \displaystyle\int_0^x g(t)\,dt \end{array} \right. ,$$

qui est continue par convergence dominée. Le théorème de Fubini permet de montrer que $G' = g$ au sens des distributions. En effet : soit $\phi \in \mathcal{D}(\mathbb{R})$. On a :

$$\langle G', \phi \rangle = -\int_{\mathbb{R}} G(x)\phi'(x)\,dx = \int_{-\infty}^{0} \left(\int_x^0 g(t)\,dt \right) \phi'(x)\,dx - \int_0^\infty \left(\int_0^x g(t)\,dt \right) \phi'(x)\,dx.$$

Soit M tel que $\phi(x) = 0$ si $|x| > M$. Pour presque-tout (t,x), on a :

$$|g(t)\phi'(x)|\mathbf{1}_{x \leq t \leq 0} \leq \|g\|_{L^\infty}\|\phi'\|_{L^\infty}\mathbf{1}_{-M \leq x \leq t \leq 0} \in L^1(\mathbb{R}^- \times \mathbb{R}^-).$$

Cela permet d'appliquer le théorème de Fubini au premier terme. On peut également l'appliquer au second terme ce qui assure que :

$$\begin{aligned}
\langle G', \phi \rangle &= - \int_{\mathbb{R}} G(x)\phi'(x)\,dx \\
&= \int_{-\infty}^{0} \int_{-\infty}^{t} g(t)\phi'(x)\,dx\,dt - \int_{0}^{\infty} \int_{t}^{\infty} g(t)\phi'(x)\,dx\,dt \\
&= \int_{\mathbb{R}} g(t)\phi(t)\,dt \\
&= \langle T, \phi \rangle.
\end{aligned}$$

On a donc $G' = f'$. On déduit qu'il existe une constante $C \in \mathbb{R}$ telle que : $f - G = C$ au sens des distributions. L'injection de $L^1_{\text{loc}}(\mathbb{R})$ dans $\mathcal{D}'(\mathbb{R})$ assure que $f(x) - G(x) = C$ pp-x. Comme les applications sont continues, on a le résultat pour tout $x \in \mathbb{R}$ ce qui achève la preuve.

□

Remarque 3.7.2. - *On peut passer la preuve du fait que $G' = g$ pour gagner du temps.*

- *Le raisonnement de l'étape 2 permet de montrer que $L^1(\mathbb{R})' = L^\infty(\mathbb{R})$.*
- *Le théorème de différentiation de Lebesgue assure alors qu'une fonction lipschitzienne est dérivable presque-partout.*

Annexe

Démontrons les résultats qu'on a utilisés sur les distributions.

Théorème 3.7.3

Soit $f \in L^1_{loc}(\mathbb{R})$ nulle au sens des distributions. Alors $f = 0$ presque-partout.

Démonstration : Soit $n > 0$. On dispose de $\chi_n \in \mathcal{D}(\mathbb{R})$ valant 1 sur $[-n, n]$. Il existe également une approximation de l'unité $(\rho_n)_n$. Puisque $f\chi_n \in L^1(\mathbb{R})$, on a (voir un livre d'intégration par exemple le livre de M. Briane et G. Pagès : *Théorie de l'intégration*) :

$$f\chi_n * \rho_k \underset{k\infty, L^1}{\to} f\chi_n. \quad (*)$$

Or si on fixe $x \in \mathbb{R}$, on a :

$$f\chi_n * \rho_k(x) = \int_{\mathbb{R}} f(t)\chi_n(t)\rho_k(x-t)\, dt.$$

Puisque $t \mapsto f(t)\chi_n(t)\rho_k(x-t) \in \mathcal{D}(\mathbb{R})$, on déduit que pour tout k :

$$f\chi_n * \rho_k = 0.$$

La convergence $(*)$ assure alors que f est nulle presque partout sur $[-n, n]$. Cela étant vrai pour tout n, on a prouvé ce qu'on voulait. □

Théorème 3.7.4

Soit $T \in \mathcal{D}'(\mathbb{R})$ une distribution de dérivée distributionnelle nulle. Alors T est constante.

Démonstration : On dispose d'une application $\chi \in \mathcal{D}(\mathbb{R})$ telle que :

$$\int_{\mathbb{R}} \chi(x)\, dx = 1.$$

Soit $\phi \in \mathcal{D}(\mathbb{R})$, on considère alors :

$$\psi := \phi - \left(\int_{\mathbb{R}} \phi(x)\, dx \right) \chi.$$

ψ est une fonction test d'intégrale nulle, on en déduit que la fonction : $x \in \mathbb{R} \mapsto \xi(x) := \displaystyle\int_{-\infty}^{x} \psi(t)\, dt$ est dans $\mathcal{D}(\mathbb{R})$. On a par hypothèse :

$$\begin{aligned} \langle T', \xi \rangle_{\mathcal{D}'} &= -\langle T, \xi' \rangle_{\mathcal{D}'} \\ &= -\langle T, \psi \rangle_{\mathcal{D}'} \\ &= 0. \end{aligned}$$

Cela assure que :

$$\langle T, \phi \rangle_{\mathcal{D}'} = \langle T, \chi \rangle_{\mathcal{D}'} \int_{\mathbb{R}} \phi(x)\, dx,$$

ce qui montre bien que T est une distribution constante. □

3.8 Convergence presque-sûre des sous-martingales bornées dans L^1

Références : P. Billingsley, *Probability and measure* – R. Durrett, *Probability : theory and examples*[10]

Leçons : 223, 262, *260, 241*

Commençons par quelques rappels qu'on peut mettre dans le plan mais qu'on ne fait pas à l'oral. On se donne $(X_n)_n$ une suite de variables aléatoires et une filtration $(\mathcal{F}_n)_n$.

Définition 3.8.1

On dit que $(X_n)_n$ est une sous-martingale si :

- *Pour tout n, X_n est $\mathcal{F}_n$-mesurable*
- *Pour tout n, $X_n \in L^1$.*
- *Pour tout n, $\mathbb{E}(X_{n+1}|X_n) \geq X_n$*

L'objet du développement est le théorème de convergence suivant.

Théorème 3.8.2

Soit $(X_n)_n$ une sous-martingale bornée dans L^1, c'est-à-dire :

$$\sup_n \mathbb{E}|X_n| < +\infty.$$

Alors il existe une variable aléatoire $X \in L^1$ telle que $(X_n)_n$ converge presque sûrement vers X.

La preuve repose sur le premier théorème d'arrêt qu'on admettra à l'oral faute de temps.

Définition 3.8.3

Une variable aléatoire T à valeurs dans $\mathbb{N} \cup \{+\infty\}$ est un temps d'arrêt relativement à la filtration $(\mathcal{F}_n)_n$ si :

$$\forall n \in \mathbb{N} \cup \{+\infty\}, (T \leq n) \in \mathcal{F}_n,$$

où F_∞ est la tribu engendrée par les $\mathcal{F}_n$.

On peut remplacer dans la définition $(T \leq n)$ par $(T = n)$ sans la changer.

Théorème 3.8.4 (*Premier théorème d'arrêt*)

Soit $(X_n)_n$ une sous-martingale. Soient S et T deux temps d'arrêt bornés par N, avec $S \leq T$ ps. Alors on a :

$$\mathbb{E}(X_S) \leq \mathbb{E}(X_T),$$

où $X_S = \sum_{k=0}^{N} X_k \mathbf{1}_{S=k}$

10. Je me suis également largement inspiré du cours suivi en M1 enseigné par M. Gradinaru.

Démonstration : Soit $k \geq 1$. On définit les accroissements de la sous-martingale par :

$$\Delta_k = X_k - X_{k-1}.$$

Remarquons tout d'abord que :

$$X_T = X_0 + \sum_{k=1}^{N} \Delta_k \mathbf{1}_{k \leq T}.$$

On a alors :

$$\begin{aligned} X_T - X_S &= \sum_{k=1}^{N} \Delta_k (\mathbf{1}_{k \leq T} - \mathbf{1}_{k \leq S}) \\ &= \sum_{k=1}^{N} \Delta_k \mathbf{1}_{S < k \leq T}. \end{aligned}$$

Ainsi :

$$\begin{aligned} \mathbb{E}(X_T - X_S) &= \sum_{k=1}^{N} \mathbb{E}(\Delta_k \mathbf{1}_{S<k\leq T}) \\ &= \sum_{k=1}^{N} \mathbb{E}(\mathbb{E}(\Delta_k \mathbf{1}_{S<k\leq T} | \mathcal{F}_{k-1})). \end{aligned}$$

Or on a : $(S < k) \in \mathcal{F}_{k-1}$ par définition d'un temps d'arrêt et $(k \leq T) = (T \leq k-1)^C \in \mathcal{F}_{k-1}$.

On a alors par propriété de mesurabilité de l'espérance conditionnelle :

$$\begin{aligned} \mathbb{E}(X_T - X_S) &= \sum_{k=1}^{N} \mathbb{E}(\mathbb{E}(\Delta_k \mathbf{1}_{S<k\leq T} | \mathcal{F}_{k-1})) \\ &= \sum_{k=1}^{N} \mathbb{E}(\mathbf{1}_{S<k\leq T} \mathbb{E}(\Delta_k | \mathcal{F}_{k-1})) \\ &\geq 0 \quad \text{puisque pour tout } k,\ \mathbb{E}(\Delta_k | \mathcal{F}_{k-1}) = \mathbb{E}(X_k | \mathcal{F}_{k-1}) - X_{k-1} \geq 0 \ . \end{aligned}$$

□

Avant de passer à la preuve du théorème, démontrons un lemme qui nous sera utile (c'est là que commence le développement).

