

# **Compléments d'agrégation**

**Yves Coudène**

Version 2.4 , 14/02/2017

## Introduction

Ce texte réunit un certain nombre de compléments d'agrégation que j'ai eu l'occasion d'enseigner à l'université de Rennes 1 pendant la période 2002-2006. Certains de ces compléments sont des développements qui peuvent effectivement être présentés à l'oral de l'agrégation ; d'autres cherchent à attirer l'attention des candidats sur plusieurs points méconnus du programme. De manière générale, j'ai cherché à montrer qu'il n'y avait pas de rupture entre les programmes du premier cycle universitaire et ceux du second cycle ; et inciter les étudiants à faire un véritable effort de synthèse.

Je tiens à remercier ici Marie-Annick Paulmier pour son aide et sa disponibilité durant la rédaction de ces compléments ; ainsi que les collègues et étudiants, dont les remarques ont beaucoup contribué à l'amélioration de ce texte.

## Sommaire

Développement asymptotique de l'intégrale $\int \frac{\sin t}{t} dt$ .....	4
Formule d'inversion de Fourier .....	5
Somme des inverses des carrés des entiers .....	6
Formule de Stirling .....	8
Théorème du point fixe de Brouwer .....	9
Sur les séries semi-convergentes .....	11
Quaternions et rotations .....	13
Convergence dans les espaces de Hilbert .....	15
Un exemple de marche aléatoire .....	18
La loi forte des grands nombres $L^2$ .....	22
Le théorème de la limite centrée .....	24
Construction du pentagone régulier .....	26
Polyèdres platoniciens .....	27
Sous-groupes de $S_4$ .....	28
Trace, formes quadratiques et extensions de corps .....	29
Continuité des racines des polynômes .....	33
Sur la diagonalisation des matrices $2 \times 2$ .....	37
Théorie de Galois .....	41
Calculs en maple .....	43
Approximation de la loi gaussienne .....	45
Exercices d'oral .....	46
Commentaires .....	47

### Développement asymptotique de l'intégrale $\int \frac{\sin t}{t} dt$

L'intégrale  $\int_0^N \frac{\sin t}{t} dt$  tend vers  $\pi/2$  lorsque  $N$  tend vers l'infini. Quitte à faire quelques calculs, on peut obtenir un développement asymptotique complet :

$$\int_0^N \frac{\sin t}{t} dt = \int_0^N \left( \int_0^\infty e^{-tx} dx \right) \sin t dt = \int_0^\infty \int_0^N e^{-tx} \sin t dt dx.$$

L'intégrale  $\int_0^N e^{-tx} \sin t dt$  se calcule explicitement à l'aide des complexes.

$$\int_0^N e^{-tx} \sin t dt = \text{Im} \left( \int_0^N e^{-tx+it} dt \right) = \frac{1}{1+x^2} - \frac{e^{-Nx}}{1+x^2} (\cos N + x \sin N)$$

On obtient donc

$$\begin{aligned} \int_0^N \frac{\sin t}{t} dt &= \frac{\pi}{2} - \int_0^\infty \frac{e^{-Nx}}{1+x^2} (\cos N + x \sin N) dx, \quad v = Nx \\ &= \frac{\pi}{2} - \frac{\cos N}{N} \int_0^\infty \frac{e^{-v}}{1+\frac{v^2}{N^2}} dv - \frac{\sin N}{N^2} \int_0^\infty \frac{ve^{-v}}{1+\frac{v^2}{N^2}} dv \end{aligned}$$

On utilise maintenant le développement en série de  $\frac{1}{1+x^2}$ , dont le reste est explicite.

$$\begin{aligned} \frac{1}{1+\frac{v^2}{N^2}} &= \sum_{k=0}^K \left( -\frac{v^2}{N^2} \right)^k + \frac{\left( -\frac{v^2}{N^2} \right)^{K+1}}{1+\frac{v^2}{N^2}} \\ \int_0^\infty \frac{e^{-v}}{1+\frac{v^2}{N^2}} dv &= \sum_{k=0}^K \frac{(-1)^k}{N^{2k}} \underbrace{\int_0^\infty v^{2k} e^{-v} dv}_{=(2k)!} + \frac{1}{N^{2K+2}} \underbrace{\int_0^\infty \frac{(-v^2)^{K+1} e^{-v}}{1+\frac{v^2}{N^2}} dv}_{|\leq (2K+2)!} \\ \int_0^\infty \frac{e^{-v}}{1+\frac{v^2}{N^2}} dv &= \sum_{k=0}^K \frac{(-1)^k}{N^{2k}} (2k)! + o\left(\frac{1}{N^{2K}}\right) \\ \int_0^\infty \frac{ve^{-v}}{1+\frac{v^2}{N^2}} dv &= \sum_{k=0}^K \frac{(-1)^k}{N^{2k}} (2k+1)! + o\left(\frac{1}{N^{2K}}\right) \end{aligned}$$

Nous parvenons au résultat final

$$\begin{aligned} \int_0^N \frac{\sin t}{t} dt &= \frac{\pi}{2} - \sum_{k=0}^K (-1)^k \left[ (2k)! \frac{\cos N}{N^{2k+1}} + (2k+1)! \frac{\sin N}{N^{2k+2}} \right] + o\left(\frac{1}{N^{2K+2}}\right) \\ &= \frac{\pi}{2} - \frac{\cos N}{N} - \frac{\sin N}{N^2} + 2 \frac{\cos N}{N^3} + 6 \frac{\sin N}{N^4} + o\left(\frac{1}{N^4}\right) \end{aligned}$$

Ce développement illustre les points suivants :

- utilisation de la transformée de Laplace,
- calcul d'intégrales réelles en passant dans le domaine complexe,
- fonction Gamma,
- obtention d'une asymptotique en mettant l'expression à évaluer sous la forme d'une intégrale dépendant d'un paramètre,
- obtention d'un reste sous forme intégrale, ce qui permet d'obtenir des équivalents, mais aussi des encadrements.

Enfin pour finir, il faut remarquer que la série obtenue est divergente.

## Formule d'inversion de Fourier

Le théorème suivant est une version ponctuelle de la formule d'inversion de Fourier, à rapprocher du théorème de Dirichlet ; on pose  $\hat{f}(t) = \int e^{-itx} f(x) dx$ .

### Théorème

Soit  $f \in L^1(\mathbf{R})$ ,  $a \in \mathbf{R}$ . On suppose que  $f$  admet une limite à droite et à gauche en  $a$  ; on suppose que  $f$  est dérivable à droite et à gauche en  $a$ . Alors,

$$\frac{1}{2} (f(a^-) + f(a^+)) = \lim_{A \rightarrow \infty} \int_{-A}^A e^{iax} \hat{f}(x) \frac{dx}{2\pi}$$

### Preuve

– Quitte à traduire la variable, on peut supposer  $a = 0$ . On a

$$\int_{\mathbf{R}} \mathbf{1}_{[-A,A]}(x) \hat{f}(x) \frac{dx}{2\pi} = \int_{\mathbf{R}} \widehat{\mathbf{1}_{[-A,A]}}(x) f(x) \frac{dx}{2\pi} = \int_{\mathbf{R}} \frac{2 \sin Ax}{x} f(x) \frac{dx}{2\pi}.$$

On va donc montrer que  $\lim_{A \rightarrow \infty} \left( \int_0^\infty \frac{2 \sin Ax}{x} f(x) \frac{dx}{2\pi} - \frac{1}{2} f(0^+) \right) = 0$ .

– Remarquons que (en faisant le changement de variable  $y = Ax$ )

$$\int_0^\infty \frac{\sin Ax}{x} dx = \int_0^\infty \frac{\sin y}{y} dy = \frac{\pi}{2},$$

et qu'ainsi  $\int_0^\infty \frac{2 \sin Ax}{x} f(x) \frac{dx}{2\pi} - \frac{1}{2} f(0^+) = \int_0^\infty 2 \sin(Ax) \frac{f(x) - f(0^+)}{x} \frac{dx}{2\pi}$ .

Sans le facteur  $1/x$ , il suffirait d'appliquer le lemme de Riemann-Lebesgue :

**Lemme** (Riemann-Lebesgue) : Soit  $g \in L^1$ . Alors,  $\lim_{A \rightarrow \infty} \int_{\mathbf{R}} e^{iAx} g(x) dx = 0$ .

– Près de 0 :

on utilise l'hypothèse  $f(x) = f(0^+) + x f'(0^+) + x \varepsilon(x)$ , avec  $\lim_{x \rightarrow 0^+} \varepsilon(x) = 0$  ;

par conséquent, il existe  $\delta > 0$  tel que  $\frac{f(x) - f(0^+)}{x}$  soit borné sur  $]0, \delta]$ . La fonction  $\frac{f(x) - f(0^+)}{x} \mathbf{1}_{]0, \delta]}(x)$  est donc intégrable et par Riemann-Lebesgue,

$$\lim_{A \rightarrow +\infty} \int_0^\delta \sin(Ax) \frac{f(x) - f(0^+)}{x} dx = 0$$

– Loin de 0 :

sur  $[\delta, +\infty[$ , on a  $0 < 1/x < 1/\delta$ , et  $\frac{f(x)}{x} \mathbf{1}_{[\delta, \infty[}(x)$  est intégrable si bien que

$$\lim_{A \rightarrow +\infty} \int_\delta^\infty \sin(Ax) \frac{f(x)}{x} dx = 0.$$

Enfin, par définition des intégrales généralisées<sup>1</sup>, on a

$$\lim_{A \rightarrow +\infty} \int_\delta^\infty \frac{\sin(Ax)}{x} f(0^+) dx = \lim_{A \rightarrow +\infty} \int_{A\delta}^\infty \frac{\sin y}{y} dy f(0^+) = 0.$$

– On démontre de même que  $\lim_{A \rightarrow \infty} \int_{-\infty}^0 \frac{2 \sin Ax}{x} f(x) \frac{dx}{2\pi} = \frac{1}{2} f(0^-)$ .

**Remarque** : On aurait pu supposer  $f \in L^1$  ou  $L^2$  et  $\exists \delta > 0, \exists K > 0$  tels que  $\forall x \in ]0, \delta[$ ,  $|f(x) - f(a^+)| \leq K|x - a|$  ;  $\forall x \in ]-\delta, 0[$ ,  $|f(x) - f(a^-)| \leq K|x - a|$ .

<sup>1</sup>  $\int_0^\infty \frac{\sin y}{y} dy = \lim_{B \rightarrow \infty} \int_0^B \frac{\sin y}{y} dy$

### Somme des inverses des carrés des entiers

La somme  $\sum \frac{1}{k^2}$  peut se calculer en évaluant de deux façons différentes l'intégrale  $\int_0^1 \int_0^1 \frac{dx dy}{1-xy}$ . Commençons par développer  $\frac{1}{1-xy}$  en série.

$$\int_0^1 \int_0^1 \frac{dx dy}{1-xy} = \iint \sum_{k \geq 0} (xy)^k dx dy = \sum_{k=0}^{\infty} \int_0^1 x^k dx \int_0^1 y^k dy = \sum_{k=0}^{\infty} \frac{1}{(k+1)^2}.$$

Effectuons ensuite dans l'intégrale le changement de variables

$$\begin{cases} x = u - t = \cos \theta - t \\ y = u + t = \cos \theta + t \end{cases}$$

La différentielle de la transformation  $(\theta, t) \mapsto (x, y)$  vaut  $\begin{pmatrix} -\sin \theta & -1 \\ -\sin \theta & +1 \end{pmatrix}$ . Son déterminant vaut  $-2 \sin \theta$ . On a donc  $dx dy = 2 |\sin \theta| d\theta dt$ .

Comme la fonction  $(x, y) \mapsto \frac{1}{1-xy}$  est invariante lorsqu'on échange les coordonnées  $x$  et  $y$ , on peut se contenter de calculer son intégrale dans le domaine  $\{(x, y) \in [0, 1]^2 \mid x \leq y\}$ . L'image réciproque de ce domaine par l'application  $(u, t) \mapsto (x, y)$  est un triangle dont les sommets sont situés aux points  $(0, 0)$ ,  $(\frac{1}{2}, \frac{1}{2})$  et  $(1, 0)$ .

Ceci correspond aux régions 
$$\begin{array}{ll} 0 \leq t \leq 1 - \cos \theta & \text{si } 0 \leq \theta \leq \frac{\pi}{3} \\ 0 \leq t \leq \cos \theta & \text{si } \frac{\pi}{3} \leq \theta \leq \frac{\pi}{2} \end{array}$$

Rappelons trois formules, qui vont intervenir dans le calcul de l'intégrale :

$$\int_0^x \frac{a}{a^2 + t^2} dt = \operatorname{atan} \left( \frac{x}{a} \right), \quad \frac{1 - \cos \theta}{\sin \theta} = \tan \left( \frac{\theta}{2} \right), \quad \operatorname{atan}(x) + \operatorname{atan}(1/x) = \frac{\pi}{2}$$

On peut maintenant effectuer le changement de variables.

$$\begin{aligned} \int_0^1 \int_0^1 \frac{dx dy}{1-xy} &= 2 \iint \frac{2 \sin \theta}{1 - \cos^2 \theta + t^2} dt d\theta \\ &= 4 \int_0^{\pi/3} \left[ \operatorname{atan} \left( \frac{t}{\sin \theta} \right) \right]_0^{1 - \cos \theta} d\theta + 4 \int_{\pi/3}^{\pi/2} \left[ \operatorname{atan} \left( \frac{t}{\sin \theta} \right) \right]_0^{\cos \theta} d\theta \\ &= 4 \int_0^{\pi/3} \operatorname{atan} \left( \frac{1 - \cos \theta}{\sin \theta} \right) d\theta + 4 \int_{\pi/3}^{\pi/2} \operatorname{atan} \left( \frac{1}{\tan \theta} \right) d\theta \\ &= 4 \int_0^{\pi/3} \frac{\theta}{2} d\theta + 4 \int_{\pi/3}^{\pi/2} \left( \frac{\pi}{2} - \theta \right) d\theta \\ &= 4 \left( \frac{\pi^2}{36} + \frac{\pi^2}{12} - \left( \frac{\pi^2}{8} - \frac{\pi^2}{18} \right) \right) \\ &= \pi^2 \left( \frac{1}{9} + \frac{1}{3} - \frac{1}{2} + \frac{2}{9} \right) \\ &= \frac{\pi^2}{6} \end{aligned}$$

Voici une application de cette formule.

### Théorème

La famille de fonctions  $e_k(x) = e^{ikx}$ ,  $k \in \mathbf{Z}$ , forme une base hilbertienne de  $L^2(\mathbf{R}/2\pi\mathbf{Z}, \frac{dx}{2\pi})$ .

### Preuve

Rappelons que les combinaisons linéaires de fonctions indicatrices d'intervalles sont denses dans  $L^2$ . Il suffit donc d'approcher ces fonctions indicatrices par des polynômes trigonométriques pour démontrer le théorème.

Pour cela, on commence par approcher la fonction  $\varphi : \mathbf{R}/2\pi\mathbf{Z} \rightarrow \mathbf{R}$  donnée par  $\varphi(x) = x$  si  $x \in [0, 2\pi[$ . Le carré de la norme  $L^2$  de  $\varphi$  est donnée par

$$\|\varphi\|^2 = \frac{1}{2\pi} \int_0^{2\pi} x^2 dx = \frac{4\pi^2}{3}$$

Les produits scalaires relativement aux  $e_k$  valent

$$(\varphi | e_0) = \frac{1}{2\pi} \int_0^{2\pi} x dx = \pi$$

$$(\varphi | e_k) = \frac{1}{2\pi} \int_0^{2\pi} x e^{-ikx} dx = \frac{1}{2\pi} \left[ \frac{x e^{-ikx}}{-ik} \right]_0^{2\pi} - \frac{1}{2\pi} \int_0^{2\pi} \frac{x e^{-ikx}}{-ik} dx = \frac{i}{k}$$

Rappelons la formule  $\|u+v\|^2 = \|u\|^2 + \|v\|^2 + 2 \operatorname{Re}(u|v)$ , valide pour tout  $u, v$  éléments d'un Hilbert. Les fonctions  $\{e_k\}$  forment une famille orthonormée de  $L^2$  :  $\|e_k\| = 1$ ,  $(e_i | e_j) = 0$  si  $i \neq j$ . Par conséquent, on a l'égalité  $\|\sum_{-n}^n (\varphi | e_k) e_k\|^2 = \sum_{-n}^n |(\varphi | e_k)|^2$ . Ceci entraîne les égalités

$$\begin{aligned} \|\varphi - \sum_{-n}^n (\varphi | e_k) e_k\|^2 &= \|\varphi\|^2 + \sum_{-n}^n (\varphi | e_k)^2 - 2 \operatorname{Re}(\varphi | (\varphi | e_k) e_k) \\ &= \|\varphi\|^2 - \sum_{-n}^n |(\varphi | e_k)|^2 \\ &= \frac{4\pi^2}{3} - \pi^2 - 2 \sum_1^n \frac{1}{k^2} \\ &= 2 \left( \frac{\pi^2}{6} - \sum_1^n \frac{1}{k^2} \right) \end{aligned}$$

La fonction  $\varphi$  est donc dans l'adhérence de l'espace vectoriel engendré par les  $e_k$ . Soit  $\alpha \in [0, 2\pi[$  ; les fonctions  $\varphi(x + \alpha)$  sont également dans cet espace. Nous avons maintenant, pour tout  $x \in [0, 2\pi[$ ,

$$\varphi(x + \alpha) - \varphi(x) + 2\pi - \alpha = 2\pi \mathbf{1}_{[0, 2\pi - \alpha[}$$

On obtient les fonctions indicatrices d'intervalles de la forme  $[0, 2\pi - \alpha[$  puis, par soustraction, toutes les fonctions indicatrices intervalles.

### Remarque

Les preuves classiques font appel au théorème de Stone-Weierstrass, ou passent par le théorème de Fejer, qui donne une convergence uniforme pour une suite explicite de polynômes trigonométriques. Ici, on obtient le résultat sans sortir du cadre de la théorie de l'intégration.

## La formule de Stirling

Voici une preuve de la formule de Stirling qui utilise la méthode de Laplace.