Définition 3.8.5

Soit $(u_n)_n$ une suite réelle. On se donne $a < b \in \mathbb{R}$ et on définit alors $\tau_1 := \inf\{k \geq 1, u_k \leq a\}$, en convenant que la borne inférieure de l'ensemble vide est $+\infty$. On définit ensuite par récurrence :

$$\tau_{2n} := \inf\{k > \tau_{2n-1}, u_k \geq b\} \quad \textit{et} \quad \tau_{2n+1} := \inf\{k > \tau_{2n}, u_k \leq a\}.$$

On peut alors définir le nombre de traversées montantes de l'intervalle $[a, b]$ par :

$$U_\infty(a, b) := \sum_{k=1}^{\infty} \mathbf{1}_{\tau_{2k}<+\infty},$$

ainsi que le nombre de traversées avant l'instant n par :

$$U_n(a, b) := \sum_{k=1}^{\infty} \mathbf{1}_{\tau_{2k}\leq n},$$

τ_{2k} est l'instant de la k-ème traversée montante.

Lemme 3.8.6

Une suite $(u_n)_n$ converge dans $\bar{\mathbb{R}}$ si et seulement si pour tout $a < b$ rationnels, $U_\infty(a, b) < +\infty$.

Démonstration : Notons $l = \liminf u_n$ et $u = \limsup u_n$.
($\Rightarrow$) Supposons que la suite $(u_n)_n$ converge, i.e $l = u$. Soient $a < b$ deux rationnels. Il y a deux cas :

- Si $a < l$, alors il existe N tel que $n \geq N \Rightarrow u_n > a$. On a donc : $U_\infty(a, b) \leq N$.
- Si $b > a \geq u = l$, alors il existe N tel que $n \geq N \Rightarrow u_n < b$. On a donc : $U_\infty(a, b) \leq N$.

($\Leftarrow$) Par contraposée : supposons que $l < u$. On se donne deux rationnels a et b tels que : $l < a < b < u$. On construit par récurrence une infinité de traversées montantes de $[a, b]$ puisque l et u sont respectivement les plus petite et plus grande valeur d'adhérence de la suite et donc la suite passe une infinité de fois en dessous de a et au dessus de b. □

On peut remplacer la suite $(u_n)_n$ par une suite de variables aléatoires $(X_n)_n$. Les τ_k sont des temps d'arrêt pour la filtration canonique associée à cette suite $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$. De plus le nombre de traversées montantes tronqué ou non est une variable aléatoire. La preuve repose sur l'inégalité des traversées montantes de Doob.

Lemme 3.8.7

Soient $a < b \in \mathbb{R}$, $n \in \mathbb{N}$. On a alors :

$$(b - a)\mathbb{E}(U_n(a, b)) \leq \mathbb{E}|X_n - a|.$$

Démonstration : On fixe $n \in \mathbb{N}$ et $k \geq 1$. On commence par appliquer le premier théorème d'arrêt aux temps d'arrêt bornés $\tau_{2k} \wedge n$ et $\tau_{2k+1} \wedge n$, où on note $\wedge$ le minimum. On a successivement :

$$
\begin{aligned}
0 &\leq \mathbb{E}(X_{\tau_{2k+1}\wedge n} - X_{\tau_{2k}\wedge n}) \\
&= \mathbb{E}((X_{\tau_{2k+1}\wedge n} - X_{\tau_{2k}\wedge n})\mathbf{1}_{\tau_{2k}\leq n}) \\
&= \mathbb{E}((X_{\tau_{2k+1}\wedge n} - X_{\tau_{2k}\wedge n})\mathbf{1}_{\tau_{2k}\leq n<\tau_{2k+1}}) + \mathbb{E}((X_{\tau_{2k+1}\wedge n} - X_{\tau_{2k}\wedge n})\mathbf{1}_{\tau_{2k+1}\leq n}) \\
&= \mathbb{E}((X_n - X_{\tau_{2k}})\mathbf{1}_{\tau_{2k}\leq n<\tau_{2k+1}}) + \mathbb{E}((X_{\tau_{2k+1}} - X_{\tau_{2k}})\mathbf{1}_{\tau_{2k+1}\leq n}) \\
&\leq \mathbb{E}((X_n - b)\mathbf{1}_{\tau_{2k}\leq n<\tau_{2k+1}}) + \mathbb{E}((X_{\tau_{2k+1}} - X_{\tau_{2k}})\mathbf{1}_{\tau_{2k+1}\leq n}) \quad \text{par définition de } \tau_{2k} \\
&\leq \mathbb{E}((X_n - b)\mathbf{1}_{\tau_{2k}\leq n<\tau_{2k+1}}) + (a-b)\mathbb{P}(\tau_{2k+1}\leq n) \quad \text{par définition des } \tau_k \\
&= \mathbb{E}((X_n - a + a - b)\mathbf{1}_{\tau_{2k}\leq n<\tau_{2k+1}}) + (a-b)\mathbb{P}(\tau_{2k+1}\leq n) \\
&= \mathbb{E}((X_n - a)\mathbf{1}_{\tau_{2k}\leq n<\tau_{2k+1}}) + (a-b)(\mathbb{P}(\tau_{2k+1}\leq n) + \mathbb{P}(\tau_{2k}\leq n<\tau_{2k+1})) \\
&= \mathbb{E}((X_n - a)\mathbf{1}_{\tau_{2k}\leq n<\tau_{2k+1}}) + (a-b)\mathbb{P}(\tau_{2k}\leq n)
\end{aligned}
$$

Or les événements suivants sont égaux :

$$(\tau_{2k} \leq n) = (U_n(a,b) \geq k).$$

On a donc :

$$0 \leq \mathbb{E}(|X_n - a|\mathbf{1}_{\tau_{2k}\leq n<\tau_{2k+1}}) + (a-b)\mathbb{P}(U_n(a,b)\geq k)$$

En sommant ces inégalités pour $k \geq 1$, on obtient :

$$
\begin{aligned}
(b-a)\sum_{k\geq 1}\mathbb{P}(U_n(a,b)\geq k) = (b-a)\mathbb{E}(U_n(a,b)) &\leq \mathbb{E}(|X_n - a|\sum_{k\geq 1}\mathbf{1}_{\tau_{2k}\leq n<\tau_{2k+1}}) \\
&\leq \mathbb{E}(|X_n - a|).
\end{aligned}
$$

(La première égalité étant un résultat classique qui découle de Fubini-Tonelli ou d'un argument de famille sommable si on préfère :

$$
\begin{aligned}
\sum_{k\geq 1}\mathbb{P}(U_n(a,b)\geq k) &= \sum_{k\geq 1}\sum_{j\geq k}\mathbb{P}(U_n(a,b)=j) \\
&= \sum_{j\geq 1}\sum_{1\leq k\leq j}\mathbb{P}(U_n(a,b)=j) \\
&= \sum_{j\geq 1} j\mathbb{P}(U_n(a,b)=j) \\
&= \mathbb{E}(U_n(a,b)). \quad)
\end{aligned}
$$

□

On peut enfin passer à la preuve du théorème de convergence.

Démonstration : Soient $a < b$ des rationnels. On a : $\sup_n \mathbb{E}|X_n - a| < +\infty$ par hypothèse. De plus :

$$U_n(a,b) \underset{n\to+\infty}{\to} U_\infty(a,b).$$

En utilisant le théorème de convergence monotone et le second lemme, on déduit que la variable aléatoire $U_\infty(a,b)$ est intégrable donc finie presque sûrement. Comme une intersection dénombrable d'événements presque sûrs est presque sûre, on a la convergence presque sûre de $(X_n)_n$ d'après le premier lemme. On note X la limite presque sûre. X est intégrable par le lemme de Fatou, ce qui achève la preuve. $\square$

Remarque 3.8.8. - *Ne pas hésiter à faire un dessin pour définir les temps τ_k, le jury appréciera presque sûrement !*

Annexe

- Une version L^1 du théorème ?

Commençons par démontrer un théorème qui renforce le précédent avec l'hypothèse d'uniforme intégrabilité (rappels faits dans l'annexe du développement sur le problème des moments).

Corollaire 3.8.9

Soit $(X_n)_n$ une sous-martingale. $(X_n)_n$ est uniformément intégrable si et seulement si elle converge presque sûrement et dans L^1 vers une variable aléatoire X.

Démonstration :

- Si $(X_n)_n$ est uniformément intégrable, alors elle est bornée dans L^1, d'après le théorème, elle converge presque sûrement vers une variable aléatoire X. Un résultat classique (redémontré dans le développement sur le problème des moments) assure qu'il y a en fait convergence L^1 grâce à l'uniforme intégrabilité.
- Si $(X_n)_n$ converge presque sûrement et dans L^1 vers X. On vérifie facilement grâce à Fubini-Tonelli que si $M > 0$:

$$\forall n, \quad \mathbb{E}(|X_n|\mathbf{1}_{|X_n|\geq M}) = \mathbb{E}(|X_n|) - \int_0^M \mathbb{P}(t < X < M)\, dt.$$

  On déduit donc que :

$$\mathbb{E}(|X_n|\mathbf{1}_{|X_n|\geq M}) \to \mathbb{E}(|X|\mathbf{1}_{|X|\geq M}),$$

  d'après les hypothèses et le théorème de convergence dominée. Ainsi, si $\epsilon > 0$, il existe $M > 0$ tel que :

$$\mathbb{E}(|X|\mathbf{1}_{|X|\geq M}) \leq \epsilon.$$

  Il existe $N > 0$ tel que :

$$\forall n \geq N,\ \mathbb{E}(|X_n|\mathbf{1}_{|X_n|\geq M}) \leq 2\epsilon.$$

  Puisque $X_0, \ldots X_{N-1}$ sont intégrables, on en déduit l'uniforme intégrabilité de la suite.

$\square$

Une application.

On considère le modèle génétique de Wright-Fisher (on pourra consulter le livre de D. Chafaï et F. Malrieu intitulé *Recueil de modèles aléatoires* pour avoir des détails sur le modèle). On fixe un entier $N > 0$ et on considère une chaîne de Markov d'espace d'états $\{0, \ldots N\}$, partant de k et vérifiant :

$$\mathcal{L}(X_{n+1}|X_n) = \mathcal{B}(N, X_n/N).$$

On note $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$. La suite $(X_n)_n$ est clairement une martingale puisque :

$$\begin{aligned}\mathbb{E}(X_{n+1}|\mathcal{F}_n) &= \mathbb{E}(X_{n+1}|X_n) \\ &= X_n \frac{X_n}{N} \\ &= X_n.\end{aligned}$$

Cette martingale est bornée par N. En vertu de ce qui précède elle converge presque sûrement et dans L^1 (car une suite bornée est uniformément intégrable) vers une variable aléatoire intégrable X. Cette variable X ne peut prendre ses valeurs que dans $\{0, N\}$ qui sont les seuls états absorbants (les autres états sont transients car ils conduisent à un état absorbant). Pour déterminer la loi de X on utilise la convergence L^1. On note $p = \mathbb{P}(X = N)$. Puisqu'on a une martingale, l'espérance de X_n est constante égale à $k = X_0$. La convergence $\mathbb{E}(X_n) \to \mathbb{E}(X) = pN$ assure que :

$$p = \frac{k}{N}.$$

3.9 Marche aléatoire sur $\mathbb{Z}^d$

Références : H. Dym et H.P. McKean, *Fourier Series and Integrals*[11]

Leçons : 260, 262, 264

Commençons par introduire quelques notations. On se donne $d \geq 1$ un entier et on note $(e_1, \ldots, e_d)$ la base canonique de $\mathbb{R}^d$. Soit $(X_i)_i$ une suite de variables aléatoire iid de loi uniforme sur $\{\pm e_1, \ldots, \pm e_d\}$. On définit alors la marche aléatoire isotrope par :

$$\begin{cases} S_0 & = 0 \\ S_n & = \sum\limits_{k=1}^{n} X_k \end{cases}.$$

Il s'agit d'une chaîne de Markov homogène. On s'intéresse à la classification des états de cette chaîne.

Définition 3.9.1

Soit $k \in \mathbb{Z}^d$. On définit le nombre de visites d'un état par :

$$N_k = \sum_{n=0}^{\infty} \mathbf{1}_{S_n = k}.$$

Théorème 3.9.2

On a la dichotomie suivante :

- *Si $d \geq 3$, alors :*
$$\mathbb{E}(N_0) < +\infty.$$
De plus :
$$|S_n| \to +\infty \quad ps.$$
On dit que la marche aléatoire est transiente.
- *Si $d \in \{1, 2\}$, alors :*
$$\mathbb{E}(N_0) = +\infty.$$
De plus, $N_0 = +\infty$ ps. On dit que l'état 0 (ou la marche aléatoire puisqu'elle est irréductible et donc tous les états le seront) est récurrent.

Démonstration :

- **Étape 1 :** D'après Fubini-Tonelli, on a :

$$\mathbb{E}(N_0) = \sum_{n=0}^{\infty} \mathbb{P}(S_n = 0) = \sum_{n=0}^{\infty} \mathbb{P}(S_{2n} = 0),$$

puisqu'on ne peut revenir en 0 qu'en un nombre pair de pas ($|S_n|_1 = n[2]$).

11. Également inspiré de la version de Corentin Kilque.

- **Étape 2 :** On note ϕ la fonction caractéristique de X_1 et ϕ_{S_n} la fonction caractéristique de S_n. Remarquons déjà que par hypothèse, on a :

$$\forall n,\ \phi_{S_n} = \phi^n.$$

On note également :

$$\forall x = (x_1, \dots, x_d) \in \mathbb{R}^d, \quad |x|_1 := |x_1| + \dots + |x_d|.$$

Remarquons que pour tout $t \in \mathbb{R}$:

$$\phi_{S_n}(t) = \sum_{k \in \mathbb{Z}^d, |k|_1 \leq n} \mathbb{P}(S_n = k) e^{it.k},$$

la linéarité de l'intégrale assure alors que :

$$\begin{aligned}\frac{1}{(2\pi)^d} \int\limits_{[-\pi,\pi]^d} \phi_{S_n}(t)\, dt &= \sum_{k \in \mathbb{Z}^d, |k|_1 \leq n} \mathbb{P}(S_n = k) \frac{1}{(2\pi)^d} \int\limits_{[-\pi,\pi]^d} e^{it.k}\, dt \\ &= \mathbb{P}(S_n = 0)\,, \text{par un simple calcul de l'intégrale en utilisant Fubini.}\end{aligned}$$

- **Étape 3 :** On donne maintenant une expression explicite de la fonction caractéristique ϕ dans notre cas. Pour tout $t \in \mathbb{R}$:

$$\phi(t) = \frac{1}{2d} \sum_{k=1}^{d} (e^{it.e_k} + e^{-it.e_k}) = \frac{1}{d} \sum_{k=1}^{d} cos(t_k) \in [-1, 1].$$

- **Étape 4 :** On peut maintenant donner une expression intégrale de la quantité $\mathbb{E}(N_0)$ qui nous intéresse pour la classification de l'état 0 :

$$\begin{aligned}\mathbb{E}(N_0) &= \sum_{n=0}^{\infty} \mathbb{P}(S_{2n} = 0) \\ &= \sum_{n=0}^{\infty} \frac{1}{(2\pi)^d} \int\limits_{[-\pi,\pi]^d} \phi^{2n}(t)\, dt \\ &= \frac{1}{(2\pi)^d} \int\limits_{[-\pi,\pi]^d} \frac{1}{1 - \phi^2(t)}\, dt,\end{aligned}$$

d'après Fubini-Tonelli.

- **Étape 5 :** Étudions maintenant l'intégrale. La fonction $t \mapsto \frac{1}{1-\phi^2(t)}$ est continue sur $[-\pi, \pi]^d \backslash \{(0, \dots, 0)\} \cup \{(\pm\pi, \dots, \pm\pi)\}$, dont y est localement intégrable. L'étude de l'intégrabilité en un point de l'ensemble $\{(\pm\pi, \dots, \pm\pi)\}$ se ramène à celle en 0 par périodicité. Or on a le développement limité suivant :

$$\begin{aligned}\phi(t) &\underset{t \to 0}{=} \frac{1}{d} \sum_{k=1}^{d} (1 - \frac{t_k^2}{2} + o(|t_k|^2)) \\ &\underset{t \to 0}{=} 1 - \frac{\|t\|_2^2}{2d} + o(\|t\|^2)\end{aligned}$$

On déduit alors que :

$$\frac{1}{1-\phi^2(t)} \underset{t\to 0}{=} \frac{d}{\|t\|^2} + o\left(\frac{1}{\|t\|^2}\right).$$

Or par intégration des fonctions radiales (voir à la fin du développement pour le détail), la fonction $t \in \mathbb{R}^d \backslash \{0\} \mapsto \frac{1}{\|t\|_2^2}$ est intégrable en 0 si et seulement si $r \in \mathbb{R}^+ \mapsto \frac{1}{r^2} r^{d-1}$ est intégrable en 0 donc si et seulement si $d \geq 3$. On déduit que :

$$\mathbb{E}(N_0) < +\infty \Leftrightarrow d \geq 3.$$

- **Étape 6 :** On conclut dans le cas $d \geq 3$. Montrons que pour tout $k \in \mathbb{Z}^d$, $\mathbb{E}(N_k) < +\infty$, sachant qu'on a le résultat pour $k = 0$ d'après ce qui précède. Soit $k \in \mathbb{Z}^d$ et $l = |k|_1$. En considérant un cas particulier (faire des pas dans chaque direction successivement par exemple), on a :

$$\mathbb{P}(S_l = -k) \geq \frac{1}{(2d)^l}.$$

On a alors pour tout $n > l$:

$$\begin{aligned}
\mathbb{P}(S_n = 0) &\geq \mathbb{P}(S_n = 0, S_l = -k) \\
&= \mathbb{P}(S_l = -k, S_n - S_l = k) \\
&= \mathbb{P}\left(S_l = -k, \sum_{j=l+1}^{n} X_j = k\right) \\
&= \mathbb{P}(S_l = -k)\mathbb{P}(S_{n-l} = k) \quad \text{(caractère iid des } X_i) \\
&\geq \frac{1}{(2d)^l}\mathbb{P}(S_{n-l} = k)
\end{aligned}$$