**Théorème**  $n! \sim n^{n+1/2} e^{-n} \sqrt{2\pi}$ .

**Preuve**

Commençons par exprimer la factorielle à l'aide de la fonction  $\Gamma$ .

$$n! = \Gamma(n+1) = \int_0^{+\infty} e^{-t} t^n dt.$$

Effectuons les changements  $t = n(1+u)$  et  $u = s/\sqrt{n}$  dans cette intégrale.

$$\begin{aligned} \Gamma(n+1) &= \int_0^{+\infty} e^{-t} t^n dt = \left( \int_{-1}^{+\infty} e^{n(\ln(1+u)-u)} du \right) n^{n+1} e^{-n} \\ &= \left( \int_{-\sqrt{n}}^{+\infty} e^{n(\ln(1+\frac{s}{\sqrt{n}})-\frac{s}{\sqrt{n}})} ds \right) n^{n+\frac{1}{2}} e^{-n} \end{aligned}$$

On utilise ensuite les deux inégalités suivantes :

$$\begin{aligned} \text{Si } u \leq 1, \ln(1+u) &\leq u - \frac{u^2}{2} + \frac{u^3}{3} \leq u - \frac{u^2}{2} + \frac{u^2}{3} = u - \frac{u^2}{6}, \\ \text{Si } u \geq 1, \ln(1+u) &\leq \frac{u-1}{2} + \ln(2) \leq \ln(2) u. \end{aligned}$$

La seconde inégalité s'obtient en comparant le logarithme à sa tangente en  $x=1$ .

Pour les valeurs de  $s$  appartenant à l'intervalle  $[-\sqrt{n}, \sqrt{n}]$ , on peut appliquer le théorème de convergence dominée. En effet, la première inégalité montre que la fonction  $e^{n(\ln(1+\frac{s}{\sqrt{n}})-\frac{s}{\sqrt{n}})}$  est majorée par  $e^{-s^2/6}$ , qui est intégrable. De plus elle converge vers  $e^{-s^2/2}$ . Par conséquent,

$$\int_{-\sqrt{n}}^{\sqrt{n}} e^{n(\ln(1+\frac{s}{\sqrt{n}})-\frac{s}{\sqrt{n}})} ds \longrightarrow \int_{\mathbf{R}} e^{-s^2/2} ds = \sqrt{2\pi}.$$

L'intégrale sur l'intervalle  $[\sqrt{n}, +\infty[$  se majore grâce à la seconde inégalité.

$$\int_{\sqrt{n}}^{+\infty} e^{n(\ln(1+s/\sqrt{n})-s/\sqrt{n})} ds \leq \int_{\sqrt{n}}^{+\infty} e^{-0,3\sqrt{n}s} ds = \frac{e^{-0,3n}}{0,3\sqrt{n}} \longrightarrow 0.$$

Voici un encadrement plus précis (cf Feller, *Introduction to Probability*).

**Théorème**  $1 \leq \frac{n!}{n^{n+1/2} e^{-n} \sqrt{2\pi}} \leq e^{\frac{1}{12n}}$

**Preuve**

Soit  $u_n = \ln \left( \frac{n!}{n^{n+1/2} e^{-n} \sqrt{2\pi}} \right)$ . En vertu de l'égalité  $\text{ath}(t) = \frac{1}{2} \ln \left( \frac{1+t}{1-t} \right)$ , on a :

$$\begin{aligned} u_{n+1} - u_n &= 1 - \frac{(2n+1)}{2} \ln \left( 1 + \frac{1}{n} \right) = 1 - (2n+1) \text{ath} \left( \frac{1}{2n+1} \right) \\ &= - \sum_{i \geq 1} \frac{1}{2i+1} \frac{1}{(2n+1)^{2i}} = - \frac{1}{3} \frac{1}{(2n+1)^2} - \frac{1}{5} \frac{1}{(2n+1)^4} - \dots \\ &\geq - \frac{1}{3} \sum_{k \geq 1} \frac{1}{(2n+1)^{2k}} = \frac{1}{12} \frac{1}{n(n+1)} = \frac{1}{12} \left( \frac{1}{n+1} - \frac{1}{n} \right) \end{aligned}$$

La suite  $u_n - \frac{1}{12n}$  est croissante, bornée, donc convergente. Sa limite est égale à celle de  $u_n$ . Au final,  $u_n - \frac{1}{12n} \leq \ln(\sqrt{2\pi}) \leq u_n$ .



## Le théorème du point fixe de Brouwer

On démontre le théorème du point fixe de Brouwer par une méthode analytique. La démonstration est basée sur la formule suivante, qui est une conséquence bien connue de la formule de Green-Riemann.

### Théorème

Soit  $R = [0, 1]^2$ ,  $f, g : [0, 1]^2 \rightarrow \mathbf{R}$  deux fonctions  $C^1$ . Alors :

$$2 \iint_R df \wedge dg = \int_{\partial R} f dg - g df$$

avec les notations habituelles :

$$df \wedge dg = \det \begin{pmatrix} \partial_x f & \partial_x g \\ \partial_y f & \partial_y g \end{pmatrix} dx dy,$$

$$f dg - g df = (f \partial_x g - g \partial_x f) dx + (f \partial_y g - g \partial_y f) dy.$$

On a utilisé les abréviations  $\partial_x = \frac{\partial}{\partial x}$  et  $\partial_y = \frac{\partial}{\partial y}$  afin d'alléger les notations. Cette formule peut se démontrer sans passer par le théorème de Green-Riemann.

### Preuve

Le terme de gauche vaut  $\int_0^1 \int_0^1 \partial_x f \partial_y g - \partial_x g \partial_y f dx dy$ .

On intègre par partie,

$$\int \partial_x f \partial_y g dy = \left[ \partial_x f g \right]_{y=0}^{y=1} - \int \partial_y \partial_x f g dy,$$

$$\int \partial_x g \partial_y f dx = \left[ g \partial_y f \right]_{x=0}^{x=1} - \int g \partial_x \partial_y f dx,$$

ce qui donne

$$\int_0^1 \int_0^1 \partial_x f \partial_y g - \partial_x g \partial_y f dx dy = \int \left[ \partial_x f g \right]_{y=0}^{y=1} dx - \int \left[ g \partial_y f \right]_{x=0}^{x=1} dy.$$

Si nous intervertissons  $f$  et  $g$  dans cette formule, nous obtenons

$$\int_0^1 \int_0^1 \partial_x f \partial_y g - \partial_x g \partial_y f dx dy = - \int \left[ \partial_x g f \right]_{y=0}^{y=1} dx + \int \left[ f \partial_y g \right]_{x=0}^{x=1} dy.$$

Il suffit de sommer les deux dernières formules pour obtenir le résultat.

Notons  $B = \{(x, y) \in \mathbf{R}^2 \mid x^2 + y^2 \leq 1\}$  la boule unité fermée de  $\mathbf{R}^2$  et  $S^1$  son bord. Le théorème du point fixe de Brouwer se déduit du lemme suivant.

### Lemme

Il n'existe pas de fonction  $s : B \rightarrow S^1$  à la fois  $C^1$  et égale à l'identité sur  $S^1$ .

### Preuve

On applique la formule donnée plus haut en se plaçant en coordonnées polaires. Posons  $s(r \cos(2\pi\theta), r \sin(2\pi\theta)) = (f(r, \theta), g(r, \theta))$ . Les fonctions  $f$  et  $g$  sont définies pour  $(r, \theta) \in [0, 1] \times [0, 1]$ . Elles satisfont la relation  $f^2 + g^2 = 1$ . Dérivons cette relation.

$$f \begin{pmatrix} \partial_r f \\ \partial_\theta f \end{pmatrix} + g \begin{pmatrix} \partial_r g \\ \partial_\theta g \end{pmatrix} = 0,$$

si bien que le déterminant  $\begin{pmatrix} \partial_r f & \partial_r g \\ \partial_\theta f & \partial_\theta g \end{pmatrix}$  est nul pour tout  $(r, \theta) \in [0, 1]^2$ . Nous avons établi la formule

$$\int_0^1 \int_0^1 df \wedge dg = 0.$$

Calculons maintenant l'intégrale sur le bord en prenant en compte les égalités

$$\begin{aligned} f(1, \theta) &= \cos(2\pi\theta), & g(1, \theta) &= \sin(2\pi\theta), \\ f(0, \theta) &= f(0, 0), & g(0, \theta) &= g(0, 0), \\ f(r, 0) &= f(r, 1), & g(r, 0) &= g(r, 1). \end{aligned}$$

Les termes associés aux bords horizontaux supérieur et inférieur du carré se compensent en raison de la périodicité en  $\theta$ . Les fonctions  $f$  et  $g$  sont constantes sur le bord vertical gauche, il reste donc

$$\int_{\partial R} f dg - g df = \int_0^1 f(1, \theta) \partial_\theta g(1, \theta) - g(1, \theta) \partial_\theta f(1, \theta) d\theta = 2\pi.$$

C'est absurde et le théorème est démontré.

### **Théorème du point fixe de Brouwer**

*Toute application continue de  $B$  dans  $B$  admet un point fixe.*

#### **Preuve**

Commençons par le cas d'une application  $h : B \rightarrow B$  qui est  $C^1$  sans point fixe. Considérons l'application  $s$  qui à un point  $x \in B$  associe le point du bord  $S^1$  qui se trouve sur la demi-droite issue de  $h(x)$  et passant par  $x$ . Cette application est égale à l'identité sur  $S^1$ . On se convainc qu'elle est bien différentiable en calculant explicitement  $s(x)$ .

$$s(x) = x + \left( \sqrt{1 - |x|^2 + \langle x, u \rangle^2} - \langle x, u \rangle \right) u, \quad \text{avec } u = \frac{x - h(x)}{\|x - h(x)\|}.$$

Ceci est en contradiction avec le lemme démontré plus haut.

Le cas d'une fonction  $h$  continue s'obtient par approximation avec une fonction  $C^1$ . Rappelons que l'ensemble des fonctions de classe  $C^1$  sur le compact  $B$  forme une algèbre qui sépare les points ; le théorème de Stone-Weierstrass permet d'affirmer que ces fonctions sont denses dans les fonctions continues.

Soit donc  $h : B \rightarrow B$  une fonction continue sans point fixe. Par compacité, on peut trouver un  $\varepsilon > 0$  tel que  $\|x - h(x)\| \geq 2\varepsilon$  pour tout  $x \in B$ . Considérons  $\tilde{h} : B \rightarrow \mathbf{R}^2$  une fonction  $C^1$  qui vérifie  $\|\tilde{h}(x) - h(x)\| < \varepsilon$  pour tout  $x \in B$ . La fonction  $x \mapsto \frac{\tilde{h}(x)}{1+\varepsilon}$  est  $C^1$  et envoie la boule  $B$  dans elle-même. Elle a donc un point fixe  $x_0$ , ce qui conduit à l'absurdité suivante :

$$\varepsilon > \|h(x_0) - \tilde{h}(x_0)\| = \|h(x_0) - (1 + \varepsilon)x_0\| \geq \|h(x_0) - x_0\| - \varepsilon \geq \varepsilon$$

Le théorème est démontré.

## Sur les séries semi-convergentes

Une série numérique de terme général  $a_n$  est dite semi-convergente si

$$\sum_{n=0}^{+\infty} |a_n| = +\infty \quad \text{et} \quad \sum_{n=0}^N a_n \quad \text{converge quand } N \rightarrow +\infty$$

Dans ce cas, il est possible, en réordonnant les termes de la série, d'obtenir une limite différente de la limite initiale. Étant donné un réel  $\alpha$  fixé, il est même possible de réordonner les termes de la série, de sorte à ce qu'elle converge vers  $\alpha$ .

Ce phénomène peut être illustré sur un cas très simple. Considérons la série

$$\sum_{n \geq 1} \frac{(-1)^{n+1}}{n}$$

Elle converge vers  $\ln(2)$ . La preuve classique consiste à utiliser le développement en série de la fonction  $x \mapsto \ln(1+x)$ .

$$\ln(1+x) = \sum_{n=1}^{+\infty} (-1)^{n+1} \frac{x^n}{n}$$

valide pour tout  $x \in ]-1, 1]$ . Justifier l'égalité en  $x = 1$  est cependant délicat car on est sur le bord du domaine de convergence. L'usage est de recourir à un théorème d'Abel ; cf par exemple Ramis-Deschamps T4 3.1.3.

Voici une preuve de l'égalité  $\ln 2 = \sum_{n \geq 1} \frac{(-1)^{n+1}}{n}$ , qui fait appel au calcul intégral. L'idée est de partir de l'égalité  $\frac{1}{n} = \int_0^1 x^{n-1} dx$  et de regrouper les termes deux à deux de façon à travailler avec des quantités positives.

$$\begin{aligned} \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} &= \sum_{k=0}^{+\infty} \left( \frac{1}{2k+1} - \frac{1}{2k+2} \right) \\ &= \sum_{k=0}^{+\infty} \int_0^1 (x^{2k} - x^{2k+1}) dx \\ &= \int_0^1 \sum_{k=0}^{+\infty} (x^{2k} - x^{2k+1}) dx \\ &= \int_0^1 \left( \frac{1}{1-x^2} - \frac{x}{1-x^2} \right) dx \\ &= \int_0^1 \frac{1}{1+x} dx \\ &= \ln(2) \end{aligned}$$

L'interversion du signe somme et du signe intégral est correcte car les fonctions  $x \mapsto x^{2k} - x^{2k+1}$  sont positives sur  $[0, 1]$ . Remarquons que la première égalité est valide car la série  $\sum \frac{(-1)^{n+1}}{n}$  est une série alternée, donc convergente.

Derrière ce calcul se cache un résultat méconnu d'intégration.

### Théorème

Soit  $(X, \mathcal{T}, \mu)$  un espace mesuré,  $(a_n(x))_{n \in \mathbf{N}}$  une suite de fonctions  $\mu$ -intégrables telle que pour  $\mu$ -presque tout  $x \in X$  la suite  $(a_n(x))_{n \in \mathbf{N}}$  soit alternée. Alors

$$\int_X \sum_{n \in \mathbf{N}} a_n(x) d\mu(x) = \sum_{n \in \mathbf{N}} \int_X a_n(x) d\mu(x).$$

### Remarque

– Comme la suite  $(|a_n(x)|)_{n \in \mathbf{N}}$  est décroissante, il suffit de vérifier que  $a_0$  est intégrable.

– Dans l'exemple précédent, la suite  $a_n$  correspond à  $a_n(x) = (-1)^{n+1} x^{n-1}$ . On a  $\sum_{n \geq 1} \int |a_n(x)| dx = +\infty$ , si bien que le théorème classique d'interversion somme-intégrale ne s'applique pas.

La preuve du théorème procède comme plus haut. On commence par se placer sur l'ensemble des  $x$  tels que  $a_0(x) \geq 0$ . On groupe les termes par deux de façon à obtenir des fonctions à valeurs positives, ce qui permet d'appliquer le théorème usuel d'interversion. On fait ensuite de même sur l'ensemble  $\{x \in X \mid a_0(x) \leq 0\}$ . Remarquons enfin que le terme général  $\int a_n d\mu$  tend vers 0, par convergence dominée.

Montrons maintenant qu'on peut réordonner les termes de la série  $\sum \frac{(-1)^{n+1}}{n}$  de façon à obtenir un résultat différent de  $\ln(2)$ . Pour cela, faisons suivre un terme positif par deux termes négatifs ; considérons les sommes partielles

$$\sum_{n \geq 0} \frac{1}{2n+1} - \frac{1}{4n+2} - \frac{1}{4n+4}.$$

$$\text{On a } \frac{1}{2n+1} - \frac{1}{4n+2} - \frac{1}{4n+4} = \frac{1}{2n+1} - \frac{1}{2(2n+1)} - \frac{1}{2(2n+2)} = \frac{1}{2(2n+1)} - \frac{1}{2(2n+2)}.$$

Par conséquent,

$$\sum_{n=0}^N \frac{1}{2n+1} - \frac{1}{4n+2} - \frac{1}{4n+4} = \frac{1}{2} \sum_{n=0}^N \frac{1}{2n+1} - \frac{1}{2n+2}.$$

La série converge donc vers  $\frac{\ln 2}{2}$ .

**Référence** : Ramis-Deschamps, T4, 1.5.2.

**Exercice** : Montrer que

$$\int_0^1 \left( \sum_{k=0}^{+\infty} x^{2k} - x^{4k+1} - x^{4k+3} \right) dx \neq \sum_{k=0}^{+\infty} \int_0^1 (x^{2k} - x^{4k+1} - x^{4k+3}) dx.$$

## Quaternions et rotations

J'explique brièvement comment employer les quaternions pour étudier les rotations de  $\mathbf{R}^3$ . Les démonstrations sont laissées en exercice.

### Quaternions

Soit  $\mathbf{C} = \mathbf{R} + i'\mathbf{R}$  l'ensemble des nombres complexes. On considère les matrices de  $M_2(\mathbf{C})$  suivantes, appelées matrices de Pauli.

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i' \\ i' & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Ces matrices obéissent aux règles de commutation suivantes.

$$\forall i, j, \quad \sigma_i \sigma_j = -\sigma_j \sigma_i \quad \text{si } i \neq j; \quad \forall i, \quad \sigma_i^2 = 1.$$

### Théorème

Les matrices  $1, \sigma_1, \sigma_2, \sigma_3, \sigma_1\sigma_2, \sigma_2\sigma_3, \sigma_3\sigma_1, \sigma_1\sigma_2\sigma_3$  forment une base de l'ensemble des matrices  $2 \times 2$  à coefficients complexes,  $M_2(\mathbf{C})$ .

L'algèbre  $M_2(\mathbf{C})$  se décompose donc en

- un espace vectoriel de dimension 3,  $\mathbf{E} = \text{Vect}(\sigma_1, \sigma_2, \sigma_3)$  ; Un vecteur  $\vec{a} \in \mathbf{E}$  s'écrit  $\vec{a} = a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3$  ;
- une sous-algèbre de dimension 4,  $\mathbf{H} = \text{Vect}(1, \sigma_1\sigma_2, \sigma_2\sigma_3, \sigma_3\sigma_1)$  ;
- un sous-espace de dimension 1 engendré par  $\sigma_1\sigma_2\sigma_3$ .

On pose  $i = \sigma_1\sigma_2\sigma_3 = i' \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ . On a bien  $i^2 = -1$ . On a également  $\sigma_1\sigma_2 = i\sigma_3, \sigma_2\sigma_3 = i\sigma_1, \sigma_3\sigma_1 = i\sigma_2$ . Attention :  $i \notin \mathbf{H}$ .

Par conséquent, tout élément de  $\mathbf{H}$  peut se mettre sous la forme

$$q = \alpha + i \vec{a}, \quad \alpha \in \mathbf{R}, \quad \vec{a} \in \mathbf{E}.$$

La loi de composition induite sur  $\mathbf{H}$  par le produit matriciel peut se récrire

$$(\alpha + i \vec{a})(\alpha' + i \vec{a}') = \alpha\alpha' + i(\alpha \vec{a}' + \alpha' \vec{a}) - \vec{a} \vec{a}'$$

$$\text{avec } \vec{a} \vec{a}' = \vec{a} \mid \vec{a}' + i \vec{a} \times \vec{a}', \quad \mid \text{ prod. scalaire, } \times \text{ prod. vectoriel.}$$

### Remarques

$i$  commute avec les vecteurs de  $\mathbf{E}$ . Le produit de deux vecteurs de  $\mathbf{E}$  donne un élément de  $\mathbf{H}$ . Enfin le carré d'un vecteur est égal au carré de sa norme.

On pose  $\bar{q} = \alpha - i \vec{a}$  et on vérifie la formule  $q\bar{q} = \alpha^2 + a_1^2 + a_2^2 + a_3^2$ . Tout élément non nul de  $\mathbf{H}$  est donc inversible. Le corps  $\mathbf{H}$  est appelé corps des quaternions. Il vient d'être construit comme sous-algèbre de  $M_2(\mathbf{C})$ .

### Rotations

La construction précédente peut être utilisée pour étudier les rotations de  $\mathbf{R}^3$ .

Pour cela, on remarque que l'exponentielle de matrices se restreint aux quaternions.

$$e^{i \vec{a}} = 1 + i \vec{a} - 1/2 \|\vec{a}\|^2 + \dots = \cos \|\vec{a}\| + i \frac{\vec{a}}{\|\vec{a}\|} \sin \|\vec{a}\| .$$

en posant  $\|\vec{a}\|^2 = a_1^2 + a_2^2 + a_3^2$ , pour  $\vec{a} \in \mathbf{E}$ . En particulier le quaternion  $e^{i \vec{a}}$  est de norme 1.

### Théorème

Soit  $\vec{a} \in \mathbf{E}$  un vecteur unitaire et  $\theta$  un nombre réel. La rotation d'axe dirigé par  $\vec{a}$  et d'angle  $\theta$  est donnée par

$$\begin{aligned} \mathbf{E} &\rightarrow \mathbf{E} \\ \vec{x} &\rightarrow e^{-\frac{1}{2}i \theta \vec{a}} \vec{x} e^{\frac{1}{2}i \theta \vec{a}} \end{aligned}$$

Ceci se vérifie en décomposant  $\vec{x}$  selon un vecteur proportionnel à  $\vec{a}$  et un vecteur orthogonal à  $\vec{a}$ . Un calcul direct redonne l'expression classique des rotations.

$$e^{-\frac{1}{2}i \theta \vec{a}} \vec{x} e^{\frac{1}{2}i \theta \vec{a}} = \cos \theta \vec{x} + (1 - \cos \theta)(\vec{x} | \vec{a}) \vec{a} + \vec{a} \times \vec{x} \sin \theta.$$

De là, il est facile de composer les rotations. On obtient par exemple une expression pour le cosinus de l'angle  $\omega$  de la rotation composée d'une rotation d'angle  $\theta$  d'axe  $\vec{a}$  suivie d'une rotation d'angle  $\phi$  d'axe  $\vec{b}$ .

$$e^{-\frac{1}{2}i \omega \vec{u}} = e^{-\frac{1}{2}i \phi \vec{b}} e^{-\frac{1}{2}i \theta \vec{a}}$$

$$e^{-\frac{1}{2}i \omega \vec{u}} = \cos(\omega/2) - i \vec{u} \sin(\omega/2)$$

$$e^{-\frac{1}{2}i \phi \vec{b}} e^{-\frac{1}{2}i \theta \vec{a}} = (\cos(\phi/2) - i \vec{b} \sin(\phi/2)) (\cos(\theta/2) - i \vec{a} \sin(\theta/2))$$

$$= \cos(\phi/2) \cos(\theta/2) - (\vec{a} | \vec{b}) \sin(\phi/2) \sin(\theta/2) \dots$$

$$\dots - i (\vec{b} \sin(\phi/2) \cos(\theta/2) + \vec{a} \cos(\phi/2) \sin(\theta/2) + \vec{b} \times \vec{a} \sin(\phi/2) \sin(\theta/2)).$$

Donc

$$\cos \frac{1}{2} \omega = \cos \frac{1}{2} \phi \cos \frac{1}{2} \theta - (\vec{a} | \vec{b}) \sin \frac{1}{2} \phi \sin \frac{1}{2} \theta.$$

### Exercices

- Montrer que le groupe des quaternions de norme 1 s'identifie à  $SU_2(\mathbf{C})$ .
- Expliquez le morphisme de  $SU_2(\mathbf{C})$  dans  $SO_3(\mathbf{R})$  donné par le théorème plus haut. Quel est son noyau ?
- Démontrer l'égalité  $(\vec{x} \times \vec{y}) \times \vec{z} = (\vec{z} | \vec{x}) \vec{y} - (\vec{z} | \vec{y}) \vec{x}$  par un calcul direct avec les quaternions.
- Soit  $\vec{u}$  un vecteur unitaire. A quoi correspond la transformation de  $\mathbf{E}$  donnée par  $\vec{x} \mapsto -\vec{u} \vec{x} \vec{u}$  ?
- Montrer que l'application  $\vec{a} \mapsto e^{i \vec{a}}$ , définie de  $\mathbf{E}$  dans l'ensemble des quaternions de norme 1, est surjective. Est-ce un morphisme ?

**Référence :** G. Casanova, *Que-sais-je ?* 1657 "L'algèbre vectorielle".

## Convergence dans les espaces de Hilbert

L'étude de la convergence dans les espaces de Hilbert peut être abordée ou bien en se basant sur la complétude de l'espace, ou bien en utilisant la compacité faible de sa boule unité. Voici un exemple qui permet de comparer ces deux approches.

### Théorème

Soit  $H$  un espace de Hilbert,  $E_i$  une suite de sous-espaces vectoriels fermés dans  $H$ , croissante pour l'inclusion :  $E_i \subset E_j$  si  $i < j$ . On note  $E_\infty$  l'adhérence de l'union des  $E_i$ ,  $i \in \mathbf{N}$ . Les projecteurs orthogonaux sur les  $E_i$  sont notés  $\pi_i$ . Alors, pour tout  $v \in H$ , on a la convergence en norme

$$\pi_i v \xrightarrow{i \rightarrow \infty} \pi_\infty v.$$

### Première preuve

Cette preuve est basée sur la complétude de  $H$ . Montrons que la suite  $\pi_i v$  est convergente, en vérifiant qu'elle est de Cauchy. Soit  $i, j$  deux entiers tels que  $i < j$ . Comme  $E_i$  est inclus dans  $E_j$ , et que les  $\pi_i$  sont des projecteurs orthogonaux, nous avons

$$\pi_i^2 = \pi_i, \quad \pi_i^* = \pi_i, \quad \pi_j \pi_i = \pi_i.$$

Ceci implique les égalités

$$\langle \pi_i v, \pi_j v \rangle = \langle \pi_j^* \pi_i v, v \rangle = \langle \pi_j \pi_i v, v \rangle = \langle \pi_i v, v \rangle = \langle \pi_i^* \pi_i v, v \rangle = \|\pi_i v\|^2.$$

En développant, on obtient

$$\|\pi_j v - \pi_i v\|^2 = \|\pi_i v\|^2 + \|\pi_j v\|^2 - 2\langle \pi_i v, \pi_j v \rangle = \|\pi_j v\|^2 - \|\pi_i v\|^2.$$

La suite  $\|\pi_i v\|^2$  est donc croissante. Comme elle est majorée par  $\|v\|^2$ , elle converge et c'est une suite de Cauchy. L'égalité précédente montre que la suite  $\pi_i v$  est aussi une suite de Cauchy, elle converge vers une limite notée  $v_\infty$ .

Vérifions que  $v_\infty$  est bien la projection de  $v$  sur  $E_\infty$ . Tous les  $\pi_i v$  sont dans  $E_\infty$ , il en va de même pour leur limite  $v_\infty$ . Ensuite, les termes de la suite  $v - \pi_n v$  sont orthogonaux à  $E_i$  dès que  $n$  est supérieur à  $i$ . La limite  $v - v_\infty$  est orthogonale à  $E_i$  pour tout  $i$ , elle appartient donc à l'orthogonal de  $E_\infty$ .

$$v_\infty \in E_\infty, \quad v - v_\infty \in E_\infty^\perp.$$

Ces propriétés caractérisent la projection de  $v$  sur  $E_\infty$ , ce qui termine la preuve.

### Seconde preuve

La suite  $\pi_i v$  est bornée par  $\|v\|$ , montrons qu'elle converge faiblement. Pour cela, il suffit de vérifier qu'elle possède une unique valeur d'adhérence. Remarquons que les espaces  $E_i, E_\infty$  sont fermés pour la topologie faible. On peut donc raisonner comme plus haut : toute valeur d'adhérence de  $\pi_i v$  est à la fois dans  $E_\infty$  et dans  $v + E_\infty^\perp$ , elle est donc nécessairement égale à  $\pi_\infty v$ .

La suite  $\pi_n v - \pi_\infty v$  converge faiblement vers zéro. Ceci implique en particulier

$$\langle \pi_\infty v - \pi_n v, v \rangle \xrightarrow{n \rightarrow \infty} 0$$

Mais cette quantité est égale au carré de la norme que nous cherchons à minorer.

$$\|\pi_\infty v - \pi_n v\|^2 = \langle (\pi_\infty - \pi_n)^* (\pi_\infty - \pi_n) v, v \rangle = \langle (\pi_\infty - \pi_n)^2 v, v \rangle = \langle (\pi_\infty - \pi_n) v, v \rangle$$

Cette norme converge donc vers zéro et le théorème est démontré.

Ces deux approches ne diffèrent pas de manière fondamentale. Il est théoriquement possible de convertir toute preuve à base de complétude en démonstration faisant appel à de la compacité faible, et vice-versa.

### **Théorème**

*Soit  $H$  un espace pré-hilbertien, c'est-à-dire un espace vectoriel muni d'une forme bilinéaire symétrique non dégénérée. Alors  $H$  est complet si et seulement si sa boule unité est compacte pour la topologie faible.*

### **Preuve**

Soit  $f_n$  une suite d'éléments de  $H$  qui est bornée, vérifions qu'elle admet une sous-suite faiblement convergente. Pour cela, on va appliquer le procédé diagonal de Cantor aux coordonnées des  $f_n$ , dans une certaine base orthonormée. Le sous-espace engendré par les  $f_n$ ,  $n \in \mathbf{N}$ , est noté  $\tilde{H}$ . Soit  $e_i$  une base hilbertienne de  $\tilde{H}$  ; cette base peut être obtenue en appliquant le procédé d'orthogonalisation de Gram-Schmidt à la suite  $\{f_i\}$ . On cherche un vecteur  $f_\infty \in \tilde{H}$  et une sous-suite  $n_k$  tel que

$$\forall i \in \mathbf{N}, \quad \langle f_{n_k} - f_\infty, e_i \rangle \xrightarrow{k \rightarrow \infty} 0.$$

La suite  $\langle f_k, e_0 \rangle$  est bornée, on peut extraire une sous-suite qui converge vers un réel  $c_0$ . De cette sous-suite, on extraie une sous-suite de  $\langle f_k, e_1 \rangle$  qui converge vers un réel  $c_1$ , et ainsi de suite. En prenant le  $k^{ieme}$  terme de la  $k^{ieme}$  sous-suite, on obtient une suite  $n_k$  telle que pour tout  $i \in \mathbf{N}$ , il existe  $c_i \in \mathbf{R}$ ,

$$\langle f_{n_k}, e_i \rangle \xrightarrow{n \rightarrow \infty} c_i.$$

Il reste à vérifier que la série  $\sum c_i e_i$  converge en norme vers une certaine fonction  $f_\infty$ . Par orthogonalité,

$$\|\sum c_i e_i\|^2 = \sum |c_i|^2.$$

Montrons que la série  $\sum |c_i|^2$  est bornée. Fixons  $N \in \mathbf{N}$ . Par définition des  $c_i$ , on peut trouver  $k \in \mathbf{N}$  tel que

$$\sum_0^N |c_i|^2 \leq \sum_0^N |\langle f_{n_k}, e_i \rangle|^2 + 1 \leq \|f_{n_k}\|^2 + 1.$$

Cette dernière quantité est bornée indépendamment de  $N$  et  $k$ . Ceci assure la convergence de la série  $\sum |c_i|^2$ , qui est donc de Cauchy. On en déduit que la suite  $\sum c_i e_i$  est aussi de Cauchy, ce qui établit sa convergence.



Montrons que  $H$  est complet si sa boule unité est faiblement compacte. Soit donc  $\{f_n\}$  une suite de Cauchy : pour tout  $\varepsilon > 0$ , on peut trouver  $N \in \mathbf{N}$  tel que pour tout  $m, n \geq N$ ,

$$\|f_m - f_n\|^2 < \varepsilon.$$

On considère une sous-suite  $\{n_k\}$  telle que  $f_{n_k}$  converge faiblement vers une limite, qu'on peut supposer nulle sans perte de généralité. Pour chaque  $m \geq N$ , on choisit  $k$  tel que  $n_k$  est supérieur à  $N$ , et pour lequel

$$|\langle f_m, f_{n_k} \rangle| < \varepsilon.$$

Ceci implique la majoration

$$\|f_m\|^2 \leq \|f_m - f_{n_k}\|^2 + 2\langle f_m, f_{n_k} \rangle \leq 3\varepsilon.$$

La suite  $\{f_n\}$  converge vers zéro et le théorème est démontré.

Donnons maintenant une application du résultat présenté plus haut.

### **Théorème de convergence des martingales $L^2$**

*Soit  $(X, \mathcal{T}, \mu)$  un espace probabilisé,  $\mathcal{T}_i$  une suite de tribus incluses dans  $\mathcal{T}$ , croissante pour l'inclusion :  $\mathcal{T}_i \subset \mathcal{T}_j$  si  $i < j$ . Supposons que les  $\mathcal{T}_i$  engendrent  $\mathcal{T}$ . Alors pour tout  $f \in L^2(X)$ , on a la convergence en norme  $L^2$*

$$E(f \mid \mathcal{T}_n) \xrightarrow[n \rightarrow \infty]{} f.$$

### **Preuve**

Rappelons que l'espérance conditionnelle de  $f$  relativement à la tribu  $\mathcal{T}_n$  est la projection orthogonale de  $f$  sur  $L^2(X, \mathcal{T}_n)$ . Notons  $\pi_\infty$  la limite de ces projections. Il faut montrer que cette limite est égale à l'identité. Comme les fonctions étagées sont denses dans les fonctions  $L^2$ , il suffit de vérifier que la famille d'ensembles suivante coïncide avec  $\mathcal{T}$  :

$$\mathcal{T}' = \{A \in \mathcal{T} \mid \pi_\infty \mathbf{1}_A = \mathbf{1}_A\}.$$

Cette classe contient l'union des  $\mathcal{T}_i$ , qui est une algèbre de parties de  $X$ . De plus, c'est une classe monotone, elle est invariante par passage au complémentaire, et si  $A_n$  est une suite croissante d'éléments de  $\mathcal{T}'$ , alors la suite  $\mathbf{1}_{A_n}$  converge vers  $\mathbf{1}_{\cup A_n}$  en norme  $L^2$  par le théorème de convergence dominée, si bien que

$$\pi_\infty(\mathbf{1}_{\cup A_n}) = \lim_{n \rightarrow \infty} \pi_\infty(\mathbf{1}_{A_n}) = \lim_{n \rightarrow \infty} \mathbf{1}_{A_n} = \mathbf{1}_{\cup A_n}.$$

Par conséquent,  $\mathcal{T}'$  contient la  $\sigma$ -algèbre engendrée par les  $\mathcal{T}_n$ . Elle coïncide avec  $\mathcal{T}$ , et la projection  $\pi_\infty$  est égale à l'identité.

Le théorème de convergence des martingales a de nombreuses conséquences en théorie des probabilités. La plus simple est sans doute la loi du 0-1 de Kolmogorov, qui permet d'obtenir l'ergodicité du décalage sur un espace produit. Enfin, le théorème de convergence des martingales  $L^2$  est en général faux si la mesure est de masse infinie. Sauriez-vous construire un contre-exemple ?

## Un exemple de marche aléatoire

Certains problèmes de la vie courante peuvent être modélisés par des marches aléatoires. On montre comment, par des raisonnements élémentaires, on peut étudier le comportement d'une de ces marches.

### 1 Le problème du collectionneur

*Nicolas, 10 ans, se lance dans une collection de cartes Pokémon. Chaque carte représente un Pokémon, et il existe 150 Pokémon différents. Les cartes sont vendues dans des paquets scellés, si bien que lorsque Nicolas achète une carte, il peut tomber sur n'importe quel Pokémon ; en particulier sur un Pokémon qu'il possède déjà.*

*En supposant qu'à chaque achat, on ait même probabilité d'obtenir chacun des 150 Pokémon, quel est le nombre moyen de cartes que Nicolas doit acheter afin d'obtenir tous les Pokémon ?*

Un raisonnement simple permet d'obtenir la réponse. Traitons le cas général où le nombre total de Pokémon est quelconque, égal à  $n$ . Notons  $t_k$  le nombre moyen de cartes à se procurer pour terminer la collection, sachant que l'on est déjà en possession de  $k$  Pokémon distincts.

Supposons que nous avons  $k$  Pokémon, tous différents, et achetons une carte supplémentaire. Avec probabilité  $p_k = k/n$ , cette carte était déjà en notre possession ; il faut donc encore acheter en moyenne  $t_k$  cartes pour terminer la collection. Avec probabilité  $1 - p_k$ , c'est une nouvelle carte ; il faut donc acheter en moyenne  $t_{k+1}$  cartes supplémentaires. Par conséquent,

$$t_k = p_k t_k + (1 - p_k) t_{k+1} + 1$$

ce qui entraîne  $t_k = t_{k+1} + \frac{1}{(1 - p_k)}$ , soit  $t_k = t_n + \sum_{l=k}^{n-1} \frac{n}{n-l}$ .

Le nombre de cartes nécessaires pour terminer la collection, sachant que l'on possède déjà  $n$  Pokémon distincts, est égal à zéro :  $t_n = 0$ . Le nombre moyen total de cartes à se procurer pour posséder l'ensemble des  $n$  Pokémon est égal à  $t_0$ , avec :

$$t_0 = n \sum_{j=1}^n \frac{1}{j}$$

Une comparaison série-intégrale, appliquée à la fonction  $1/x$ , donne l'encadrement  $\ln(n) < \sum_{j=1}^n 1/j < \ln(n) + 1$ . La différence  $\sum 1/j - \ln(n)$  est décroissante ; elle converge donc vers une certaine constante  $\gamma$ , appelée constante d'Euler. Si bien que le nombre recherché est de l'ordre de  $n(\ln(n) + \gamma)$ . En première approximation, on peut prendre  $\gamma = 0.577$ . Pour  $n = 150$ , on obtient  $t_0 \simeq 838$ . Nicolas peut espérer terminer sa collection après avoir acheté 838 cartes.