Il ne reste qu'à sommer sur n pour avoir le résultat. On déduit donc que presque sûrement :

$$\forall k \in \mathbb{Z}^d,\ N_k < +\infty.$$

On a successivement :

$$\begin{aligned}
\mathbb{P}(|S_n| \to +\infty) &= \mathbb{P}\left(\bigcap_{A\in\mathbb{N}} \bigcup_{n_0\in\mathbb{N}} \bigcap_{n\geq n_0} \{|S_n|_1 \geq A\}\right) \\
&= \lim_{A\to+\infty} \mathbb{P}\left(\bigcup_{n_0\in\mathbb{N}} \bigcap_{n\geq n_0} \{|S_n|_1 \geq A\}\right), \quad \text{par continuité décroissante de la probabilité} \\
&= \lim_{A\to+\infty} \mathbb{P}\left(\liminf_n \{|S_n|_1 \geq A\}\right) \\
&= \lim_{A\to+\infty} \mathbb{P}\left(\bigcap_{|k|_1 < A} \{N_k < +\infty\}\right) \\
&= 1, \quad \text{comme intersection d'événements presque-sûrs.}
\end{aligned}$$

- **Étape 7 :** On conclut dans le cas $d \in \{1,2\}$, i.e $\mathbb{E}(N_0) = +\infty$. On note T le dernier temps d'atteinte de 0 :

$$T = \sup\{n \geq 0,\, S_n = 0\}.$$

T est bien défini car l'ensemble en question contient 0 donc est non vide. On a alors :

$$\begin{aligned}
\mathbb{P}(T < +\infty) &= \sum_{n=0}^{\infty} \mathbb{P}(T = n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(S_n = 0,\, \forall k > n,\, S_k \neq 0) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(S_n = 0,\, \forall k > n,\, S_k - S_n \neq 0) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(S_n = 0)\mathbb{P}(T = 0).
\end{aligned}$$

Puisque par hypothèse :

$$\sum_{n=0}^{\infty} \mathbb{P}(S_n = 0) = +\infty,$$

on a nécessairement $\mathbb{P}(T = 0) = 0$, donc : $\mathbb{P}(T < +\infty) = 0$. Or on a l'égalité des événements suivants :

$$(T = +\infty) = \limsup_n \{S_n = 0\} = (N_0 = +\infty),$$

on a bien $N_0 = +\infty$ presque sûrement.

$\square$

Annexe

-Preuve du résultat utilisé d'intégration des fonctions radiales

On pourra trouver dans le livre de O. Garet et A. Kurtzmann : *De l'intégration aux probabilités.*

Proposition 3.9.3

Soit $\phi : \mathbb{R} \to \mathbb{R}$ une fonction mesurable et $\|.\|$ une norme sur $\mathbb{R}^d$. Alors la fonction $\phi \circ \|.\|$ est intégrable si et seulement si la fonction $t \in \mathbb{R}^+ \mapsto t^{d-1}\phi(t)$ est intégrable. De plus on a l'égalité suivante :

$$\int_{\mathbb{R}^d} \phi(\|x\|)\, dx = V \int_{\mathbb{R}^+} \phi(t)\, d\, t^{d-1}\, dt,$$

Où on note V le volume de la boule unité pour la norme $\|.\|$.

Démonstration : On note λ_d la mesure de Lebesgue sur $\mathbb{R}^d$. D'après le lemme de transfert, la fonction $\phi \circ \|.\|$ est intégrable par rapport à λ_d si et seulement si ϕ est intégrable par rapport à la mesure image m de λ_d par $\|.\|$. On aura alors dans ce cas :

$$\int_{\mathbb{R}^d} \phi(\|x\|)\, dx = \int_{\mathbb{R}^+} \phi(t)\, dm(t).$$

Il nous reste à identifier m. Soit $a > 0$, on a par définition :

$$\begin{aligned} m([0,a]) &= \lambda_d(B(0,a)) \\ &= \lambda_d(a\, B(0,1)) \\ &= a^d \lambda_d(B(0,1)) \quad \text{(par propriété d'échelle de la mesure de Lebesgue)} \\ &= a^d V \\ &= \int_{[0,a]} V dt^{d-1}\, dt \end{aligned}$$

Puisque les intervalles de la forme $[0,a]$ engendrent la tribu borélienne $\mathcal{B}(\mathbb{R}^+)$ et forment une famille stable par intersections finies, on déduit du lemme d'identification des mesures (conséquences des résultats sur les classes monotones) que m est la mesure à densité $t \in \mathbb{R}^+ \mapsto V dt^{d-1}$: le résultat suit du lemme de transfert. □

-Quelques résultats sur la classification des états d'une chaîne de Markov.

Cela peut être utile pour avoir un peu de recul sur ce qu'on manipule mais je ne pense pas que ce soit nécessaire pour faire le développement ! On peut parler de tout cela dans les leçons 230, 241, 243 (quand on aime les probas, on ne compte pas !)

Définition 3.9.4

Soit $(X_n)_n$ une chaîne de Markov homogène d'espace d'états E dénombrable et soit $x \in E$.

- *x est dit récurrent si $\mathbb{P}_x(\tau_x < +\infty) = 1$, où $\mathbb{P}_x$ est une probabilité sous laquelle notre* *chaîne de Markov part de l'état x et :*

$$\tau_x = \inf\{n \geq 1,\ X_n = x\}.$$

- *x est dit transient si $\mathbb{P}_x(\tau_x < +\infty) < 1$.*

Définition 3.9.5

Soient $s \in [0,1]$, $x, y \in E$. On définit :

- $U(x,y,s) = \sum_{n=1}^{+\infty} \mathbb{P}_x(\tau_y = n)s^n$ *et donc* $U(x,y) = \mathbb{P}_x(\tau_y < +\infty)$.
- $G(x,y,s) = \sum_{n=0}^{+\infty} \mathbb{P}_x(X_n = y)s^n$ *et donc* $G(x,y) = \mathbb{E}_x(N_y)$ *où* $N_y = \sum_{n=0}^{\infty} \mathbf{1}_{X_n = y}$.

x est donc récurrent si et seulement si $U(x,x) = 1$.

Lemme 3.9.6

Soient $s \in [0,1[$ et $x \in E$. On a :

$$G(x,x,s) = 1 + G(x,x,s)U(x,x,s).$$

Démonstration : On a successivement :

$$\begin{aligned}
G(x,x,s) &= \sum_{n=0}^{+\infty} \mathbb{P}_x(X_n = x)s^n \\
&= 1 + \sum_{n=1}^{+\infty} \mathbb{P}_x(X_n = x)s^n \\
&= 1 + \sum_{n=1}^{+\infty} \left(\sum_{k=0}^{n} \mathbb{P}_x(X_n = x\,,\ \tau_x = k) \right) s^n \\
&= 1 + \sum_{n=1}^{+\infty} \left(\sum_{k=0}^{n} \mathbb{P}_x(X_1, \ldots, X_{k-1} \neq x,\ X_k = x,\ X_n = x) \right) s^n \\
&= 1 + \sum_{n=1}^{+\infty} \left(\sum_{k=0}^{n} \mathbb{P}_x(X_1, \ldots, X_{k-1} \neq x,\ X_k = x)\mathbb{P}_x(X_n = x | X_k = x) \right) s^n \quad \text{(prop. de Markov)} \\
&= 1 + \sum_{n=1}^{+\infty} \left(\sum_{k=0}^{n} \mathbb{P}_x(\tau_x = k)\mathbb{P}_x(X_{n-k} = x) \right) s^n \\
&= 1 + G(x,x,s)U(x,x,s) \quad \text{(par produit de Cauchy).}
\end{aligned}$$

$\square$

Corollaire 3.9.7

x est récurrent si et seulement si $G(x,x)=+\infty$.

Démonstration : Soit $s \in [0,1[$, on a :

$$G(x,x,s)=\frac{1}{1-U(x,x,s)}.$$

En laissant tendre s vers 1 (par convergence monotone), on a l'équivalence voulue. □

Remarquons qu'on a étudié la fonction G dans le développement.

Définition 3.9.8

On définit :

- $H(x,y)=\mathbb{P}_x(\limsup(X_n=y))=\mathbb{P}_x(X_n=y\,\textit{pour une infinité de } n)=\mathbb{P}_x(N_y=+\infty)$.
- $H^{(m)}(x,y)=\mathbb{P}_x(X_n=y\,\textit{pour au moins m entiers } n\geq 1)$

Énonçons enfin la loi de type $0-1$, qu'on a redémontrée dans le développement.

Proposition 3.9.9

- *x est récurrent si et seulement si $H(x,x)=1$*
- *x est transient si et seulement si $H(x,x)=0$.*

Lemme 3.9.10

Soient $m\geq 1$, $x,y\in E$, alors :

$$H^{(m+1)}(x,y)=U(x,y)H^{(m)}(y,y).$$

Démonstration : On a :

$$\begin{aligned}
H^{(m+1)}(x,y) &= \sum_{k=1}^{\infty}\mathbb{P}_x(\tau_y=k\,,\,X_n=y \text{ pour au moins m entiers } n\geq k+1)\\
&= \sum_{k=1}^{\infty}\mathbb{P}_x(\tau_y=k)\mathbb{P}_x(X_n=y \text{ pour au moins m entiers } n\geq k+1|\tau_y=k)\\
&= \sum_{k=1}^{\infty}\mathbb{P}_x(\tau_y=k)\mathbb{P}_x(X_n=y \text{ pour au moins m entiers } n\geq k+1|X_k=y) \quad \text{(Markov)}\\
&= \sum_{k=1}^{\infty}\mathbb{P}_x(\tau_y=k)\mathbb{P}_y(X_n=y \text{ pour au moins m entiers } n\geq 1)\\
&= U(x,y)H^{(m)}(y,y).
\end{aligned}$$

□

On peut maintenant démontrer la proposition.