## 2 Interprétation en terme de marche aléatoire

Cherchons à formaliser les raisonnements précédents. Pour cela, il faut se donner un univers, composé des résultats qui peuvent être obtenus à l'issue de l'épreuve, d'une probabilité sur cet ensemble, et d'une variable aléatoire qui représente la quantité à étudier.

### L'univers $\Omega$

Dans chacun des deux exemples, on peut modéliser la situation par une suite de nombres  $X_0, X_1, \dots$  compris entre 0 et  $n$ . Dans le premier exemple, le terme de rang  $i$  de la suite, noté  $X_i$ , correspond au nombre de Pokémon différents en possession de Nicolas, après l'achat de  $i$  cartes.

L'univers  $\Omega$  de tous les résultats possibles correspond donc à l'ensemble des suites indexées par les entiers, à valeurs dans  $\{0, \dots, n\}$ .

$$\Omega = \{0, \dots, n\}^{\mathbf{N}}.$$

La variable aléatoire  $X_i : \Omega \rightarrow \mathbf{R}$  est donnée par la projection sur la  $i^{\text{ème}}$  coordonnée.

### La probabilité $P$

Cherchons à définir sur  $\Omega$  une probabilité  $P$  rendant compte du phénomène étudié. Rappelons que le cylindre  $[X_0 = k_0, X_1 = k_1, \dots, X_l = k_l]$  est le sous-ensemble de  $\Omega$  composé des suites qui débutent par  $k_0, \dots, k_l$ . Pour définir une probabilité  $P$  sur  $\Omega$ , il suffit de spécifier sa valeur sur les cylindres  $[X_0 = k_0, X_1 = k_1, \dots, X_l = k_l]$ , et ce pour tout  $l$  et tout  $l$ -uplet  $k_0, \dots, k_l$ .

Nous allons définir la probabilité  $P$  par une récurrence sur la taille des cylindres. Pour cela, rappelons la notion de *probabilité conditionnelle*. Soit  $A, B$  deux sous-ensembles de  $\Omega$ , et  $P$  une probabilité sur  $\Omega$ . La probabilité de  $A$  sachant  $B$  est donnée par

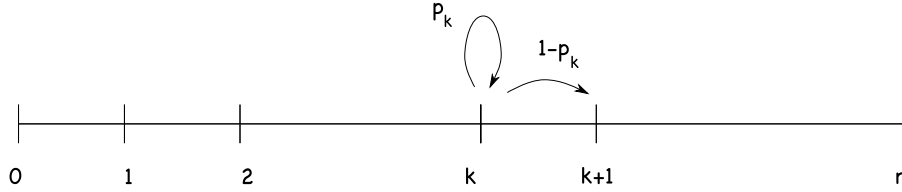
$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Prenons pour  $A$  l'événement  $(X_{i+1} = k_{i+1})$  et pour  $B$  l'événement  $[X_0 = k_0, X_1 = k_1, \dots, X_i = k_i]$ . L'intersection de  $A$  et de  $B$  correspond à l'événement  $[X_0 = k_0, X_1 = k_1, \dots, X_{i+1} = k_{i+1}]$ , si bien qu'il suffit de spécifier les valeurs  $P(X_{i+1} = k_{i+1} | X_0 = k_0, X_1 = k_1, \dots, X_i = k_i)$  pour déterminer  $P$ .

Dans nos exemples, la valeur prise par  $X_{i+1}$  ne dépend pas des valeurs prises par  $X_0, \dots, X_{i-1}$ , mais juste de la valeur prise par  $X_i$  (C'est la *propriété de Markov*). En effet, si la valeur de  $X_i$  est connue, disons égale à  $k$ , la variable aléatoire  $X_{i+1}$  ne peut prendre que les deux valeurs  $k$  et  $k+1$ , et les probabilités associées à ces deux valeurs sont complètement déterminées.

$$\begin{aligned} P(X_{i+1} = l | X_i = k) &= p_k && \text{si } l = k \\ &= 1 - p_k && \text{si } l = k + 1 \\ &= 0 && \text{sinon} \end{aligned}$$

avec  $p_k = \frac{k}{n}$ .



Nous décidons donc de munir  $\Omega$  de la probabilité  $P$  définie par les relations précédentes, car nous pensons qu'elle correspond aux situations que nous sommes en train d'étudier.

Remarquons que la probabilité  $P$  n'est pas entièrement déterminée par les relations données plus haut ; les quantités  $P(X_0 = k_0)$  n'ont pas été spécifiées. Elles n'interviendront pas dans la suite.

### La variable aléatoire $T$

On s'intéresse au temps  $T$  nécessaire pour atteindre la valeur  $n$ . Cette quantité est une variable aléatoire définie sur  $\Omega$  qui peut s'exprimer en fonction des  $X_i$ .

$$T = \text{Card}\{i \in \mathbf{N} \mid X_i < n\} = \sum_{i=0}^{\infty} \mathbf{1}_{\{X_i < n\}}.$$

Rappelons maintenant comment est défini le concept d'*espérance conditionnelle*. Soit  $T$  une variable aléatoire définie sur  $\Omega$ , prenant des valeurs entières, et soit  $A$  un sous-ensemble de  $\Omega$ . L'espérance conditionnelle de  $T$  sachant  $A$  est donnée par une des deux expressions équivalentes suivantes.

$$E(T|A) = \frac{E(T\mathbf{1}_A)}{P(A)} = \sum_{i=0}^{\infty} iP(T = i|A).$$

La quantité que nous cherchons à calculer correspond au temps moyen nécessaire pour atteindre la valeur  $n$ , sachant que nous sommes parti de la valeur 0. Il s'agit donc de l'espérance de  $T$ , sachant que  $X_0 = 0$ .

$$t_0 = E(T|X_0 = 0).$$

### 3 Calcul des $t_k$

On peut maintenant définir précisément les quantités  $t_k$  et dériver la relation de récurrence qui permet de les calculer. Les quantités  $t_k$  correspondent au temps moyen nécessaire pour atteindre la valeur  $n$ , sachant qu'on part de la valeur  $k$ . On a donc

$$t_k = E(T|X_0 = k).$$

Le résultat suivant se déduit de la définition de l'espérance conditionnelle.

#### **Théorème**

Soit  $T$  une variable aléatoire définie sur  $\Omega$ ,  $A$  un sous-ensemble de  $\Omega$ . Soit  $B_0, \dots, B_n$  une partition de  $\Omega$  : les  $B_i$  sont des sous-ensembles de  $\Omega$  disjoints deux à deux, et l'union de tous ces sous-ensembles est égale à  $\Omega$ . On a alors la relation suivante.

$$E(T|A) = \sum_l E(T|B_l \cap A) P(B_l|A).$$

Soit  $k$  un entier strictement inférieur à  $n$ , considérons les événements

$$A = (X_0 = k), \quad B_l = (X_1 = l).$$

Posons  $T' = \sum_{i \geq 1} \mathbf{1}_{\{X_i < n\}}$  si bien que  $T = \mathbf{1}_{\{X_0 < n\}} + T'$ . La relation précédente devient

$$\begin{aligned} E(T|X_0 = k) &= \sum_l E(T|X_1 = l, X_0 = k) P(X_1 = l|X_0 = k) \\ &= \sum_l E(\mathbf{1}_{\{X_0 < n\}} + T'|X_1 = l, X_0 = k) P(X_1 = l|X_0 = k) \\ &= \sum_l E(T'|X_1 = l, X_0 = k) P(X_1 = l|X_0 = k) + 1 \end{aligned}$$

La quantité  $T'$  ne dépend que des variables  $X_1, X_2, \dots$  mais pas de la variable  $X_0$ . Par conséquent, on a l'égalité  $E(T'|X_1 = l, X_0 = k) = E(T'|X_1 = l)$ , ce qui donne

$$E(T|X_0 = k) = \sum_l E(T'|X_1 = l) P(X_1 = l|X_0 = k) + 1.$$

La quantité  $E(T'|X_1 = l)$  correspond au temps moyen nécessaire pour atteindre  $n$ , sachant qu'on est parti de  $l$ . Elle est donc égale à  $t_l$  (Il s'agit du caractère *stationnaire* de la marche aléatoire). L'équation précédente devient

$$t_k = p_k t_k + (1 - p_k) t_{k+1} + 1.$$

C'est la relation attendue.

## La loi forte des grands nombres $L^2$

Voici une preuve de la loi forte des grands nombres, pour des variables aléatoires dans  $L^2$ .

### Théorème

Soit  $(\Omega, \mathcal{T}, \mu)$  un espace mesuré, soient  $(X_i)_{i \in \mathbf{N}}$  des fonctions intégrables définies de  $\Omega$  dans  $\mathbf{R}$ , de carrés intégrables. On suppose que pour tout  $i, j$ ,  $i \neq j$ ,  $\int_{\Omega} X_i d\mu = 0$ ,  $\int_{\Omega} X_i^2 d\mu = 1$  et  $\int_{\Omega} X_i X_j d\mu = 0$ . Alors

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow +\infty]{} 0 \quad p.s.$$

### Preuve

• Par l'inégalité de Markov,

$$\mu\left(\left|\frac{1}{m} \sum_{i=1}^m X_i\right| > \varepsilon\right) \leq \frac{1}{m^2 \varepsilon^2} E\left(\left(\sum_{i=1}^m X_i\right)^2\right)$$

$$E\left(\left(\sum_{i=1}^m X_i\right)^2\right) = E\left(\sum_{i,j} X_i X_j\right) = E\left(\sum_{i=1}^m X_i^2\right) + 2E\left(\sum_{i < j} X_i X_j\right) = m$$

$$\text{donc } \mu\left(\left|\frac{1}{m} \sum_{i=1}^m X_i\right| > \varepsilon\right) \leq \frac{1}{m \varepsilon^2}$$

• On veut appliquer le lemme de Borel-Cantelli, mais  $\sum \frac{1}{m} = +\infty$ . Remplaçons  $m$  par  $m^2$ .

$$\mu\left(\left|\frac{1}{m^2} \sum_{i=1}^{m^2} X_i\right| > \varepsilon\right) \leq \frac{1}{m^2 \varepsilon^2}.$$

Comme  $\sum \frac{1}{m^2} < +\infty$ , le lemme de Borel-Cantelli montre que

presque partout,  $\forall \varepsilon > 0, \exists M \in \mathbf{N}, \forall m \geq M, \left|\frac{1}{m^2} \sum_{i=1}^{m^2} X_i\right| < \varepsilon$ .

On en déduit, en prenant  $\varepsilon = 1/K, K \in \mathbf{N}^*$ ,

$$\frac{1}{m^2} \sum_{i=1}^{m^2} X_i \xrightarrow[m \rightarrow +\infty]{} 0 \quad p.p.$$

• Soit  $n \in \mathbf{N}$ . Soit  $m$  la partie entière de  $\sqrt{n}$ . On a :  $m^2 \leq n \leq (m+1)^2 - 1$   
On utilise à nouveau le lemme de Borel-Cantelli.

$$\mu\left(\frac{1}{n} \left|\sum_{i=m^2+1}^n X_i\right| > \varepsilon\right) \leq \frac{1}{n^2 \varepsilon^2} E\left(\sum_{i=1}^n X_i^2\right) \leq \frac{1}{n^2 \varepsilon^2} (n - m^2) \leq \frac{2}{n^{3/2} \varepsilon^2}$$

Comme  $\sum \frac{1}{n^{3/2}} < +\infty$ , on en déduit que

$$\frac{1}{n} \sum_{i=m^2+1}^n X_i \xrightarrow[n \rightarrow +\infty]{} 0 \quad p.p.$$

Ceci termine la preuve du théorème.

On a utilisé le lemme de Borel-Cantelli ; rappelons son énoncé et sa preuve.

**Lemme de Borel-Cantelli**, partie facile

Soit  $(\Omega, \mathcal{T}, \mu)$  un espace mesuré,  $A_i, i \in \mathbf{N}$  des sous-ensembles mesurables de  $\Omega$  tels que  $\sum_{i \in \mathbf{N}} \mu(A_i) < \infty$ . Alors, presque tout  $\omega \in \Omega$  n'appartient qu'à un nombre fini d'ensembles  $A_i$ .

**Preuve**

Il suffit d'intégrer l'égalité  $\text{Card}\{i \in \mathbf{N} \mid \omega \in A_i\} = \sum_{i=0}^{\infty} \mathbf{1}_{A_i}(\omega)$ , et de remarquer qu'une fonction intégrable est finie presque partout.

**Remarques**

- L'hypothèse  $\int X_i d\mu = 0$  est une hypothèse de centrage.
- On n'a pas eu besoin de supposer que  $\mu$  est une mesure de probabilité. Cette hypothèse n'est utile que si les variables aléatoires  $X_i$  ne sont pas centrées.
- Si les  $X_i$  sont centrées et que la suite des  $(X_i)_{i \in \mathbf{N}}$  est stationnaire, on peut remplacer les hypothèses sur les  $X_i$  par une condition dite de "corrélations sommables" :

$$\text{La Série } \sum_i \int X_1 X_i d\mu \text{ est absolument convergente.}$$

- Lorsque  $(\Omega, \mathcal{T}, \mu)$  est un espace probablisé, et que les  $X_i$  ne sont pas centrés, on peut remplacer les hypothèses sur les  $X_i$  par une hypothèse de "décorrélation" :

$$\int X_i X_j d\mu - \int X_i d\mu \int X_j d\mu = 0.$$

La moyenne des  $X_i$  converge alors presque sûrement, dès que la moyenne des  $E(X_i)$  converge, et ces deux moyennes sont égales. C'est l'énoncé classique de la loi des grands nombres pour des variables aléatoires  $L^2$ .

- Lorsque  $(\Omega, \mathcal{T}, \mu)$  est un espace probablisé, et que les  $X_i$  sont des fonctions bornées par une constante  $C$ , on peut simplifier la seconde partie de la preuve en utilisant la majoration  $\left| \sum_{i=m^2+1}^n X_i \right| \leq C(n - m^2) \leq 2C\sqrt{n}$

Voici un corollaire de la loi forte des grands nombres, qui s'obtient en considérant des fonctions de la forme  $\mathbf{1}_A(X_i)$ .

**Corollaire**

Soient  $(\Omega, \mathcal{T}, \mu)$  un espace probablisé,  $(X_i)_{i \in \mathbf{N}}$  des variables aléatoires indépendantes équidistribuées, et  $A \subset \mathbf{R}$  un borélien. Alors, pour presque tout  $\omega \in \Omega$ ,

$$\frac{1}{n} \text{Card} \left\{ i \in \{1, 2, \dots, n\} \mid X_i(\omega) \in A \right\} \xrightarrow{n \rightarrow +\infty} P(X_1 \in A).$$

Remarquons pour terminer, qu'il est possible d'obtenir une vitesse dans la loi des grands nombres lorsque les variables aléatoires sont dans  $L^2$ . Il suffit de remplacer  $\varepsilon$  par  $n^{-0,2}\varepsilon$  dans la preuve du théorème pour obtenir

$$\frac{1}{n^{0,8}} \sum_{i=1}^n X_i \longrightarrow 0 \quad p.p.$$

La normalisation optimale est en fait en  $\frac{1}{n^{1/2+\varepsilon}}$  ; c'est une application du théorème des trois séries, mais cela ne semble pas pouvoir se déduire de la preuve donnée plus haut.

## Le théorème de la limite centrée

Voici une preuve du théorème de la limite centrée pour des variables de Bernoulli.

**Théorème** Soit  $(X_i)_{i \in \mathbf{N}}$  une suite de variables aléatoires indépendantes, suivant toutes une loi de Bernoulli de paramètre  $p \in ]0, 1[$ .

$P(X_i = 1) = p$ ,  $P(X_i = 0) = 1 - p$  ;  $E(X_i) = p$ ,  $V(X_i) = p(1 - p) = \sigma^2$ .

On pose  $S_n = \sum_{i=1}^n X_i$ . Alors

$$P\left(\frac{S_n}{n} \in \left[E(X) + \frac{a}{\sqrt{n}}, E(X) + \frac{b}{\sqrt{n}}\right]\right) \xrightarrow{n \rightarrow +\infty} \frac{1}{\sqrt{2\pi}\sigma} \int_a^b e^{-\frac{x^2}{2\sigma^2}} dx$$

### Preuve

La variable aléatoire  $S_n$  suit une loi binomiale de paramètres  $n$  et  $p$ . La probabilité qu'il s'agit d'évaluer vaut donc  $\sum_l C_n^l p^l (1-p)^{n-l}$ , avec un indice  $l$  entier variant entre  $np + a\sigma\sqrt{n}$  et  $np + b\sigma\sqrt{n}$ . On traite le cas  $a = 0$ .

**Lemme** Soit  $k$  un réel dépendant de  $n$ , tel que  $np + k \in \mathbf{N}$  et  $0 \leq k \leq b\sqrt{n}$ .

$$C_n^{np+k} p^{np+k} (1-p)^{n(1-p)-k} \sim \frac{1}{\sqrt{2\pi n} \sigma} e^{-\frac{k^2}{2n\sigma^2}}$$

### Preuve du lemme

On utilise la formule de Stirling :  $n! \sim n^n e^{-n} \sqrt{2\pi n}$ .

$$\begin{aligned} C_n^{np+k} &= \frac{n!}{(np+k)! (n(1-p)-k)!} \\ &\sim \frac{n^n}{(np+k)^{np+k} (n(1-p)-k)^{n(1-p)-k}} \frac{\sqrt{n}}{\sqrt{2\pi} \sqrt{np+k} \sqrt{n(1-p)-k}} \end{aligned}$$

$$\begin{aligned} C_n^{np+k} &\sim \left( \left( p + \frac{k}{n} \right)^{np+k} \left( 1 - p - \frac{k}{n} \right)^{n(1-p)-k} \sqrt{2\pi n \sigma} \right)^{-1} \left( \left( 1 + \frac{k}{pn} \right) \left( 1 + \frac{k}{(1-p)n} \right) \right)^{-\frac{1}{2}} \\ &\sim \left( p^{np+k} (1-p)^{n(1-p)-k} \left( 1 + \frac{k}{pn} \right)^{np+k} \left( 1 - \frac{k}{(1-p)n} \right)^{n(1-p)-k} \sqrt{2\pi n \sigma} \right)^{-1} \end{aligned}$$

Il s'agit maintenant d'étudier l'expression

$$\left( 1 + \frac{k}{pn} \right)^{np+k} \left( 1 - \frac{k}{(1-p)n} \right)^{n(1-p)-k}$$

Pour cela, on passe au logarithme et on effectue un développement à l'ordre 2.

$$\begin{aligned} &(np+k) \ln\left(1 + \frac{k}{pn}\right) + (n(1-p)-k) \ln\left(1 - \frac{k}{(1-p)n}\right) \\ &= (np+k) \left( \frac{k}{pn} - \frac{k^2}{2p^2 n^2} \right) + (n(1-p)-k) \left( -\frac{k}{(1-p)n} - \frac{k^2}{2(1-p)^2 n^2} \right) + o\left(\frac{k^3}{n^2}\right) \\ &= \frac{k^2}{n} \left( \frac{1}{p} + \frac{1}{1-p} - \frac{1}{2p} - \frac{1}{2(1-p)} \right) + o\left(\frac{k^3}{n^2}\right) = \frac{k^2}{2p(1-p)n} + o\left(\frac{k^3}{n^2}\right) \quad CQFD \end{aligned}$$

Soit  $t \in [0, b]$  ; prenons  $k$  tel que  $np+k$  soit égal à la partie entière de  $np + \sqrt{nt}$  et posons  $f_n(t) = \sqrt{n} C_n^{np+k} p^{np+k} (1-p)^{n(1-p)-k}$ . Les  $f_n(t)$  sont des fonctions en escalier décroissantes sur  $[0, b]$ . D'après le lemme, elles convergent vers la fonction  $f(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{t^2}{2\sigma^2}\right)$ . On veut montrer que

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{b\sqrt{n}} f_n\left(\frac{k}{\sqrt{n}}\right) - \int_0^b f(x) dx \xrightarrow{n \rightarrow +\infty} 0$$

Le lemme de Dini permet d'affirmer que la suite  $f_n$  converge uniformément vers  $f$  sur l'intervalle  $[0, b]$ . La différence  $\frac{1}{\sqrt{n}} \sum_k f_n\left(\frac{k}{\sqrt{n}}\right) - f\left(\frac{k}{\sqrt{n}}\right)$  est majorée par  $\|f - f_n\|_\infty$  ; elle tend donc vers 0. Enfin, la série  $\frac{1}{\sqrt{n}} \sum_{k=1}^{b\sqrt{n}} f\left(\frac{k}{\sqrt{n}}\right)$  est une somme de Riemann associée à la fonction  $f$  ; elle converge donc vers  $\int_0^b f(x) dx$ .