Démonstration : Remarquons tout d'abord que :

$$H^{(1)}(x,x) = \mathbb{P}_x(\tau_x < +\infty) = U(x,x).$$

La lemme précédent permet alors de montrer par récurrence que :

$$\forall m \geq 1,\ H^{(m)}(x,x) = (U(x,x))^m.$$

On obtient le résultat voulu en laissant tendre m vers $+\infty$ puisque :

$$H^{(m)}(x,x) \underset{m\to+\infty}{\to} H(x,x),$$

par continuité décroissante des probabilités. □

-Étude plus fini du temps de retour en 0 de la marche aléatoire symétrique sur $\mathbb{Z}$.

Il est classique que :

$$\mathbb{P}(S_{2n} = 0) = \binom{2n}{n}\frac{1}{4^n},$$

car l'événement est déterminé de manière unique par les n instants où la marche va à droite. On peut donc calculer la fonction G :

$$\forall s \in [0,1[,\ G(0,0,s) = \sum_{n=0}^{+\infty}\binom{2n}{n}\frac{1}{4^n}s^{2n} = \frac{1}{\sqrt{1-s^2}} \quad \text{(développement en série des fonctions puissances).}$$

D'après un lemme précédent :

$$\forall s \in [0,1[,\ U(0,0,s) = 1 - \frac{1}{G(0,0,s)} = 1 - \sqrt{1-x^2} = \sum_{n=1}^{+\infty}\binom{2n}{n}\frac{1}{4^n(2n-1)}s^{2n},$$

en utilisant le développement en série des fonctions puissances. Cela assure que pour tout $n \geq 1$:

$$\mathbb{P}(\tau_0 = 2n) = \binom{2n}{n}\frac{1}{4^n(2n-1)}.$$

De plus par convergence monotone :

$$\mathbb{E}(\tau_0) = \lim_{s\to1^-}\frac{d\,U(0,0,s)}{ds} = +\infty.$$

Le temps de retour en 0 est fini presque sûrement (puisque 0 est récurrent) mais n'est pas intégrable (on dit que 0 est récurrent nul).

Remarque 3.9.11. - *Selon la vitesse à laquelle on va, on peut passer l'étape 6 ou 7, par exemple en énonçant seulement le résultat de transience en dimension supérieure à 3 dans le développement.*

- *Dans le cas où la marche n'est pas symétrique sur $\mathbb{R}$, la loi forte des grands nombres implique la divergence presque-sûre vers $\pm\infty$ selon que la marche a tendance à aller vers la gauche ou la droite.*

3.10 Théorème de stabilité de Liapunov

Références : F. Rouvière, *Petit guide du calcul différentiel* – https://agreg-maths.fr/uploads/versions/1241/lyapunov.pdf

Leçons : 220, 221, *215*.

3.11 Nombre de zéros d'une EDO

Références : H. Queffélec et C. Zuily, *Analyse pour l'agrégation* – `https://agreg-maths.fr/uploads/versions/531/Densit%C3%A9%20dans%20C%C2%B0%20des%20fonctions%20continues%20nulle%20part%20d%C3%A9rivables.pdf`

Leçons : 221, 224, *220*.

3.12 Formule des compléments

Références : J. Bernis et L. Bernis, *Analyse pour l'agrégation de mathématiques* – https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/formule_complements.pdf

Leçons : 235, 236, 245, 265.

3.13 Lax-Milgram et application

Références : : F. Hirsch et G. Lacombe, *Éléments d'analyse fonctionnelle*– H. Brézis, *Analyse fonctionnelle* – https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/lax_milgram.pdf

Leçons : 208, 213, 222, *205.*

3.14 Équation de Schrödinger sur $\mathbb{R}$

Références : J. Rauch, *Partial differential equations* – `https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/schrodinger.pdf`

Leçons : 222, 239, 250, *201.*

3.15 Théorème taubérien fort

Références : X. Gourdon, *Analyse* – http://perso.eleves.ens-rennes.fr/~afontain/agregation3.html

Leçons : 209, 230, 235, 241, 243.

3.16 Inversion de Fourier

Références : J.M. Bony, *Cours d'analyse - Théorie des distributions et analyse de Fourier* – https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/inversion_fourier.pdf

Leçons : 234, 235, 236 ,239, 250.

3.17 Prolongement de la fonction Zêta

Références : B. Candelpergher, *Calcul intégral* – `http://perso.eleves.ens-rennes.fr/people/Antoine.Diez/zeta.pdf`

Leçons : 207, 239, 245, 265.

3.18 Théorème de Féjèr

Références : H. Queffélec et C. Zuily, *Analyse pour l'agrégation* – `https://agreg-maths.fr/uploads/versions/1169/fejer.pdf`

Leçons : 202, 209, 234, 246.

3.19 Quelques ordres moyens

Références : G. Tenenbaum, Introduction à la théorie analytique et probabiliste des nombres – `http://www.unige.ch/~alaurent/Agregation/Liste_DVP/Ordres_moyens.pdf`

Leçons : 224, 230.

3.20 Densité des fonctions continues nulle part dérivables

Références : H. Queffélec et C. Zuily, *Analyse pour l'agrégation* – https://agreg-maths.fr/uploads/versions/531/Densit%C3%A9%20dans%20C%C2%B0%20des%20fonctions%20continues%20nulle%20part%20d%C3%A9rivables.pdf

Leçons : 201, 205, 228.

4 Développements d'Algèbre et de Géométrie

4.1 Théorème de la base de Burnside

Références : M. Zavidovique, *Un max de maths* – http://perso.eleves.ens-rennes.fr/people/Antoine.Mouzard/DEV/baseburnside.pdf

Leçons : 101, 103, 104, 108, 151.

4.2 Coniques passant par 5 points

Références : J.D. Eiden, *Géométrie analytique classique* – https://agreg-maths.fr/uploads/versions/30/conique_5_points_scourte.pdf

Leçons : 152, 162, 171, 181, *151*.

4.3 Polygones réguliers constructibles

Références : D.J. Mercier, *Cours de Géométrie* – `http://www.unige.ch/~alaurent/Agregation/Liste_DVP/Polygones_reguliers_constructibles.pdf`

Leçons : 102, 121, 125, 151, 182, 183.

4.4 Réciprocité quadratique

Références : : P. Caldero et J. Germoni, *Histoires hédonistes de groupes et de géométrie* - Tome premier – `https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/reciprocite_quadratique.pdf`

Leçons : 101, 104, 120, 121, 123, 126, 170, 190.

4.5 Quaternions et rotations

Références : D. Perrin, *Cours d'algèbre* – A. Jeanneret et D. Lines, *Invitation à l'algèbre* – `http://perso.eleves.ens-rennes.fr/people/Antoine.Diez/quaternions.pdf`

Leçons : 101, 106, 108, 161, 182, 183.

4.6 Autour des endomorphismes semi-simples

Références : X. Gourdon, *Algèbre* – `http://perso.eleves.ens-rennes.fr/~tuntr932/agreg.html`

Leçons : 122, 141, 153, 154, 155.

4.7 Nombre de polynômes irréductibles sur les corps finis

Références : : S. Francinou, H. Gianella, *Exercices de mathématiques pour l'agrégation*, Algèbre 1 – `https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/polynomes_corps_finis.pdf`

Leçons : 123, 125, 141, 190.

4.8 Théorème de structure des groupes abéliens finis

Références : : G. Peyré, *L'algèbre discrète de la transformée de Fourier* – URLhttps://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/structure.pdf

Leçons : 102, 104, 107, 110, 120.

4.9 Théorème des deux carrés

Références : : D. Perrin, *Cours d'Algèbre* – `https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/deux_carres.pdf`

Leçons : 121, 122, 126.

4.10 Décomposition de Dunford effective

Références : P. Caldero et J. Germoni, *Histoires hédonistes de groupes et de géométrie* - Tome premier – `http://perso.eleves.ens-rennes.fr/people/Antoine.Diez/dunfordnewton.pdf`

Leçons : 153, 155, 157.

4.11 Frobenius Zolotarev

Références : V. Beck, J. Malick,G. Peyré, *Objectif Agrégation* – `http://www.unige.ch/~alaurent/agreg.html`

Leçons : 103, 105, 106, 123, *152*.

4.12 Théorème de structure des polynômes symétriques

Références : E. Ramis, C. Deschamps, J. Odoux, *Cours de mathématiques spéciales*, tome 1 – `http://perso.eleves.ens-rennes.fr/people/Antoine.Diez/polynomessymetriques.pdf`

Leçons : 144

4.13 Réduction de Frobenius

Références : X. Gourdon, *Algèbre* – P. Caldero, J. Germoni, *Histoires hédonistes de groupes et de géométrie*, Tome premier – `https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/frobenius.pdf`

Leçons : 150, 151, 153, 154, 159.

4.14 Sous-groupes distingués et noyaux de caractères

Références : : F. Ulmer, *Théorie des groupes*– G. Peyré, *L'algèbre discrète de la transformée de Fourier*– https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/sous_groupes_distingues_table_caracteres.pdf

Leçons : 103, 107.

4.15 Nombre de matrices diagonalisables sur les corps finis

Références : S. Francinou, H. Gianella, *Exercices de mathématiques pour l'agrégation*, Algèbre 1 – `https://perso.ens-rennes.fr/math/people/salim.rostam/files/agreg/D%c3%a9veloppements/Nombre_matrices_diagonalisables.pdf`

Leçons : 101, 104, 123, 150, 155, 190.

4.16 Irréductibilité des polynômes cyclotomiques sur $\mathbb{Z}$

Références : D. Perrin, *Cours d'algèbre* – `http://www.unige.ch/~alaurent/Agregation/Liste_DVP/Irreductibilite_polynomes_cyclotomiques.pdf`

Leçons : 102, 141, 144, 190.

4.17 Table des caractères de S_4

Références : G. Peyré, *L'algèbre discrète de la transformée de Fourier*– http://perso.eleves.ens-rennes.fr/~flemonni/agregation/developpements/Table_S4.pdf

Leçons : 101, 104, 105, 107, 160, 161.

4.18 Théorème de Sophie Germain

Références : S. Francinou, H. Gianella, S. Nicolas, *Oraux X-ENS*, Algèbre 1 – `https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/sophie_germain.pdf`

Leçons : 120, 121, 126, 142.

4.19 Algorithme de Berlekamp

Références : V. Beck, J. Malick, G. Peyré, *Objectif Agrégation* – https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/berlekamp.pdf

Leçons : 123, 141, 142, *151*.

5 Développements mixtes

5.1 Sous-groupes compacts de $GL_n(\mathbb{R})$

Références : F. Testard, *Analyse mathématique* – https://agreg-maths.fr/uploads/versions/1242/kakutani.pdf

Leçons : 106, 150, 181, 203, *101, 160, 253.*

5.2 Image de l'exponentielle

Références : M. Zavidovique, *Un max de maths.* – `http://www.unige.ch/~alaurent/Agregation/Liste_DVP/Image_exponentielle.pdf`

Leçons : 153, 156, 204, 214.

5.3 Convergence de méthodes itératives

Références : L. Dumas, *Modélisation à l'oral de l'agrégation* – `https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/methodes_iteratives.pdf`

Leçons : 157, 162, 226, 233.

5.4 Lemme de Morse

Référence : F. Rouvière, *Petit guide du calcul différentiel*

Leçons : 158, 170, 171, 214, 215.

Théorème 5.4.1 (*Lemme de Morse*)

Soit U un ouvert de $\mathbb{R}^n$ contenant 0 et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^3$. On suppose également que $df_0 = 0$ et que la Hessienne $d^2 f_0$ est non dégénérée (ie. inversible) et de signature $(p, n-p)$. Alors il existe ϕ un $\mathcal{C}^1$-difféomorphisme entre deux voisinages V_0 et $W_0 = \phi(V_0)$ de 0 dans $\mathbb{R}^n$ tel que $\phi(0) = 0$ et :

$$\forall x \in V_0,\ f(x) - f(0) = \phi_1(x)^2 + \cdots + \phi_p(x)^2 - \phi_{p+1}(x)^2 - \cdots - \phi_n(x)^2.$$

Pour démontrer ce résultat, nous allons avoir besoin d'un lemme.

Lemme 5.4.2

On se donne $A_0 \in \mathcal{S}_n(\mathbb{R})$ une matrice symétrique et inversible. Alors il existe V un voisinage de A_0 dans $\mathcal{S}_n(\mathbb{R})$ et une application $\rho \in \mathcal{C}^1(V, GL_n(\mathbb{R}))$ telle que :

$$\forall A \in V, \quad A = {}^t(\rho(A))A_0\rho(A).$$

Démonstration : On considère l'application :

$$f : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \to \mathcal{S}_n(\mathbb{R}) \\ M & \mapsto {}^tMA_0M \end{cases}.$$

Il s'agit d'une application de classe $\mathcal{C}^1$ car les coordonnées à l'arrivée sont polynômiales en celles de M. Calculons la différentielle de f en I_n. Soit $H \in \mathcal{M}_n(\mathbb{R})$. On a :

$$f(I_n + H) \underset{H \to 0}{=} A_0 + {}^tHA_0 + A_0H + o(\|H\|),$$

puisque A_0 est symétrique. Ainsi :

$$\forall H \in \mathcal{M}_n(\mathbb{R}), \quad df_{I_n}.H = {}^tHA_0 + A_0H.$$

On en déduit que $\ker df_{I_n} = A_0^{-1}\mathcal{A}_n(\mathbb{R})$, qui est de dimension $\frac{n(n-1)}{2}$. On considère un supplémentaire F de $\ker df_{I_n}$ dans $\mathcal{M}_n(\mathbb{R})$, et ψ la restriction de f à F. Puisque $I_n \in F$, la différentielle de ψ en I_n est la restriction à F de la différentielle de f en I_n.

Par choix de F, $d\psi_{I_n}$ est injective et donc inversible par le théorème du rang (puisque la dimension de $\mathcal{S}_n(\mathbb{R})$ est égale à la dimension de F). Le théorème d'inversion locale assure qu'il existe U un voisinage de I_n dans F, qu'on peut supposer inclus dans $\mathrm{GL}_n(\mathbb{R})$ puisque I_n est inversible et que $\mathrm{GL}_n(\mathbb{R})$ est un ouvert, et V un voisinage de A_0 dans $\mathcal{S}_n(\mathbb{R})$ tels que ψ soit un $\mathcal{C}^1$-difféomorphisme de U sur V. L'application $\rho := \psi^{-1}$ convient puisqu'on a alors :

$$\forall A \in V, \quad A = \psi(\rho(A)) = {}^t(\rho(A))A_0\rho(A).$$

□

On peut maintenant passer à la preuve du lemme de Morse.

Démonstration : Quitte à restreindre U, on peut supposer que U est une boule ouverte centrée en 0. On peut ainsi utiliser la formule de Taylor avec reste intégral à l'ordre 2 (ce qui est possible puisque les boules sont convexes). Ainsi pour tout $x \in U$:

$$f(x) - f(0) = \int_0^1 (1-t)\, {}^t x \,\mathrm{Hess} f(tx)\, x\, dt,$$

puisque 0 est un point critique de f. En posant $Q(x) := \displaystyle\int_0^1 (1-t)\,\mathrm{Hess} f(tx)\, dt \in \mathcal{S}_n(\mathbb{R})$, on a :

$$\forall x \in U, \quad f(x) - f(0) = {}^t x\, Q(x)\, x.$$

De plus Q est de classe $\mathcal{C}^1$ d'après le théorème de régularité sous l'intégrale, qui s'applique puisque f est supposée $\mathcal{C}^3$. Or un calcul direct donne $Q(0) = \frac{1}{2}\mathrm{Hess} f(0)$ qui est inversible et de signature $(p, n-p)$ par hypothèse. Le lemme précédent assure alors qu'il existe V un voisinage de $Q(0)$ dans $\mathcal{S}_n(\mathbb{R})$ et une application $\rho \in \mathcal{C}^1(V, \mathrm{GL}_n(\mathbb{R}))$ tels que :

$$\forall A \in V, \quad A = {}^t(\rho(A))\, Q(0)\, \rho(A).$$

La continuité de Q et le fait qu'elle soit à valeurs $\mathcal{S}_n(\mathbb{R})$ assure qu'il existe un voisinage W de 0 dans $\mathbb{R}^n$ tel que $Q(W) \subset V$. Il s'ensuit que :

$$\forall x \in W, \quad f(x) - f(0) = {}^t(\rho(Q(x))x)\, Q(0)\, \rho(Q(x))x.$$

Le théorème d'inertie de Sylvester assure qu'il existe $A \in \mathrm{GL}_n(\mathbb{R})$ telle que $Q(0) = {}^t A\, \Delta_{p,n-p}\, A$, où :

$$\Delta_{p,n-p} := \begin{pmatrix} I_p & 0 \\ 0 & -I_{n-p} \end{pmatrix}.$$

Si on pose $\phi : x \in W \mapsto A\rho(Q(x))x \in \mathbb{R}^n$, on définit une application de classe $\mathcal{C}^1$ par composition de telles applications. De plus on a :

$$\forall x \in W, \quad f(x) - f(0) = {}^t(\phi(x))\, \Delta_{p,n-p}\, \phi(x) \quad (*).$$

Calculons la différentielle de ϕ en 0. Si $h \in W$:

$$\phi(h) \underset{h \to 0}{=} A\rho(Q(h))h = A\rho(Q(0))h + o(\|h\|).$$

Ainsi la jacobienne de ϕ en 0 est la matrice $A\rho(Q(0)) = A \in \mathrm{GL}_n(\mathbb{R})$. Le théorème d'inversion locale assure que ϕ est un $\mathcal{C}^1$-difféomorphisme entre deux ouverts V_0 et $\phi(V_0)$ voisinages de 0. De plus l'égalité $(*)$ assure que :

$$\forall x \in V_0,\ f(x) - f(0) = \phi_1(x)^2 + \cdots + \phi_p(x)^2 - \phi_{p+1}(x)^2 - \cdots - \phi_n(x)^2.$$

□

Remarque 5.4.3. *On trouvera une belle application sur la stabilité des systèmes Hamiltoniens dans le livre de J. Bernis et L. Bernis : Analyse pour l'agrégation de mathématiques.*

5.5 Théorème des extrema liés

Références : F. Rouvière, *Petit guide de calcul différentiel*– J. Lafontaine, *Introduction aux variétés différentielles* – https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/extrema_lies.pdf

Leçons : 159, 214, 215, 219.