### Remarques

– Le caractère décroissant de  $f_n$  provient de la décroissance de la suite  $P_m = C_n^m p^m (1-p)^{n-m}$  lorsque  $m > np$ . Pour démontrer cette propriété classique de la loi binomiale, il suffit de calculer le quotient  $P_{m+1}/P_m$ . On obtient  $\frac{P_{m+1}}{P_m} = (\frac{n+1}{m+1} - 1) \frac{p}{1-p}$ , quantité qui est inférieure à 1 dès que  $m+1 \geq p(n+1)$ .

– On peut se passer du lemme de Dini. Pour cela, on commence par montrer que  $\int f_n \rightarrow \int f$  à l'aide du théorème de convergence dominée. Pour obtenir la domination, on remarque que la suite  $(f_n(0))_{n \in \mathbf{N}}$  est convergente ; l'encadrement  $0 \leq f_n(x) \leq f_n(0)$  montre que la suite  $f_n$  est uniformément bornée sur  $[0, b]$ .

Il faut ensuite montrer que  $\frac{1}{\sqrt{n}} \sum_{k=0}^{b\sqrt{n}} f_n(\frac{k}{\sqrt{n}}) - \int_0^b f_n(x) dx \xrightarrow{n \rightarrow +\infty} 0$ .

Comme  $f_n$  est décroissante, on a  $\frac{1}{\sqrt{n}} f_n(\frac{k+1}{\sqrt{n}}) \leq \int_{\frac{k}{\sqrt{n}}}^{\frac{k+1}{\sqrt{n}}} f_n(x) dx \leq \frac{1}{\sqrt{n}} f_n(\frac{k}{\sqrt{n}})$ .

Il s'agit donc de majorer la somme  $\frac{1}{\sqrt{n}} \sum_{k=1}^{b\sqrt{n}} f_n(\frac{k}{\sqrt{n}}) - f_n(\frac{k+1}{\sqrt{n}})$ .

Cette somme vaut  $\frac{1}{\sqrt{n}}(f_n(\frac{1}{\sqrt{n}}) - f_n(b))$  car les termes de la série se simplifient deux à deux. Elle est donc majorée par  $\frac{1}{\sqrt{n}} f_n(0)$ , qui converge vers 0.

Le corollaire suivant est très utile en pratique.

### Corollaire

On considère une épreuve aléatoire modélisée par un espace probabilisé  $(\Omega, \mathcal{T}, P)$ . Un expérimentateur cherche à évaluer la probabilité  $P(A)$  associée à un événement  $A \in \mathcal{T}$ .

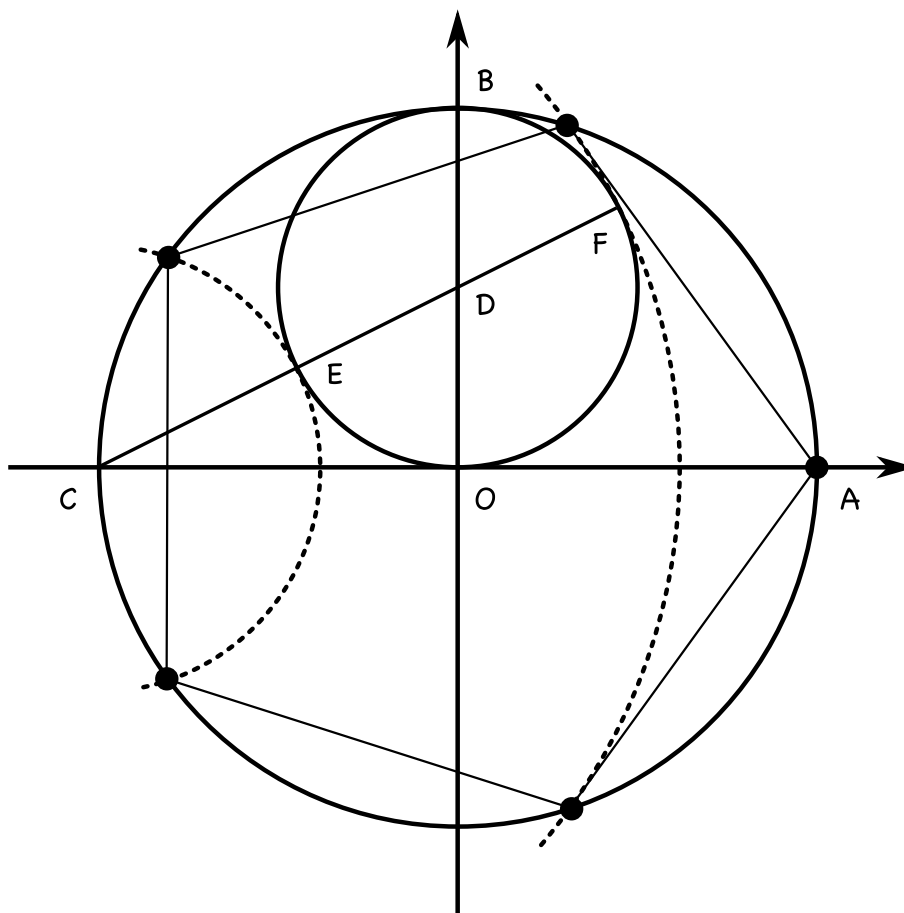
Pour cela, il répète  $n$  fois l'épreuve considérée, de manière indépendante. Considérons le nombre moyen de fois où l'événement  $A$  s'est trouvé réalisé au cours de ces  $n$  épreuves. On s'attend à ce que ce nombre soit proche de la probabilité  $P(A)$ . Le théorème de la limite centrée affirme que ce nombre appartient à l'intervalle  $[P(A) - \frac{t\sigma}{\sqrt{n}}, P(A) + \frac{t\sigma}{\sqrt{n}}]$  avec probabilité  $\frac{1}{\sqrt{2\pi}} \int_{-t}^t e^{-x^2/2} dx$ .

L'écart-type  $\sigma$  vaut  $\sqrt{P(A)(1-P(A))}$  et n'est à priori pas connu de l'expérimentateur. Il est tout de même majoré par  $\frac{1}{2}$ , ce qui permet de donner un intervalle de confiance pour  $P(A)$ , qui dépend du nombre  $n$  d'épreuves effectuées, et du risque d'erreur  $\alpha = 1 - \frac{1}{\sqrt{2\pi}} \int_{-t}^t e^{-x^2/2} dx$  qu'on veut bien tolérer.

Par exemple, un risque  $\alpha = 0,01$  donne un intervalle de confiance pour  $P(A)$  égal à  $[f_n - \frac{1,29}{\sqrt{n}}, f_n + \frac{1,29}{\sqrt{n}}]$ , où  $f_n$  est la fréquence observée de l'événement  $A$  lors des  $n$  épreuves.

## Construction du pentagone régulier

Voici une construction du pentagone régulier à la règle et au compas. Soit  $A$ ,  $B$  et  $C$  les points de coordonnées  $(1, 0)$ ,  $(0, 1)$  et  $(-1, 0)$ . On trace le cercle de diamètre  $OB$  puis la droite reliant le centre  $D$  de ce cercle au point  $C$ . Soit  $E$  et  $F$  les deux points d'intersection de ce cercle et de cette droite. Il suffit maintenant de tracer les cercles de centre  $C$  et passant par  $E$  et  $F$ , pour obtenir quatre des sommets du pentagone.



On renvoie à l'article “Construction-a-la-regle-et-au-compas” de l’encyclopédie *Wikipedia* pour d’autres constructions classiques en géométrie élémentaire.

## Polyèdres platoniciens

On s'intéresse aux polyèdres convexes (enveloppe convexe d'un nombre fini de points de  $\mathbf{R}^3$ ) qui satisfont les propriétés suivantes :

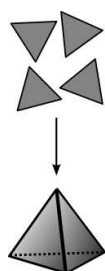
- Les faces sont des polygones réguliers à  $n$  côtés. Toutes les arêtes ont même longueur.
- Le nombre d'arêtes en chaque sommet du polyèdre ne dépend pas du sommet considéré. Notons  $k$  ce nombre d'arêtes.

On va montrer qu'il n'y a qu'un nombre fini de possibilités pour  $n$  et  $k$ , ce qui donne un nombre fini de possibilités pour le nombre  $\mathbf{S}$  de sommets, le nombre  $\mathbf{A}$  d'arêtes et le nombre  $\mathbf{F}$  de faces d'un tel polyèdre.

Calculons les valeurs possibles pour  $n, k, S, A$  et  $F$ .

Comme le polyèdre est convexe, on a *la formule d'Euler*,  $\mathbf{S} - \mathbf{A} + \mathbf{F} = 2$ .

Soit  $\tilde{S}, \tilde{A}, \tilde{F}$  le nombre total de sommets, d'arêtes et de faces avant assemblage.



$$\begin{aligned}\tilde{S} &= 12 \\ \tilde{A} &= 12 \\ \tilde{F} &= 4\end{aligned}$$

On a les relations

$$\tilde{S} = \tilde{A} = n\tilde{F}$$

$$\begin{aligned}S &= 4 \\ A &= 6 \\ F &= 4\end{aligned}$$

$$\begin{aligned}S &= \tilde{S}/k \\ A &= \tilde{A}/2 \\ F &= \tilde{F}\end{aligned}$$

Ce qui donne  $\left(n\left(\frac{1}{k} - \frac{1}{2}\right) + 1\right)F = 2$ .

On doit donc avoir  $1 - n\left(\frac{1}{2} - \frac{1}{k}\right) > 0$ , c'est-à-dire  $1 > n\left(\frac{1}{2} - \frac{1}{k}\right)$ .

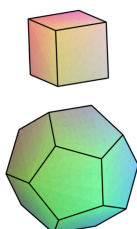
Comme  $k \geq 3$ , on a :  $\frac{1}{2} - \frac{1}{k} \geq 1/6$ , d'où l'encadrement  $1 > n\left(\frac{1}{2} - \frac{1}{k}\right) \geq n/6$ .

Par conséquent,  $3 \leq n < 6$ .

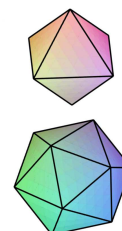
La relation  $\frac{1}{n} > \frac{1}{2} - \frac{1}{k}$  implique  $\frac{2n}{n-2} > k$ , ce qui laisse comme possibilités

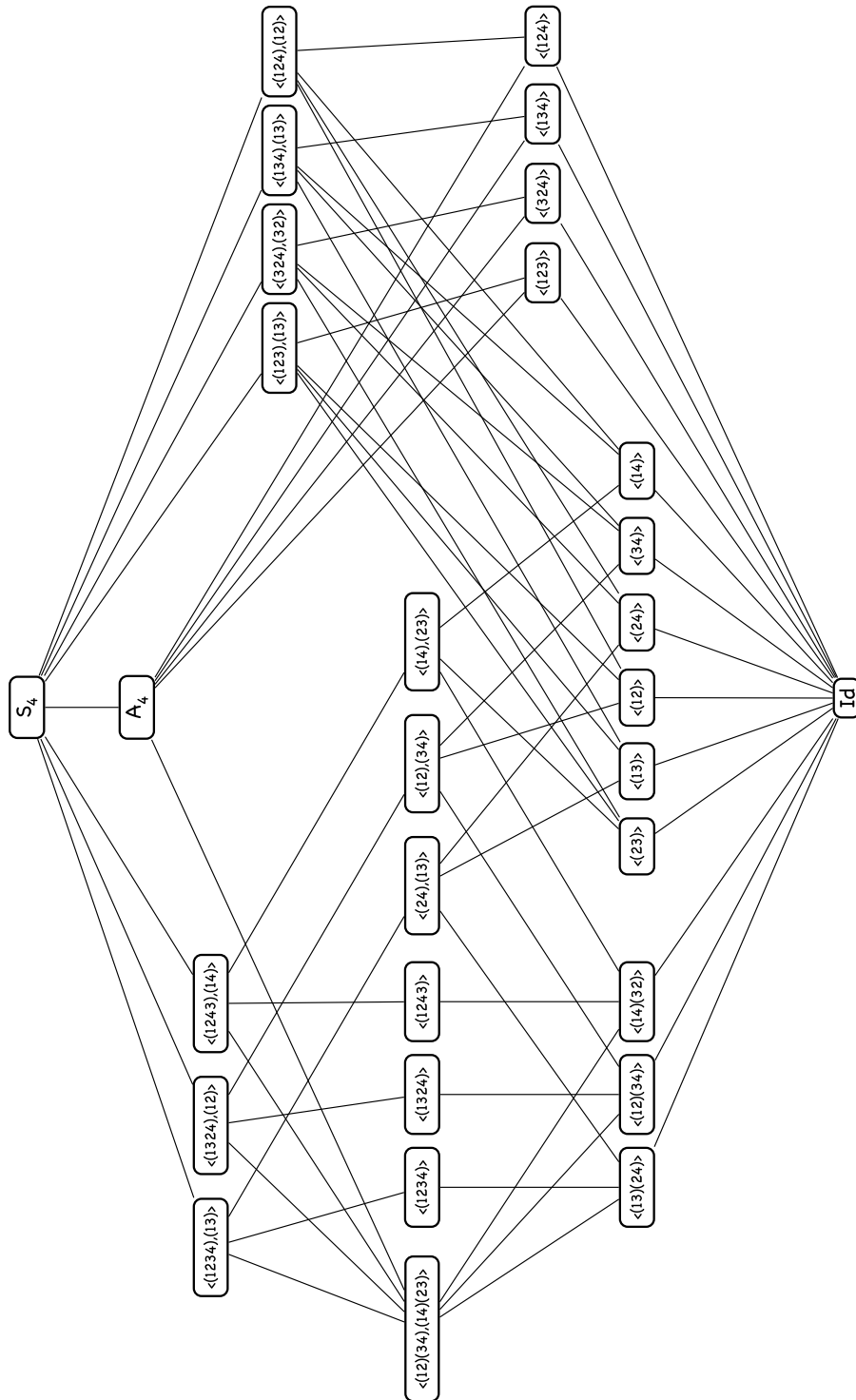
$$n = 4, k = 3 ; n = 5, k = 3 ; n = 3, k = 3, 4, 5.$$

De là, on en déduit les valeurs de  $S, A, F$ .  $F = \frac{2}{1 - (\frac{1}{2} - \frac{1}{k})n}$ ,  $S = \frac{n}{k} F$ ,  $A = \frac{n}{2} F$ .



n	k	S	A	F	
3	3	4	6	4	tétraèdre
3	4	6	12	8	octaèdre
3	5	12	30	20	icosaèdre
4	3	8	12	6	cube
5	3	20	18	12	dodécaèdre



Sous-groupes de  $S_4$ 

## Trace, formes quadratiques et extensions de corps

Ce document porte sur les notions de dimension d'espace vectoriel, extensions de corps, trace de matrices, polynôme minimal et caractéristique, théorème de Cayley-Hamilton, formes quadratiques, nombres irrationnels, déterminant de Vandermonde, déterminant de Gram.

On passe un certain temps en premier cycle à démontrer que tout espace vectoriel de dimension finie admet une base ; de même, on démontre que toute forme quadratique admet une base orthogonale. Cependant, à ce niveau du cours, tous les espaces vectoriels de dimension finie considérés arrivent avec une base naturelle, et toutes les formes quadratiques utilisées sont données dans une base orthogonale.

Au niveau du programme de l'agrégation, il existe essentiellement deux exemples d'espaces vectoriels de dimension finie pour lesquels il n'y a pas de base donnée à priori :

- Le premier est formé par les solutions d'une équation différentielle sur un intervalle borné. Sous des hypothèses ad hoc, le théorème de Cauchy-Lipschitz affirme que l'espace vectoriel des solutions est de dimension finie, mais il peut être difficile d'exhiber une base explicite de solutions. Cet espace vectoriel est muni d'une forme quadratique naturelle,  $(f, g) \rightarrow \int fg$ , qui ne fait référence à aucune base particulière.
- Le second exemple est de nature algébrique. Considérons un corps  $L$  engendré par un nombre fini de nombres algébriques sur  $\mathbf{Q}$ . Ce corps est un espace vectoriel de dimension finie sur  $\mathbf{Q}$ . Les nombres algébriques qui ont permis de le définir ne forment pas une base de  $L$  en général. Il existe une forme quadratique naturelle sur  $L$ , qui est donnée par la trace,  $(x, y) \rightarrow \text{Trace}_{L/\mathbf{Q}}(xy)$ .

Dans ces deux exemples, il peut même être difficile de déterminer la dimension de l'espace vectoriel considéré. On va voir comment l'existence d'une forme quadratique peut aider à résoudre ce problème.

### Définitions

Dans la suite, on s'intéresse au second exemple. Soit donc  $K$  un corps de caractéristique différente de 2,  $L$  une extension de dimension finie de  $K$ . Notons  $n$  la dimension de  $L$  en tant que  $K$ -espace vectoriel. Dans la suite, il peut être utile de faire intervenir une extension algébriquement close de  $K$  contenant  $L$  ; elle sera abusivement notée  $\bar{K}$ . On peut se restreindre à  $K = \mathbf{Q}$  et  $\bar{K} = \mathbf{C}$  si on veut.

Pour tout élément  $y \in L$ , on considère l'application  $K$ -linéaire donnée par

$$A_y : \begin{array}{l} L \rightarrow L \\ x \rightarrow xy \end{array}$$

La *trace* de cette application linéaire est un élément de  $K$  qui est noté  $\text{tr}_{L/K}(y)$ , ou encore s'il n'y a pas d'ambiguïté sur les corps considérés,  $\text{tr}(y)$ . On vérifie immédiatement que l'application  $(x, y) \rightarrow \text{tr}(xy)$  est une application  $K$ -bilinéaire symétrique définie sur  $L$ .

Remarquons que si l'on se donne une base de  $L$  sur  $K$ , l'application qui associe à  $y$  la matrice de  $A_y$  dans cette base, réalise un plongement de  $L$  dans l'algèbre  $M_n(K)$ . Bien sûr, ce plongement n'est pas surjectif.

Voici une propriété des applications de la forme  $A_y$ , qui ne sont pas vraies pour toutes les applications  $K$ -linéaires

**Lemme** Le polynôme minimal  $P_m$  de  $A_y$  est irréductible. Le polynôme caractéristique  $P_c$  de  $A_y$  est égal à une puissance de son polynôme minimal.

### Preuve

Commençons par montrer que le polynôme minimal de  $A_y$  est irréductible sur  $K$ . Pour cela, rappelons que le polynôme minimal de  $y$  est le plus petit polynôme non nul  $P$  de  $K[X]$  (au sens de la division) qui vérifie  $P(y) = 0$ . Ce polynôme est irréductible. S'il était le produit de deux polynômes non constants de  $K[X]$ ,  $y$  serait racine d'un de ces deux polynômes, et donc  $P$  ne serait pas minimal. Maintenant le polynôme minimal de  $A_y$  est égal au polynôme minimal de  $y$  en vertu des deux relations  $P(A_y) = A_{P(y)}$  et  $P(y) = P(A_y)1$ .

L'anneau  $K[X]$  étant factoriel, on peut décomposer  $P_c$  sous la forme d'un produit  $P_c = (P_m)^l H$ , avec  $l$  un entier et  $H$  un polynôme premier à  $P_m$ . Remarquons que  $H$  n'a pas de racines en commun avec  $P_m$ ; cela découle, par exemple, du théorème de Bezout. Si  $H$  est non constant,  $P_c$  aurait une racine dans  $\bar{K}$  qui ne serait pas racine de  $P_m$ . C'est absurde car les racines de  $P_c$  sont les valeurs propres de  $A_y$  (dans  $\bar{K}$ ) et ces valeurs propres sont toutes racines du polynôme minimal  $P_m$ .

Voici une conséquence de ce lemme. Si  $y$  est non nul,  $A_y$  ne peut pas être nilpotente. En effet, s'il existe un entier  $k$  tel que  $(A_y)^k = 0$ , le polynôme minimal de  $A_y$  divise  $X^k$ ; comme il est irréductible, il est égal à  $X$ , donc  $A_y = 0$ .

Autre conséquence : si la caractéristique de  $K$  est nulle,  $A_y$  est diagonalisable sur  $\bar{K}$ . De fait, en caractéristique 0 (plus généralement si le corps est parfait), les polynômes irréductibles ont leurs racines simples. Une matrice dont le polynôme minimal a ses racines simples est diagonalisable (sur  $\bar{K}$ ).

## Indépendance et trace

Voici comment utiliser la trace dans des questions d'indépendance linéaire.

### Théorème

On se donne  $p_1, p_2, \dots, p_k$  des nombres premiers distincts. Alors les nombres réels  $\sqrt{p_1}, \sqrt{p_2}, \dots, \sqrt{p_k}$  sont linéairement indépendants sur  $\mathbf{Q}$ .

### Preuve

Rappelons que si  $q$  est une forme bilinéaire symétrique définie sur un espace vectoriel  $L$ , et si  $v_1, v_2, \dots, v_k$  sont des vecteurs de  $L$ , alors la famille des  $v_i$  est libre si le déterminant de la matrice de terme général  $q(v_i, v_j)$  est non nul (une combinaison linéaire entre les  $v_i$  donne tout de suite une combinaison linéaire sur les colonnes de cette matrice). Un tel déterminant est appelé déterminant de Gram.

Ici, on considère comme  $\mathbf{Q}$ -espace vectoriel le corps engendré par les  $\sqrt{p_i}$ , et comme forme quadratique  $(x, y) \rightarrow \text{Tr}_{L/\mathbf{Q}}(xy)$ . Il faut donc calculer les traces  $\text{tr}_{L/\mathbf{Q}}(\sqrt{p_i p_j})$ .

- Si  $i = j$ , on a  $\text{tr}(p_i) = p_i \text{tr}(1) = np_i$ , où  $n$  est la dimension de  $L$  sur  $\mathbf{Q}$ .
- Si  $i \neq j$ , on remarque que le polynôme minimal de  $\sqrt{p_i p_j}$  est égal à  $X^2 - p_i p_j$ . Son polynôme caractéristique est donc égal à  $(X^2 - p_i p_j)^l$ , pour un certain entier  $l \in \mathbf{N}$ . La trace de  $\sqrt{p_i p_j}$  est égale au coefficient de  $X^{2l-1}$  dans ce polynôme, elle est donc nulle.

Par conséquent, la matrice de terme général  $\text{tr}_{L/\mathbf{Q}}(\sqrt{p_i p_j})$  est diagonale, et ses termes diagonaux sont des entiers non nuls ; son déterminant est donc non nul. Les  $\sqrt{p_1}, \sqrt{p_2}, \dots, \sqrt{p_k}$  sont linéairement indépendants sur  $\mathbf{Q}$ .

On peut calculer la dimension de  $\mathbf{Q}[\sqrt{p_1}, \dots, \sqrt{p_k}]$  ; pour cela, on considère la famille des  $\sqrt{\prod_{i \in I} p_i}$  où  $I$  est une partie quelconque de l'ensemble  $\{1 \dots k\}$ . (on pose  $\prod_{i \in I} p_i = 1$  si  $I$  est vide). Cette famille possède  $2^k$  éléments ; l'espace vectoriel qu'elle engendre coïncide avec  $\mathbf{Q}[\sqrt{p_1}, \dots, \sqrt{p_k}]$ . Il suffit de vérifier que c'est une famille libre, ce qui se fait comme plus haut, en considérant les traces. La dimension recherchée est donc égale à  $2^k$ .

### Non-dégénérescence de la trace

Suffit-il de calculer le déterminant de la matrice de terme général  $\text{tr}(x_i x_j)$  pour savoir si la famille  $\{x_1, \dots, x_k\}$  est libre ? Par exemple, si  $K$  est un sous-corps de  $\mathbf{C}$ , et si la forme quadratique associée à  $q$  est définie positive, La réponse est oui. Ceci provient du fait que la restriction d'une forme quadratique définie positive à un sous-espace vectoriel est encore définie positive.

Si la forme quadratique n'est pas positive, ce n'est plus forcément vrai. Dès l'instant où la forme admet un vecteur isotrope ( $q(x, x) = 0$ ), on obtient un contre-exemple en considérant la famille libre composée de l'unique élément  $x$ , pour laquelle on a  $\det(q(x, x)) = 0$ . Cependant, si la forme quadratique est non-dégénérée, on a tout de même le résultat suivant.

*Pour toute base  $\{x_1, \dots, x_n\}$ , le déterminant  $\det(q(x_i, x_j))$  est non nul.*

Ce résultat n'est pas difficile ; cf par exemple Ramis-Deschamps Tome 2 1.1.2 prop 2 cor I. Cette référence comporte d'autres informations sur les déterminants de la forme  $\det(q(x_i, x_j))$ , appelés déterminants de Gram. Au final, on obtient le critère suivant.

*Soit  $q$  une forme bilinéaire symétrique non-dégénérée. Pour qu'une famille génératrice  $\{x_1 \dots x_n\}$  soit une base, il faut et il suffit que  $\det(q(x_i, x_j))$  soit non nul.*

Dans le cas d'extensions de corps, il est facile de trouver des familles génératrices, si bien que la méthode présentée plus haut permet effectivement de déterminer la dimension de l'extension, si la trace est non dégénérée.

### Théorème

Soit  $K$  un corps de caractéristique 0, et  $L$  une extension finie de  $K$ . Alors  $(x, y) \rightarrow \text{tr}_{L/K}(xy)$  est une forme bilinéaire symétrique non dégénérée.

**Première preuve**

Il s'agit de montrer que pour tout  $y \in L$  non nul, on peut trouver un  $x \in L$  tel que  $tr(xy) \neq 0$ . Prenons  $x = y^{-1}$ . On a alors  $tr(xy) = tr(1) = \dim_K L$ . Comme la caractéristique de  $K$  est non nulle, cet entier est non nul.

**Seconde preuve**

Sous la seule hypothèse de séparabilité, en utilisant le théorème de l'élément primitif : l'extension  $L$  est de la forme  $L \simeq K[\theta]$ , pour un certain  $\theta \in L$ .

Le polynôme minimal de  $A_\theta$  est égal au polynôme minimal  $P_m$  de  $\theta$  ; il est donc irréductible, et ses racines sont simples (dans  $\bar{K}$  ; c'est la séparabilité). Il est de degré  $n$ , en vertu de l'isomorphisme  $K[\theta] \simeq K[X]/P_m$ . Comme il divise le polynôme caractéristique de  $A_\theta$ , il est en fait égal à ce polynôme. On conclut que le polynôme caractéristique de  $A_\theta$  a ses racines simples. Par conséquent les valeurs propres  $\lambda_i$  de  $A_\theta$  sont de multiplicité 1.

Comme  $1, \theta, \theta^2, \dots, \theta^{n-1}$  forme une base de  $K[\theta]$  sur  $K$ , il suffit de montrer que la matrice  $\left( tr(\theta^i \theta^j) \right)_{\substack{0 \leq i \leq n-1 \\ 0 \leq j \leq n-1}}$  est non dégénérée, c'est-à-dire que son déterminant est non nul.

$$\det\left( tr(A_\theta^{i+j}) \right) = \det\left( \sum \lambda_k^{i+j} \right) = \det\left( \sum \lambda_k^i \lambda_k^j \right) = \det({}^t B B) = \det(B)^2$$

où  $B = (\lambda_i^j)_{\substack{0 \leq i \leq n-1 \\ 0 \leq j \leq n-1}}$ .

Le déterminant de  $B$  est un déterminant de Van Der Mond, qui peut être calculé explicitement :  $\det B = \prod_{i>j} (\lambda_i - \lambda_j) \neq 0$ . Ceci termine la preuve.

Voici enfin un exemple d'extension pour laquelle la trace n'est pas définie positive. Soit  $\alpha$  une racine dans  $\mathbf{C}$  de l'équation  $X^4 + 1$  ( $\alpha = e^{i\pi/4}$  par exemple) ; l'extension  $\mathbf{Q}[\alpha]$  est de degré 4 sur  $\mathbf{Q}$ . Le polynôme minimal de  $\alpha^2$  est égal à  $X^2 + 1$ , son polynôme caractéristique est donc égal à  $(X^2 + 1)^2$ . Par conséquent, la trace  $tr_{\mathbf{Q}[\alpha]/\mathbf{Q}}(\alpha^2)$  est nulle. Le nombre  $\alpha$  est un vecteur isotrope de la trace dans l'extension  $\mathbf{Q}[\alpha]$ .

Toutes ces considérations sont classiques en théorie de Galois. Voici deux références pour en savoir plus : Ian Stewart, *Galois Theory* et *Algebraic Theory of Numbers*.



## Continuité des racines des polynômes

Lorsqu'on varie continûment les coefficients d'un polynôme à coefficients réels ou complexes, ses racines se déplacent-elles de manière continue dans le plan complexe ?

Pour les polynômes de degré 2 unitaires à coefficients réels, les formules bien connues suivantes :

$$\lambda = \frac{-b + \sqrt{b^2 - 4c}}{2}, \quad \mu = \frac{-b - \sqrt{b^2 - 4c}}{2},$$

permettent d'exprimer les racines du polynôme  $P(X) = X^2 + bX + c$  en fonction de ses coefficients, dès que  $b^2 - 4c > 0$ . Ces deux fonctions  $\lambda$  et  $\mu$  dépendent de manière continue des coefficients sur l'ensemble des polynômes à coefficients réels dont le discriminant est strictement positif. Malheureusement, il n'y a pas d'extension continue de la fonction racine carrée à l'ensemble du plan complexe. Peut-on malgré tout montrer que les racines d'un polynôme dépendent continûment de ses coefficients ?

### Contre-exemples à la continuité des racines

Soit  $\mathcal{E}_2$  l'ensemble des polynômes unitaires de degré 2 à coefficients complexes ; supposons qu'il existe deux fonctions continues  $\lambda, \mu : \mathcal{E}_2 \rightarrow \mathbf{C}^2$  telles que pour tout  $P \in \mathcal{E}_2$ ,

$$P(X) = (X - \lambda(P))(X - \mu(P)).$$

Intéressons nous à la famille de polynômes suivante, paramétrée par  $\theta \in \mathbf{R}/2\pi\mathbf{Z}$ .

$$P_\theta(X) = X^2 - e^{i\theta}.$$

Posons  $\tilde{\lambda}(\theta) = \lambda(P_\theta)$ ,  $\tilde{\mu}(\theta) = \mu(P_\theta)$ . Ces deux fonctions sont définies sur  $\mathbf{R}/2\pi\mathbf{Z}$ . On devrait avoir

$$\forall \theta \in \mathbf{R}/2\pi\mathbf{Z}, \quad P_\theta(X) = (X - \tilde{\lambda}(\theta))(X - \tilde{\mu}(\theta)).$$

### Proposition

*Il n'existe pas de fonctions continues  $\tilde{\lambda}, \tilde{\mu}$  définies de  $\mathbf{R}/2\pi\mathbf{Z}$  dans  $\mathbf{C}$  telles que*

$$X^2 - e^{i\theta} = (X - \tilde{\lambda}(\theta))(X - \tilde{\mu}(\theta))$$

### Preuve

Pour chaque  $\theta \in ]0, 2\pi[$ , on doit avoir  $\tilde{\lambda}(\theta) \in \{e^{i\frac{\theta}{2}}, -e^{i\frac{\theta}{2}}\}$ . L'image de  $]0, 2\pi[$  par  $\tilde{\lambda}$  est donc contenue dans  $S^1 \setminus \{-1, 1\}$ . Cet espace admet deux composantes connexes. Supposons, sans perte de généralité, que  $\tilde{\lambda}(]0, 2\pi[)$  est dans le demi-cercle supérieur. On a alors l'égalité

$$\forall \theta \in ]0, 2\pi[, \quad \tilde{\lambda}(\theta) = e^{i\frac{\theta}{2}}.$$

Comme  $\tilde{\lambda}$  est continue, on devrait avoir  $\tilde{\lambda}(0) = 1$ ,  $\tilde{\lambda}(2\pi) = -1$ , si bien que  $\tilde{\lambda}$  n'est pas périodique de période  $2\pi$ . Ceci termine la preuve.

On peut construire de la même façon une famille de polynômes à coefficients réels dont les racines complexes ne dépendent pas continûment des polynômes. Posons

$$\forall \theta \in \mathbf{R}/2\pi\mathbf{Z}, \quad P_\theta(X) = X^4 + (8 - 2\cos\theta)X^2 + 8\sin(\theta)X + 8\cos(\theta) + 17$$

Ces polynômes admettent pour racines les nombres  $2i + e^{-i\frac{\theta}{2}}$ ,  $-2i + e^{i\frac{\theta}{2}}$ ,  $2i - e^{i\frac{\theta}{2}}$  et  $-2i - e^{-i\frac{\theta}{2}}$ . En raisonnant comme précédemment, on montre qu'il n'est pas possible de construire une fonction dépendant continûment de  $\theta \in \mathbf{R}/2\pi\mathbf{Z}$ , à valeurs dans cet ensemble de racines. On pourra construire en exercice un exemple de polynôme de degré 3 dont les racines ne dépendent pas de manière continue de ses coefficients.

Les problèmes qui apparaissent dans les exemples précédents sont dus à un échange des racines, après une rotation de  $\pi$  sur l'angle  $\theta$ . Nous allons voir que cette permutation est la seule obstruction à la continuité des racines en fonction du polynôme.

### Racines à permutation près

Nous voulons considérer l'ensemble des racines d'un polynôme de degré  $n$  à permutation près. Notons  $S_n$  le groupe des permutation de l'ensemble  $\{1, \dots, n\}$ . Ce groupe agit sur  $\mathbf{C}^n$  par échange des indices. Soit  $(x_1, \dots, x_n) \in \mathbf{C}^n$  et  $\sigma \in S_n$  :

$$\sigma(x_1, \dots, x_n) = (x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

Le quotient de  $\mathbf{C}^n$  par cette action est noté  $\mathbf{C}^n/S_n$ . Il est muni de la topologie quotient. Notons  $\mathcal{E}_n$  l'ensemble des polynômes unitaires de degré  $n$ , à coefficients complexes.

### Théorème

L'application  $\psi : \mathbf{C}^n \rightarrow \mathcal{E}_n$  définie par

$$\psi(\lambda_1, \dots, \lambda_n) = \prod_{i=1}^n (X - \lambda_i).$$

*est continue, passe au quotient et donne un homéomorphisme de  $\mathbf{C}^n/S_n$  sur  $\mathcal{E}_n$ . Les racines, à permutation près, dépendent donc continûment du polynôme.*

### Preuve

L'application  $\psi$  est bien continue car les coefficients d'un polynôme s'expriment de manière polynomiale en fonction des racines de ce polynôme. Tout polynôme de degré  $n$  à coefficients complexes possède  $n$  racines, l'application  $\psi$  est donc surjective. Comme deux ensembles de racines définissent le même polynôme si et seulement si ils se déduisent l'un de l'autre par permutation, on obtient bien une bijection après être passé au quotient.

Il reste à montrer que si  $\bar{F}$  est un fermé de  $\mathbf{C}^n/S_n$ , alors son image est fermée dans  $\mathcal{E}_n$ . Comme  $\pi^{-1}(\bar{F})$  est fermé, il suffit de vérifier que pour tout fermé  $F$  de  $\mathbf{C}^n$ , l'ensemble  $\psi(F)$  est fermé. Nous allons utiliser le lemme suivant.