5.6 Simplicité de $SO_3(\mathbb{R})$

Références : P. Caldero et J. Germoni, *Histoires hédonistes de groupes et de géométrie* - Tome premier – D. Perrin, *Cours d'Algèbre* – `http://www.unige.ch/~alaurent/Agregation/Liste_DVP/SO3_simple.pdf`

Leçons : 103, 106, 108, 160, 161, 204.

5.7 Méthode du gradient à pas optimal

Références : J. Bernis et L. Bernis, *Analyse pour l'agrégation de mathématiques* – `http://www.unige.ch/~alaurent/Agregation/Liste_DVP/Methode_gradient_optimal.pdf`

Leçons : 162, 219, 226, 233, 253.

5.8 Homéomorphisme de exp : $\mathcal{S}_n(\mathbb{R}) \to \mathcal{S}_n^{++}(\mathbb{R})$

Références : P. Caldero, J. Germoni, *Histoires hédonistes de groupes et de géométrie*, Tome premier – `https://www.math.univ-toulouse.fr/~ckilque/documents_agregation/developpements_individuels/exponentielle_matrices_symetriques.pdf`

Leçons : : 150, 155, 156, 158, 160, 203.

6 Les bonus !

Je mets dans cette partie d'autres résultats qui peuvent parfois faire office de développement mais que j'ai pas testés. Pour l'instant je n'ai tapé qu'un document qui concerne le transport optimal discret. Il s'agit d'un problème d'optimisation convexe qui utilise le théorème de Krein-Milman en dimension finie que je redémontre (l'utilisation du théorème de séparation de Hahn-Banach n'est pas nécessaire donc il ne faut pas avoir peur si vous ne le connaissez pas !). Je trouve la preuve très jolie, le résultat rigolo et ça rentre dans plein de leçons. Je l'ai d'ailleurs mis dans mon plan d'Algèbre aux oraux (leçon 159 sur les formes linéaires et la dualité) et ça a intrigué le jury qui m'a demandé de raconter un peu le problème !

6.1 Un problème de transport optimal discret

Référence : Un sujet de devoir de prépa de N. Tosel et mon stage de Master 1 avec F. Bolley.

Leçons : 105, 106, 151, 159, 181, 219, 253.

Cadre du problème

On considère n tas de sable de même volume disons 1 par commodité et n fossés tous de volume 1 également. Le but étant de déplacer de manière optimale, c'est-à-dire en minimisant un certain coût, l'ensemble des n tas de sable dans les fossés. **A priori on s'autorise à repartir un même tas de sable dans plusieurs fossés.** On peut aussi imaginer le transport de marchandises dans des entrepôts sachant que chaque stock contient la même quantité de marchandises et chaque dépôt la même capacité de réception. Le but étant de minimiser le coût de transport qu'on peut donner sous la forme d'un coût par kilogramme de marchandises transportées d'un stock donné à un entrepôt donné. Modélisons mathématiquement le problème.

Modélisation d'un plan de déplacement (stratégie) :

Une stratégie de déplacement est donnée par une matrice de $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ où $m_{i,j}$ représente la quantité de sable déplacée du tas i au fossé j. Il est naturel de supposer qu'en plus, on a :

- $m_{i,j} \geq 0$
- $\forall i,\ \sum\limits_{j=1}^{n} m_{i,j} = 1$: il ne reste pas de sable dans chaque tas, tout le sable a été déplacé lorsque de l'opération puisque le volume de sable à déplacer a le même volume que l'ensemble des fossés.
- $\forall j,\ \sum\limits_{i=1}^{n} m_{i,j} = 1$ pour la même raison, chaque fossé est rempli entièrement.

On modélise donc une stratégie de déplacement par une matrice bistochastique (i.e. une matrice vérifiant les 3 conditions précédentes) : on note $\mathcal{B}_n$ leur ensemble.

Modélisation du coût de transport :

On introduit la matrice $C \in \mathcal{M}_n(\mathbb{R})$ dont le coefficient $c_{i,j}$ représente le coup volumique de transport du tas i au trou j. Ainsi le coût total de transport, qui dépend de C et de la stratégie M est :

$$C(M) = \sum_{i,j \leq n} c_{i,j} m_{i,j} = \mathrm{Tr}(C^T M) := C.M$$

Objectif :

Existe-t-il une stratégie qui réalise (et si oui laquelle) : $\inf_{M \in \mathcal{B}_n} C.M \quad ?$.

Le théorème de Krein-Milman

On commence par définir la notion de point extrémal d'un ensemble convexe.

Définition 6.1.1

Si K est un convexe, on dit que $M \in K$ est extrémal si $M = \frac{C_1+C_2}{2}$, avec C_1 et C_2 dans $K \Rightarrow C_1 = C_2 = M$.

On note Extr(K) leur ensemble.

Passons maintenant au fameux théorème de Krein-Milman, énoncé ici en dimension finie seulement car la preuve est bien plus simple.

Théorème 6.1.2

Soit K un convexe compact non vide d'un $\mathbb{R}$-espace vectoriel de dimension finie. Alors K est l'enveloppe convexe de ses points extrémaux.

On commence par démontrer l'existence d'un hyperplan d'appui.

Théorème 6.1.3

Soit $K \subset \mathbb{R}^n$ un compact convexe et soit $c \in \partial K$. Alors il existe $\phi \in (\mathbb{R}^n)^* \backslash \{0\}$ telle que :

$$\forall x \in K,\, \phi(x) \leq \phi(c).$$

C'est-à-dire que l'hyperplan affine $c + \ker(\phi)$ sépare c et K au sens large.

Démonstration : Quitte à se placer dans l'espace affine engendré par K, on se ramène au cas où $\mathring{K} \neq \emptyset$. Il suffit alors d'appliquer le théorème de Hahn-Banach géométrique (cas ouvert) aux convexes $\mathring{K}$ et $\{c\}$. En fait si K est d'intérieur vide, il est contenu dans un hyperplan. En effet, l'espace affine engendré par K ne peut pas être de dimension n sinon K contiendrait un simplexe engendré par une famille génératrice et ne serait donc pas d'intérieur vide.

Si on n'a pas envie d'utiliser le théorème de Hahn-Banach, on peut ruser en utilisant le théorème de projection. En effet, puisque $c \in \partial K$, on dispose d'une suite $(x_n)_n$ d'éléments qui ne sont pas dans K et qui converge vers c (puisque $\partial K \subset (\mathring{K})^C = \overline{K^C}$). On note y_n la projection de x_n sur K. Par propriété des angles obtus, on a :

$$\forall n \in \mathbb{N},\, \forall z \in K,\, \langle x_n - y_n, z - y_n \rangle \leq 0. \quad (*)$$

Posons alors pour tout n :

$$u_n = \frac{x_n - y_n}{\|x_n - y_n\|}.$$

Par compacité de la sphère unité en dimension finie, il existe $u \in \mathbb{R}^n$ de norme 1 telle qu'une sous-suite de $(u_n)_n$ converge vers u. On vérifie que $H := c + u^\perp$ est un hyperplan d'appui. L'inégalité $(*)$ assure que :

$$\forall z \in K,\, \langle u_n, z - y_n \rangle \leq 0.$$

Par continuité de l'application projection, on a $y_n \to p(c) = c$. On déduit que :

$$\forall z \in K,\ \langle u, z - c\rangle \leq 0.$$

La forme linéaire $\langle u, .\rangle$ convient. $\square$

Énonçons et démontrons un lemme avant de passer à la preuve du théorème de Krein-Milman.

Lemme 6.1.4

Soit $H = \phi^{-1}(\lambda)$ un hyperplan d'appui à K en $c \in \partial K$. $H \cap K$ est convexe compact et un point $d \in H \cap K$ est extrémal dans K si et seulement si il l'est dans $H \cap K$.

Démonstration : On raisonne par contraposée. Il est clair que si d n'est pas extrémal dans $H \cap K$, alors a fortiori il ne l'est pas dans K. Supposons que $d = \frac{x+y}{2}$ avec $x, y \in K$ et $x \neq y$, n'est pas extrémal dans K. On sait déjà que :

$$\phi(x) \leq \lambda \quad \text{et} \quad \phi(y) \leq \lambda.$$

On déduit par linéarité que :

$$\phi(d) = \phi\left(\frac{x+y}{2}\right) \leq \lambda = \phi(d) \quad (\text{car}\, d \in H).$$

D'où $\phi(x) = \phi(y) = \lambda$ donc $x, y \in H \cap K$. Cela montre exactement que d n'est pas extrémal dans $H \cap K$. $\square$

On peut enfin passer à la preuve théorème de Krein-Milman !

Démonstration : On suppose K non vide sinon il n'y a pas beaucoup d'intérêt. On raisonne par récurrence sur d la dimension du sous-espace affine engendré par K (c'est-à-dire si $x_0 \in K$: $\dim(\mathrm{Vect}(K - x_0))$).

- Si $d = 0$, alors $K = \{x_0\}$ et le résultat est trivial.
- Si le résultat est vrai pour les dimensions inférieures à $d-1$ et K " de dimension " d. Soit $x \in \partial K$, montrons que x s'écrit comme combinaison convexe de points extrémaux. Soit $H = \phi^{-1}(\lambda)$ un hyperplan d'appui à K en x. Si x est extrémal dans K il n'y a rien à montrer. Sinon x n'est pas extrémal dans $H \cap K$ d'après le lemme précédent. On peut appliquer l'hypothèse de récurrence à $H \cap K$ qui est bien de dimension inférieure à $d-1$. Ainsi x s'écrit comme combinaison convexe de point extrémaux dans $H \cap K$ qui sont des point extrémaux de K. Cela montre que :
$$\partial K \subset \mathrm{Conv}(\mathrm{Extr}(K)).$$
Si maintenant $x \in \mathring{K}$, alors une droite issue de x rencontre la frontière de K en deux points. Ainsi x est une combinaison convexe de deux points de la frontière qui sont eux même combinaison de points extrémaux. D'où :
$$K \subset \mathrm{Conv}(\mathrm{Extr}(K)),$$
ce qui achève la preuve puisque l'autre inclusion est immédiate.