**Lemme**

Soit  $P(X) = X^n + \sum_{k=1}^{n-1} a_k X^k$  un polynôme unitaire à coefficients complexes de degré  $n$ , et soit  $\lambda$  une racine de  $P$ . Alors

$$|\lambda| \leq \max \left\{ 1, \sum_{k=0}^{n-1} |a_k| \right\}$$

**Preuve**

Si  $|\lambda|$  est plus petit que 1, la conclusion est vérifiée. Sinon, nous avons la majoration  $|\lambda|^k \leq |\lambda|^{n-1}$  pour tout  $k < n$ , ce qui implique

$$|\lambda^n| = \left| - \sum_{k=0}^{n-1} a_k \lambda^k \right| \leq \sum_{k=0}^{n-1} |a_k| |\lambda|^{n-1} = \left( \sum_{k=0}^{n-1} |a_k| \right) |\lambda|^{n-1}.$$

Le lemme est démontré.

Terminons maintenant la preuve du théorème. Soit  $F$  un fermé de  $\mathbf{C}^n$  et  $\{P_k\}$  une suite dans  $\psi(F)$  qui converge vers un certain polynôme  $P \in \mathcal{E}_n$ . Posons  $P_k = \psi(\lambda_k)$  avec  $\lambda_k \in F$ . Les coefficients des  $P_k$  sont tous majorés par une constante qui ne dépend pas de  $k$ . D'après le lemme, il en va de même pour tous les coefficients des  $\lambda_k$ . On peut donc extraire une sous-suite  $\lambda_{k_l} \in \mathbf{C}^n$  qui converge vers un vecteur  $\lambda \in F$ . Par continuité de  $\psi$ , nous avons l'égalité

$$P = \psi(\lambda) \in \psi(F),$$

ce qui montre que  $\psi(F)$  est fermé.

**Remarques**

– D'après ce qui précède, l'espace  $\mathbf{C}^n/S_n$  est un espace topologique séparé homéomorphe à  $\mathbf{C}^n$ .

– Plutôt que de travailler avec la topologie quotient sur  $\mathbf{C}^n/S_n$ , on aurait pu définir une distance sur cet espace. Munissons  $\mathbf{C}^n$  de la norme euclidienne  $\|\cdot\|$ , et considérons sur  $\mathbf{C}^n$  la fonction

$$\forall x, y \in \mathbf{C}^n, \quad d(x, y) = \min \{ \|x - \sigma(y)\| \mid \sigma \in S_n \}.$$

Cette fonction passe au quotient. On vérifie qu'on obtient une distance sur  $\mathbf{C}^n/S_n$  et que la projection canonique  $\pi : \mathbf{C}^n \rightarrow \mathbf{C}^n/S_n$  est continue.

**Décomposition en polynômes symétriques élémentaires**

Le théorème précédent est à rapprocher du théorème de décomposition des polynômes symétriques en polynômes symétriques élémentaires. Rappelons comment sont définis ces polynômes symétriques élémentaires.

$$\sigma_k(\lambda_1, \dots, \lambda_n) = \sum_{\substack{S \subset \{1, \dots, n\} \\ \#S=k}} \prod_{i \in S} \lambda_i$$

Montrons qu'on peut exprimer toute fonction continue invariante par permutation des coordonnées, en fonction des polynômes symétriques élémentaires  $\sigma_i$ .

### Corollaire

Soit  $f : \mathbf{C}^n \rightarrow \mathbf{C}$  une fonction continue, invariante par l'action de  $S_n$  sur les coordonnées. Alors il existe une fonction continue  $g : \mathbf{C}^n \rightarrow \mathbf{C}$  telle que

$$\forall (\lambda_1, \dots, \lambda_n) \in \mathbf{C}^n, \quad f(\lambda_1, \dots, \lambda_n) = g(\sigma_1(\lambda_1, \dots, \lambda_n), \dots, \sigma_n(\lambda_1, \dots, \lambda_n))$$

### Preuve

Notons  $\bar{\psi} : \mathbf{C}^n/S_n \xrightarrow{\sim} \mathcal{E}_n$  l'homéomorphisme donné par le théorème, et soit  $\bar{f} : \mathbf{C}^n/S_n \rightarrow \mathbf{C}$  l'application obtenue à partir de  $f$  en passant au quotient. Définissons une application  $p : \mathbf{C}^n \xrightarrow{\sim} \mathcal{E}_n$  en posant

$$p(\sigma_1, \dots, \sigma_n) = X^n + \sum X^{n-k} (-1)^k \sigma_k.$$

Cette application satisfait la relation  $p(\sigma_1(\lambda_i), \dots, \sigma_n(\lambda_i)) = \prod (X - \lambda_i)$ . Il suffit donc de poser  $g = \bar{f} \circ \bar{\psi}^{-1} \circ p$  pour conclure.

### Polynômes dont toutes les racines sont réelles

Exprimer de manière continue les racines d'un polynôme appartenant à un certain sous-ensemble  $E \in \mathcal{E}_n$  revient donc à chercher une section pour la projection  $\pi : \mathbf{C}^n \rightarrow \mathbf{C}^n/S_n$  au-dessus du sous-ensemble  $E' = \pi(\psi^{-1}(E))$ , c'est-à-dire une application continue  $s : E' \rightarrow \mathbf{C}^n$  telle que  $\pi \circ s = id$ .

Intéressons nous aux polynômes à valeurs réelles dont toutes les racines sont réelles. Définissons une section  $s : \mathbf{R}^n/S_n \rightarrow \mathbf{R}^n$  comme suit : le vecteur  $s(\lambda)$  est obtenu à partir de  $\lambda$  en ordonnant ses coordonnées de manière croissante.

### Proposition

L'application  $s : \mathbf{R}^n/S_n \rightarrow \mathbf{R}^n$  est continue.

### Preuve

Soit  $\lambda_k$  une suite de  $\mathbf{R}^n/S_n$  qui converge vers un certain  $\lambda \in \mathbf{R}^n/S_n$ . On veut montrer que  $s(\lambda_k)$  converge vers  $s(\lambda)$ . Comme les coefficients de  $s(\lambda_k)$  sont bornés, il suffit de vérifier que toute valeur d'adhérence  $\mu$  de  $s(\lambda_k)$  est en fait égale à  $s(\lambda)$ . Les coefficients de  $s(\lambda_k)$  sont ordonnés de manière croissante, il en va de même pour les coefficients de  $\mu$ . De plus,  $\mu$  se projette sur  $\lambda$  par le biais de l'application  $\pi : \mathbf{R}^n \rightarrow \mathbf{R}^n/S_n$ . Le vecteur  $\mu$  est donc obtenu à partir de  $\lambda$  en ordonnant ses coordonnées, i.e.  $\mu = s(\lambda)$ .

On en déduit la continuité des racines lorsque celles-ci sont réelles, ordonnées par ordre croissant.

### Proposition

Considérons l'ensemble des polynômes unitaires de degré  $n$  à coefficients réels dont toutes les racines sont réelles. Alors l'application qui à un tel polynôme associe ses racines, ordonnées par ordre croissant, est continue.

## Sur la diagonalisation des matrices 2x2

On sait que toute matrice  $A$ , à coefficients réels ou complexes, dont les valeurs propres sont toutes distinctes, est diagonalisable. Peut-on réaliser cette diagonalisation de manière continue ? En d'autres termes, peut-on choisir la matrice conjuguant  $A$  à une matrice diagonale, de façon à ce que celle-ci dépende continûment de  $A$  ? Le but de ce texte est de démontrer que cela n'est pas possible sur l'ouvert des matrices dont les valeurs propres sont toutes distinctes.

### 1 Etude locale

Remarquons d'abord que si  $M$  est conjuguée à une matrice diagonale  $D$  par le biais d'une matrice  $U \in GL_n$ ,

$$U^{-1}MU = D$$

alors les coefficients diagonaux de  $D$  sont des valeurs propres de  $M$  et les vecteurs colonnes de  $U$  sont des vecteurs propres de  $M$ . Réciproquement, si  $U$  est une matrice inversible dont les colonnes sont des vecteurs propres de  $M$ , alors  $U^{-1}MU$  est diagonale.

Par conséquent, diagonaliser  $M$  continûment revient peu ou prou à faire un choix pour les vecteurs propres de  $M$ , qui dépend continûment de  $M$ .

Ce choix est toujours possible localement, au voisinage d'une matrice dont toutes les valeurs propres sont distinctes. C'est une application classique du théorème d'inversion locale.

Pour simplifier, on va se restreindre au cas des matrices 2x2, et donner des expressions explicites pour ces conjugaisons. Intéressons nous au cas des matrices à coefficients réels et notons  $\mathcal{U}$  l'ouvert de  $M_2(\mathbf{R})$  correspondant aux matrices ayant leurs deux valeurs propres distinctes.

$$\mathcal{U} = \{M \in M_2(\mathbf{R}) \mid \Delta(P_c(M)) \neq 0\}$$

où  $P_c(M)$  est le polynôme caractéristique de  $M$  et  $\Delta$  est son discriminant :  $\Delta(x^2 + \alpha x + \beta) = \alpha^2 - 4\beta$ . Cet ouvert  $\mathcal{U}$  a deux composantes connexes  $\mathcal{U}_+$  et  $\mathcal{U}_-$  correspondant à  $\Delta > 0$  et  $\Delta < 0$ . Considérons le cas  $\Delta > 0$ , c'est à dire le cas où  $M$  a ses deux valeurs propres réelles.

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Les deux valeurs propres de  $M$  sont données par les expressions :

$$\lambda_+ = \frac{1}{2}(a + d + \sqrt{(a - d)^2 + 4bc}), \quad \lambda_- = \frac{1}{2}(a + d - \sqrt{(a - d)^2 + 4bc})$$

Elles dépendent continûment de  $a, b$  et  $c$  sur l'ouvert  $\mathcal{U}_+$ . Les vecteurs propres associés à  $\lambda_+$  sont proportionnels à  $\begin{pmatrix} a - b \\ a - \lambda_+ \end{pmatrix}$  ; remarquons que ce vecteur est lui-même proportionnel à  $\begin{pmatrix} a - \lambda_+ \\ c \end{pmatrix}$ . Les vecteurs propres associés à  $\lambda_-$  sont proportionnels à  $\begin{pmatrix} a - b \\ a - \lambda_- \end{pmatrix}$  ; remarquons que ce vecteur est lui-même proportionnel au vecteur propre  $\begin{pmatrix} a - \lambda_- \\ c \end{pmatrix}$ .

On peut donc former différentes matrices susceptibles de diagonaliser  $M$  à partir de ces vecteurs ; par exemple, la matrice  $\begin{pmatrix} a-b & b \\ a-\lambda_+ & a-\lambda_- \end{pmatrix}$ , dont le déterminant est égal à  $b(\lambda_+ - \lambda_-)$ , ou  $\begin{pmatrix} b & d-\lambda \\ a-\lambda_+ & c \end{pmatrix}$  de déterminant égal à  $bc + (a - \lambda_+)^2$  (car  $\lambda_+ + \lambda_- = a + d$ ), ou encore  $\begin{pmatrix} d-\lambda_+ & b \\ c & a-\lambda_- \end{pmatrix}$  de déterminant égal à  $-bc - (a - \lambda_-)^2$ . Par conséquent,

- Sur  $\mathcal{U}_+ \cap \{b \neq 0\}$ ,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} b & b \\ a-\lambda_+ & a-\lambda_- \end{pmatrix}^{-1} \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \begin{pmatrix} b & b \\ a-\lambda_+ & a-\lambda_- \end{pmatrix}$$

- Sur  $\mathcal{U}_+ \cap \{a \neq \lambda_+, |bc| < (a - \lambda_+)^2\}$ ,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} b & d-\lambda_- \\ a-\lambda_+ & c \end{pmatrix}^{-1} \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \begin{pmatrix} b & d-\lambda_- \\ a+\lambda_+ & c \end{pmatrix}$$

- Sur  $\mathcal{U}_+ \cap \{a \neq \lambda_-, |bc| < (a - \lambda_-)^2\}$ ,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d-\lambda_+ & b \\ c & a-\lambda_- \end{pmatrix}^{-1} \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \begin{pmatrix} d-\lambda_+ & b \\ c & a-\lambda_- \end{pmatrix}$$

Les trois ouverts précédents recouvrent  $\mathcal{U}_+$  ; on est donc parvenu à diagonaliser  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , au moins localement. Le problème est que les trois matrices qui réalisent ces conjugaisons ne coïncident pas sur l'intersection de ces ouverts, si bien qu'il n'est pas possible de les "recoller" afin de former une solution globale continue qui conjugue  $M$  à une matrice diagonale.

On pourrait penser que cela est dû à un mauvais choix, au niveau des vecteurs propres. Il n'en est rien.

### Théorème

Soit  $\mathbf{K} = \mathbf{R}$  et  $\mathcal{U} = \{M \in M_2(\mathbf{R}) \mid \Delta(P_c(M)) > 0\}$ ,

ou  $\mathbf{K} = \mathbf{C}$  et  $\mathcal{U} = \{M \in M_2(\mathbf{C}) \mid \Delta(P_c(M)) \neq 0\}$ .

Il n'existe pas de fonction continue  $f : \mathcal{U} \rightarrow GL_2(\mathbf{K})$  telle que, pour tout  $M \in \mathcal{U}$ ,  $f(M)^{-1}Mf(M)$  soit diagonale.

**Remarque :** ce théorème est en fait vrai en toute dimension.

### 2 Le cas réel

Remarquons que si une telle fonction  $f$  existait, alors on pourrait diagonaliser continûment les matrices symétriques à l'aide de matrices de  $SO_2$ . En effet, si  $v_1, v_2$  sont les deux vecteurs colonnes de  $f(M) : f(M) = (v_1, v_2)$ , alors  $\tilde{f}(M) = \left(\frac{v_1}{|v_1|}, \frac{v_2}{|v_2|}\right)$  conjugue encore  $M$  à une matrice diagonale. Si  $M$  est symétrique, ses vecteurs propres  $v_1$  et  $v_2$  sont orthogonaux ; par conséquent  $\tilde{f}(M) \in O_2(\mathbf{R})$ . La fonction  $\frac{\tilde{f}(M)}{\det f(M)} \in SO_2(\mathbf{R})$  réalise donc la conjugaison recherchée. Le théorème précédent découle donc de l'énoncé suivant.

### Théorème

Soit  $Sym(\mathbf{R}^2) = \{M \in M_2(\mathbf{R}) \mid {}^t M = M\}$ .

Il n'existe pas de fonction continue  $f : \mathcal{U} \cap \text{Sym}(\mathbf{R}^2) \longrightarrow SO_2(\mathbf{R})$  telle que, pour tout  $M \in \mathcal{U} \cap \text{Sym}(\mathbf{R}^2)$ ,  $f(M)^{-1}Mf(M)$  soit diagonale.

Au lieu de considérer l'ensemble de toutes les matrices symétriques, on peut même se restreindre à la classe de conjugaison d'une matrice diagonale  $A_0 \in \mathcal{U}$ . Posons

$$\mathcal{O}_{A_0} = \{UA_0U^{-1} \mid U \in SO_2(\mathbf{R})\}.$$

Les matrices de la forme  $f(A)^{-1}Af(A)$ ,  $A \in \mathcal{O}_{A_0}$ , sont diagonales et conjuguées à  $A_0$  ; elles ont donc les mêmes valeurs propres que  $A_0$ . Il n'existe qu'un nombre fini de telles matrices, elles sont obtenues en permutant les termes diagonaux de  $A_0$ . Comme  $\mathcal{O}_{A_0}$  est connexe, on voit que  $f(A)^{-1}Af(A)$  est constant. Quitte à multiplier  $f$  par une matrice de permutation, on peut donc supposer que  $f(A)^{-1}Af(A)$  est égale à  $A_0$ .

### Théorème

Soit  $A_0 \in \mathcal{U}$  une matrice diagonale.

Il n'existe pas de fonction continue  $f : \mathcal{O}_{A_0} \longrightarrow SO_2(\mathbf{R})$  tel que

$$f(A)^{-1}Af(A) = A_0.$$

### Lemme

Soit  $A_0 \in \mathcal{U}$  une matrice diagonale.

Si  $A = UA_0U^{-1} = VA_0V^{-1}$ , alors  $UV^{-1}$  est diagonale.

### Preuve du lemme

$UV^{-1}$  doit commuter avec  $A_0$ . La matrice  $UV^{-1}$  doit donc laisser invariant les sous-espaces propres de  $A_0$  ; ceux-ci sont engendrés par les vecteurs de la base canonique. La matrice  $UV^{-1}$  est donc diagonale.

### Preuve du théorème

Soit  $\mathcal{D}$  le sous-ensemble des matrices diagonales de  $M_2(\mathbf{R})$ . Considérons la projection  $\pi$  de  $SO_2$  sur  $\mathcal{O}_{A_0}$  donnée par

$$\begin{aligned} \pi : SO_2(\mathbf{R}) &\longrightarrow \mathcal{O}_{A_0} \\ U &\longmapsto UA_0U^{-1} \end{aligned}$$

Le lemme montre que les "fibres" de cette projection s'identifient naturellement à  $\mathcal{D} \cap SO_2(\mathbf{R})$  :

$$\pi(u) = \pi(v) \iff uv^{-1} \in \mathcal{D} \cap SO_2(\mathbf{R}).$$

L'existence de  $f$  permettrait d'établir un homéomorphisme entre  $SO_2(\mathbf{R})$  et  $\mathcal{O}_{A_0} \times (\mathcal{D} \cap SO_2(\mathbf{R}))$  :

$$\begin{aligned} \mathcal{O}_{A_0} \times (\mathcal{D} \cap SO_2(\mathbf{R})) &\longrightarrow SO_2(\mathbf{R}) \\ (A, D) &\longmapsto f(A)D \end{aligned}$$

On peut écrire explicitement l'inverse de cette application :

$$\begin{aligned} \mathcal{O}_{A_0} \times (\mathcal{D} \cap SO_2(\mathbf{R})) &\longleftarrow SO_2(\mathbf{R}) \\ (UA_0U^{-1}, f(UA_0U^{-1})^{-1}U) &\longleftarrow U \end{aligned}$$

La matrice  $f(UA_0U^{-1})^{-1}U$  est bien diagonale car elle commute avec  $A_0$ . En effet, d'après la définition de  $f$ , on doit avoir l'égalité

$$f(UA_0U^{-1})^{-1}UA_0U^{-1}f(UA_0U^{-1}) = A_0.$$

On est parvenu à une absurdité. Il n'existe pas d'homéomorphisme entre  $SO_2(\mathbf{R})$  et  $\mathcal{O}_{A_0} \times (\mathcal{D} \cap SO_2(\mathbf{R}))$  car  $SO_2(\mathbf{R})$  est connexe tandis que  $\mathcal{D} \cap SO_2(\mathbf{R}) = \{Id, -Id\}$  n'est pas connexe.