$\square$

Propriétés de l'ensemble des matrices bistochastiques

Compacité

Théorème 6.1.5

$\mathcal{B}_n$ est un ensemble compact de $\mathcal{M}_n(\mathbb{R})$.

Démonstration : En effet, si on définit les applications suivantes continues :

$$p_{.j} : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \to \mathbb{R} \\ M = (m_{i,j}) & \mapsto \sum_{i=1}^{n} m_{i,j} \end{cases} .$$

$$p_{i.} : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \to \mathbb{R} \\ M = (m_{i,j}) & \mapsto \sum_{j=1}^{n} m_{i,j} \end{cases} .$$

Alors : $\mathcal{B}_n = \bigcap_{i=1}^{n} p_{i.}^{-1}(\{1\}) \cap \bigcap_{j=1}^{n} p_{.j}^{-1}(\{1\}) \cap \mathcal{M}_n(\mathbb{R}^+)$
Ainsi $\mathcal{B}_n$ est fermé comme intersection de fermés.
De plus, $\mathcal{B}_n$ est borné, par exemple pour la norme du maximum des coefficients par 1. Donc $\mathcal{B}_n$ compact.

$\square$

Convexité et points extrémaux

Théorème 6.1.6

$\mathcal{B}_n$ est un ensemble convexe.

Démonstration : En effet, si $\epsilon \in [0,1]$ et $B, B' \in \mathcal{B}_n$, alors $\epsilon B + (1-\epsilon)B' \in \mathcal{B}_n$.

$\square$

En fait on connait exactement les points extrémaux du convexe compact $\mathcal{B}_n$.

Théorème 6.1.7

Les points extrémaux de $\mathcal{B}_n$ sont les matrices de permutations données par :

$$\forall \sigma \in S_n,\ P_\sigma = (\boldsymbol{1}_{i=\sigma(j)}).$$

Démonstration : Commençons par montrer que les matrices de permutations sont bien des points extrémaux de $\mathcal{B}_n$. Si $P_\sigma = \frac{C+C'}{2}$, avec $C, C' \in \mathcal{B}_n$. Alors :

$$\forall j \in \{1,\dots,n\}\,,\ \frac{c_{\sigma(j),j} + c'_{\sigma(j),j}}{2} = 1$$

$$\forall k \neq j,\ \frac{c_{\sigma(j),k} + c'_{\sigma(j),k}}{2} = 0$$

Donc $C = C' = P_\sigma$. Étudions maintenant la réciproque.

On note J_n l'espace vectoriel des matrices dont la somme de chaque ligne et de chaque colonne est nulle. On exhibe un isomorphisme entre J_n et $\mathcal{M}_{n-1}(\mathbb{R})$.

$$S : \begin{cases} \mathcal{M}_{n-1}(\mathbb{R}) & \to J_n \\ A & \mapsto \begin{pmatrix} A & U \\ V & w \end{pmatrix} \end{cases}$$

où :

1. $V_j = -\sum\limits_{k=1}^{n-1} A_{k,j}$
2. $U_i = -\sum\limits_{l=1}^{n-1} A_{i,l}$
3. $w = -\sum\limits_{j=1}^{n-1} V_j$

Il s'agit bien d'un isomorphisme puisque :

- Injectivité : Si $A \neq A'$, alors $S(A) \neq S(A')$.
- Surjectivité : Si $M \in \mathcal{B}_n$, alors si $A = (m_{i,j})_{i,j \leq n-1}$, alors $M = S(A)$.

Conclusion : $\dim(J_n) = (n-1)^2$.

On note $C_{i,j}$ la forme linéaire sur $\mathcal{M}_n(\mathbb{R})$ qui renvoie le coefficient (i,j).

Définition 6.1.8

Soit $\alpha, \beta : \{1,\dots,n\}$ deux applications vérifiant : $\forall i,\ \alpha(i) < \beta(i)$.
$C_{\alpha,\beta} = \{M \in \mathcal{M}_n(\mathbb{R}),\ \forall (i,j),\ j \notin \{\alpha(i), \beta(i)\} \Rightarrow C_{i,j}(M) = 0\}$.

Remarquons que :

$$\dim(C_{\alpha,\beta}) \geq 2n.$$

Démontrons le lemme clé suivant.

Lemme 6.1.9

On considère $B \in \mathcal{B}_n$ telle qu'il existe α, β comme précédemment avec :

$$\forall i,\ C_{i,\alpha(i)}(B), C_{i,\beta(i)}(B) > 0.$$

Alors B n'est pas extrémale dans $\mathcal{B}_n$.

Démonstration : On cherche $C \in C_{\alpha,\beta}$ et $\epsilon > 0$ tels que :

- $B = \frac{B-\epsilon C}{2} + \frac{B+\epsilon C}{2}$
- $C \neq 0$
- $B \pm \epsilon C \in \mathcal{B}_n$

Choisissons maintenant grâce à ce qui précède C et ϵ.

1. On a :
$$\dim(J_n \cap C_{\alpha,\beta}) \geq 2n + (n-1)^2 - \dim(J_n + C_{\alpha,\beta}) > 0.$$
Il existe donc $C \in J_n \cap C_{\alpha,\beta} \backslash \{0\}$.
2. Choisissons ϵ.
   - Fixons i. Si $j \notin \{\alpha(i), \beta(i)\}$, alors $C_{i,j}(B \pm \epsilon C) = B_{i,j} \geq 0$.
   - Sinon on peut choisir $\epsilon > 0$ assez petit tel que pour tout i :
$$B_{i,\alpha(i)} \pm \epsilon C_{i,\alpha(i)} \quad \text{et} \quad B_{i,\beta(i)} \pm \epsilon C_{i,\beta(i)} \geq 0$$

Cela assure que si $B \in \mathcal{B}_n$ est telle qu'il existe α et β comme précédemment avec :
$$\forall i,\ C_{i,\alpha(i)}(B), C_{i,\beta(i)}(B) > 0$$
Alors B n'est pas extrémale dans $\mathcal{B}_n$.

□

Achevons maintenant la preuve du fait qu'une matrice extrémale est de permutation. On raisonne par récurrence sur la dimension n. On note :

P_n : "Une matrice extrémale de $\mathcal{B}_n$ est de permutation".

- **Initialisation :** $\mathcal{B}_1 = \{(1)\}$.
- **Hérédité :** Si P_{n-1} est vraie. Soit $B \in \mathcal{B}_n$. D'après ce qui précède il n'existe pas deux applications α et β comme précédemment. On déduit facilement que M est de la forme suivante :
$$\begin{pmatrix} * & * & 0 & * & * & * \\ * & * & \vdots & * & * & * \\ * & * & 0 & * & * & * \\ 0 & \dots & 1 & 0 & \dots & 0 \\ * & * & 0 & * & * & * \\ * & * & \vdots & * & * & * \\ * & * & 0 & * & * & * \end{pmatrix}.$$
La matrice à laquelle on enlève la ligne i et la colonne j est dans $\mathcal{B}_{n-1}$ et reste extrémale sinon B ne le serait pas. Par hypothèse de récurrence, c'est une matrice de permutation. Donc B est elle même une matrice de permutation.

Cela achève la preuve du théorème.

□

Résolution du problème initial

Théorème 6.1.10

Il existe une matrice de permutation P_σ réalisant :

$$\inf_{M \in \mathcal{B}_n} C.M.$$

Démonstration : On veut minimiser la fonction $M \in \mathcal{B}_n \mapsto C.M$. Cette application est continue car c'est la restriction d'une forme linéaire sur $\mathcal{M}_n(\mathbb{R})$ qui est donc continue (dimension finie).

Puisque $\mathcal{B}_n$ est un ensemble compact, on déduit que la fonction atteint bien son infimum C^* en une matrice $M \in \mathcal{B}_n$.

Par l'absurde : si le minimum n'est pas atteint en une matrice de permutation.
Alors d'après le théorème de Krein-Milman :

$$M = \sum_{i=1}^{s} \lambda_i P_{\sigma_i},$$

où $\sigma_i \in S_n$, $\lambda_i \in [0,1]$ avec $\sum \lambda_i = 1$. On a donc :

$$C.M = \sum_{i=1}^{s} \lambda_i C.P_{\sigma_i} > C^*,$$

ce qui est absurde.

$\square$

Conclusion : On a vu que le problème de transport optimal de tas de sable (ou de n'importe quoi d'autre) admet une solution qui est réalisée par le transport de chaque tas de sable entier vers un fossé. Ainsi indépendamment du coût de transport, il est inutile de découper les tas de sable pour espérer optimiser le coût.

Remarque 6.1.11. *En fait on a résolu un problème de transport optimal discret pour des probabilités qui chargent un ensemble fini de points avec la même masse, précisément on a montré que le problème de Kantorovitch (" s'autoriser à découper les tas de sable ") est équivalent au problème de Monge (" envoyer chaque tas dans un fossé ").*