### Remarques

- La preuve se généralise à des matrices de taille quelconque.
- L'application  $\pi$  est un revêtement à deux feuillet non trivial de  $S^1$  par  $S^1$ .

### 3 Le cas complexe

Les énoncés précédents se généralisent au cas complexe en remplaçant  $SO_2(\mathbf{R})$  par  $SU_2(\mathbf{C})$  et les matrices symétriques par les matrices hermitiennes.

Les arguments précédents établiraient un homéomorphisme entre  $SU_2(\mathbf{C})$  et  $\mathcal{O}_{A_0} \times (\mathcal{D} \cap SU_2(\mathbf{C}))$ . Mais  $SU_2(\mathbf{C})$  est homéomorphe à  $S^3$  qui est simplement connexe ; l'homéomorphisme est donné par

$$\begin{aligned} \{(a, b) \in \mathbf{C}^2 \mid |a|^2 + |b|^2 = 1\} &\longrightarrow SU_2(\mathbf{C}) \\ (a, b) &\longmapsto \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \end{aligned}$$

tandis que  $\mathcal{D} \cap SU_2(\mathbf{C}) = \left\{ \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \mid \theta \in \mathbf{R} \right\}$  est homéomorphe au cercle  $S^1$ , qui n'est pas simplement connexe.

### Remarques

- on peut démontrer que  $\mathcal{O}_{A_0}$  est homéomorphe à la sphère  $S^2$ . La projection  $\pi : SU_2(\mathbf{C}) \longrightarrow \mathcal{O}_{A_0}$  est la fibration de Hopf.
- La preuve se généralise en dimension quelconque. Le groupe  $\mathcal{D} \cap SU_n(\mathbf{C})$  est maintenant homéomorphe à un tore  $S^1 \times S^1 \times \dots \times S^1$ .



## Théorie de Galois

### Théorème

Soit  $K$  un corps de caractéristique nulle (ou un corps fini),  $L$  une extension finie de  $K$ , supposée normale :  $\forall \sigma \in \text{hom}_K(L, \overline{K}), \sigma(L) \subset L$ . Alors on a une bijection

$$\begin{array}{ccc} \{\text{sous-corps de } L \text{ contenant } K\} & \longleftrightarrow & \{\text{sous-groupes de } \text{Aut}_K(L)\} \\ & & L' \longrightarrow \text{Aut}_{L'}(L) \\ L^H = \{x \in L \mid \forall \sigma \in H, \sigma x = x\} & \longleftarrow & H \end{array}$$

De plus,  $\text{Card } \text{Aut}_K(L) = \dim_K L$ .

Pour les applications, on remarque que les extensions de décomposition sont normales. Pour la preuve, on utilise les trois résultats suivants :

- Théorème de l'élément primitif : si  $\text{car } K = 0$  ou  $K$  fini,  $L$  extension finie de  $K$  Alors il existe  $\theta \in L$  tel que  $L = K[\theta]$ .
- Polynômes symétriques : si  $P \in K[X_1, \dots, X_n]$  est un polynôme symétrique, il existe  $\tilde{P} \in K[X_1, \dots, X_n]$  tel que  $P(X_1, \dots, X_n) = \tilde{P}(S_1, \dots, S_n)$  où on a noté  $S_k(X_1, \dots, X_n) = \sum_{i_1 < \dots < i_k} X_{i_1} \dots X_{i_k}$ .
- Si la caractéristique de  $K$  est nulle ou si  $K$  est fini, tout polynôme irréductible a ses racines simples dans  $\overline{K}$ .

### Notations

si  $P(X) = \sum a_n X^n \in L[X]$ , on pose  $(\sigma P)(X) = \sum \sigma(a_n) X^n$ . On a alors

$$\forall x \in L, \forall \sigma \in \text{Aut}_K(L), \sigma(P(x)) = (\sigma P)(\sigma x).$$

En particulier, si  $P \in K[X]$  et  $P(x) = 0$ , alors  $\forall \sigma \in \text{Aut}_K(L), P(\sigma x) = 0$ .

### Lemme

Soit  $P_{L/K}^x$  le polynôme minimal de  $x \in L$  et  $\Gamma_x = \{\sigma x \mid \sigma \in \text{Aut}_K(L)\}$ . Alors

$$P_{L/K}^x(X) = \prod_{y \in \Gamma_x} (X - y).$$

### Preuve du lemme

– Cas  $x = \theta$

Les  $\sigma\theta$  étant tous distincts et annulant  $P_{L/K}^\theta \in K[X]$ , on a dans  $L[X]$ ,

$$\prod_{\sigma \in \text{Aut}_K(L)} (X - \sigma\theta) \mid P_{L/K}^\theta.$$

En particulier  $\text{Card}(\text{Aut}_K L) \leq \dim_K L$ . Mais si  $\theta'$  est une racine de  $P_{L/K}^\theta$ ,

$$K[\theta'] \simeq K[X]/P_{L/K}^{\theta'} = K[X]/P_{L/K}^\theta \simeq L$$

Cet isomorphisme envoie  $\theta'$  sur  $\theta$ . Par conséquent, il existe  $\sigma \in \text{Aut}_K(L)$  tel que  $\sigma\theta = \theta'$ , ce qui implique  $\text{Card } \text{Aut}_K L \geq d^\circ P_{L/K}^\theta = \dim_K L$ .

Les deux polynômes ont donc même degré, l'un divise l'autre ; ils sont égaux.

– *Cas général*

Il existe  $P \in K[X]$  tel que  $P(\theta) = x$ . Posons

$$P_c^x(X) = \prod_{\sigma \in \text{Aut}_{K/L}} (X - \sigma(x)) \in L[X]$$

En fait,  $P_c^x \in K[X]$ . En effet, ses coefficients sont des polynômes symétriques en les  $\sigma\theta$ , ils s'expriment donc en fonctions des polynômes symétriques élémentaires des  $\sigma\theta$ , c'est à dire en fonction des coefficients de  $P_{L/K}^\theta$ . On a donc  $P_{L/K}^x | P_c^x$  dans  $K[X]$ .  $P_{L/K}^x$  a ses racines simples, donc  $P_{L/K}^x | \prod_{y \in \Gamma_x} (X - y)$  dans  $\overline{K}[X]$ . Les  $y \in \Gamma_x$  étant racines de  $P_{L/K}^x$ , on a égalité.

**Corollaire du lemme :**  $L^{\text{Aut}_{K/L}} = K$  car si  $x \in L^{\text{Aut}_{K/L}}$ ,  $\Gamma_x = \{x\}$ .

**Preuve du théorème**

– Si  $K \subset L' \subset L$ ,  $L'$  est de caractéristique 0 ou fini,  $L$  est normal sur  $L'$ , on peut donc appliquer le corollaire :  $L^{\text{Aut}_{L'/L}} = L'$ .

– Pour l'autre sens, on a clairement  $H \subset \text{Aut}_{L^H}(L)$ .

Soit  $\varphi \in \text{Aut}_{L^H}(L)$  : si  $x \in L$  est tel que  $\forall h \in H$ ,  $h(x) = x$  alors  $\varphi(x) = x$ . On considère  $Q(X) = \prod_{\sigma \in H} (X - \sigma\theta)$ .  $Q$  est invariant par  $H$  donc  $Q \in L^H[X]$ .

Ses coefficients sont invariants par  $L^H$ , donc par  $\varphi$ .

$$\varphi(Q(\theta)) = \varphi(0) = 0 = (\varphi Q)(\varphi(\theta)) = Q(\varphi(\theta)).$$

Donc  $\varphi\theta = \sigma\theta$  pour un certain  $\sigma \in H$  ;  $\varphi = \sigma \in H$ , ce qu'il fallait démontrer.

• **Preuve du théorème de l'élément primitif** en caractéristique 0

Il suffit de le démontrer pour les extensions de la forme  $K[\alpha, \beta]$ . On cherche  $\theta$  sous la forme  $\theta = \beta + c\alpha$ , avec  $c$  entier.

Soit  $P^\alpha(X) = \prod_{i=0}^n (X - \alpha_i)$ ,  $P^\beta(X) = \prod_{i=0}^m (X - \beta_i)$  les polynômes minimaux de  $\alpha$  et  $\beta$ . On a :  $P^\alpha, P^\beta \in K[X]$ ,  $\alpha_0 = \alpha$ ,  $\beta_0 = \beta$ ,  $\beta_i, \alpha_i \in \overline{K}$ . Considérons le polynôme  $P(X) = P^\beta(\theta - cX)$ , où  $c$  est un entier à déterminer. On a les égalités  $P(\alpha) = P^\beta(\beta) = 0$ ,  $P^\alpha(\alpha) = 0$  et  $\alpha$  est racine simple de  $P^\alpha$ .

Remarquons que si  $c \neq \frac{\beta_i - \beta_j}{\alpha - \alpha_j}$ ,  $\forall i, j$ , alors  $P(\alpha_j) \neq 0 \quad \forall j = 1 \text{ à } n$ .

On choisit un tel  $c$  appartenant à  $K$ . Maintenant,  $\text{pgcd}(P, P^\alpha) = X - \alpha$  dans  $\overline{K}[X]$  car  $\alpha$  est la seule racine commune. De plus  $P, P^\alpha \in K[\theta][X]$ , il en va donc de même du pgcd, celui-ci se calculant par exemple par l'algorithme d'Euclide. Donc  $\alpha \in K[\theta]$ ,  $\beta = \theta - c\alpha \in K[\theta]$ ,  $K[\alpha, \beta] = K[\theta]$ .

• **Preuve** irréductibilité  $\Rightarrow$  racines simples

$\text{pgcd}_{K[X]}(P, P') = 1$  car  $P$  est irréductible,  $d^\circ P' < d^\circ P$ ,  $P' \not\equiv 0$  donc  $P$  et  $P'$  n'ont pas de racines en commun dans  $\overline{K}$ .

• **Remarque finale**

L'hypothèse car  $K = 0$  ou  $K$  fini peut être remplacée par l'hypothèse l'extension  $L/K$  est séparable :  $\forall x \in L$ ,  $P_{L/K}^x$  a ses racines simples. Les preuves sont inchangées.

## Calculs en MAPLE

### SERIES NUMERIQUES

Quelques séries que MAPLE sait calculer.

```

> sum((-1)^(n+1)/n,n=1..infinity);
      ln(2)
> sum(1/(n*2^n),n=1..infinity);
      ln(2)
> sum(1/n^2,n=1..infinity);
      1/6 π^2
> sum(1/n!,n=0..infinity);
      e^1
> sum(n/2^n,n=1..infinity);
      2
> sum(1/(1+n^2),n=-infinity..infinity);
      π coth(π)
> sum(1/(n*(n+1)),n=1..infinity);
      1
> sum(1/n,n=1..infinity);
      ∞

```

Quelques autres où il sèche.

```

> sum(sin(n)/n,n=1..infinity);
      ∑_{n=1}^∞ sin(n)/n
> sum(exp(-n^2),n=0..infinity);
      ∑_{n=0}^∞ e^{-n^2}
> sum((-1)^n,n=1..infinity);
      undefined
> sum(1/n^3,n=1..infinity);
      ζ(3)

```

Enfin, le résultat obtenu n'est pas toujours très clair.

```

>eval(sum((n^4+3*n-3)/(n^6+3*n^5+n^4+4*n^3+n),n=1..infinity));
-∞ signum ( 16/5 + 1/15 ∑_{α=RootOf(-Z^3+3-Z^2+1)} -23 + 15 α^2 + 38 α ) - 3 γ -
∑_{α=RootOf(-Z^3+3-Z^2+1)} ( 23/15 - α^2 - 38/15 α ) Ψ(1-α) + (1/10 - 4/5 i) Ψ(1-i) +
(1/10 + 4/5 i) Ψ(1+i)

```

## CALCUL D'INTEGRALES

Quelques intégrales que MAPLE sait calculer.

```

> int(1/(x*(x+1)),x=1..infinity);
      ln(2)
> int(1/(1+x^2),x=-infinity..infinity);
      pi
> int(sin(x)/x, x=0..infinity);
      1/2 pi
> int((sin(x))^3/x^3,x=0..infinity);
      3/8 pi
> int(1/x,x=0..infinity);
      infinity
> int(exp(-x^2/2),x=-infinity..infinity);
      sqrt(2) sqrt(pi)
> int(exp(I*x^2)*cos(x),x=-infinity..infinity);
      sqrt(pi) 4sqrt(-1)e^{-1/4 i}
> int(exp(-x)*x^8,x=0..infinity);
      40320

```

et quelques autres qui lui posent problème.

```

> simplify((evalc(int(exp(-abs(x))/(1+x^2),x=0..infinity)))));
      sin(1) Ci(1) - cos(1) Si(1) + 1/2 cos(1) pi
> int(sin(x)/sinh(x),x=0..infinity);
      integral_0^infinity sin(x)/sinh(x) dx
> int(1/sqrt(1+x^4),x=-infinity..infinity);
      1/2 beta(1/4, 1/4)

```

## Approximation de la loi gaussienne

On cherche à évaluer la loi gaussienne et tout particulièrement le reste

$$f(x) = \int_x^{\infty} e^{-\frac{1}{2}t^2} dt$$

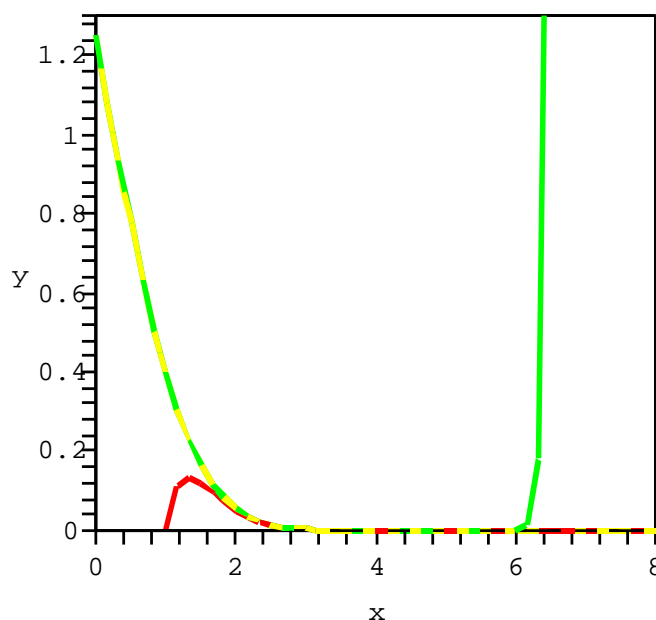
On peut pour cela utiliser le développement en série convergente obtenu en intégrant le développement de l'exponentielle :

$$\sqrt{\frac{\pi}{2}} + \sum_{k=0}^{\infty} (-1)^{k+1} \frac{x^{2k+1}}{2^k k! (2k+1)}$$

Cependant cette série converge lentement ; dès que  $x$  est grand le nombre de termes à calculer devient rédhibitoire. Le développement asymptotique divergent suivant donne de meilleurs résultats :

$$e^{-\frac{x^2}{2}} \left( \frac{1}{x} + \sum_{k=1}^{\infty} \frac{(-1)^k (2k-1)!}{2^{k-1} (k-1)!} \frac{1}{x^{2k+1}} \right)$$

Le graphique suivant compare la courbe de  $f$  à celles obtenues en gardant les termes  $k < 2$  de la série divergente et  $k < 50$  de la série convergente.



Dès que  $x$  est supérieur à 5, deux termes de la série divergente donnent une meilleure approximation que cinquante termes de la série convergente. Les deux développements permettent d'obtenir des encadrements de  $f$  ; on peut démontrer que  $f$  est coincée entre deux sommes consécutives de l'une quelconque des séries. Par exemple,

$$\left(1 - \frac{1}{x^2}\right) \frac{1}{x} e^{-\frac{x^2}{2}} < f(x) < \frac{1}{x} e^{-\frac{x^2}{2}}$$

### Exercices d'oral

- Diagonaliser la matrice  $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ .
- Décomposer en polynômes symétriques élémentaires  $X_1^3 + X_2^3$ .
- Donner l'axe et l'angle de la rotation  $\frac{1}{3} \begin{bmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \\ 1 & -2 & 2 \end{bmatrix}$ .
- Montrer que  $\int_0^{+\infty} \left| \frac{\sin(t)}{t} \right| dt = +\infty$ .
- Calculer  $\sum_{n \geq 1} \frac{(-1)^n}{n}$ . On justifiera soigneusement le calcul.
- Factoriser dans  $\mathbf{R}[X]$  le polynôme  $X^4 + 1$ .
- Montrer que les polynômes  $X^4 + X^3 + X^2 + X + 1$  et  $X^4 + 6X^2 + 9X + 3$  sont irréductibles sur  $\mathbf{Q}$ .
- Calculer  $\int_{-\infty}^{+\infty} \frac{e^{itx}}{1+x^2} dx$ .
- Calculer  $\int_{-\infty}^{+\infty} \frac{\sin t}{t} dt$ .
- Décomposer en éléments simples la fraction rationnelle  $\frac{x}{1+x^2}$ .
- Quelle est la dimension de l'espace vectoriel des matrices strictement triangulaires supérieures ?
  - $X^4 + 1$  est-il un polynôme cyclotomique ?
  - Calculer  $\int_0^1 \frac{dx}{\sqrt{1+x^2}}$ .
  - Calculer la somme  $\sum_{n=1}^{+\infty} \frac{1}{n^2}$ .
  - Quelle est la dimension de l'espace vectoriel des matrices antisymétriques de  $M_n(\mathbf{R})$  ?
  - Calculer  $\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx$ .
  - Calculer le volume de la sphère.
  - Quel est le nom de la quadrique  $x^2 + y^2 - z^2 = 1$  ?
  - Calculer  $\sum \frac{(-1)^n}{2n+1}$ . On justifiera soigneusement le calcul.
  - Identifier la quadrique  $xy + yz + zx = 1$ .
  - Soit  $f \in L^1(\mathbf{R}, dx)$ . Montrer que  $\lim_{t \rightarrow +\infty} \int_{\mathbf{R}} e^{itx} f(x) dx = 0$ .
  - Calculer  $\sum_{n \geq 1} \frac{n}{2^n}$ .
  - Décomposer en cycle la permutation  $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 1 & 7 & 6 & 5 & 4 & 2 \end{pmatrix}$ .
  - Un segment de longueur  $a$  étant donné, construire à la règle et au compas un segment de longueur  $\sqrt{a}$ .
  - Soit  $f \in C^1(\mathbf{R}_+, \mathbf{R})$  décroissante, telle que  $\lim_{x \rightarrow +\infty} f(x) = 0$ . Montrer que  $\forall t > 0, \int_0^N f(x) e^{itx} dx$  a une limite quand  $N \rightarrow +\infty$ .
  - Trouver un vecteur  $\begin{pmatrix} n_1 \\ n_2 \end{pmatrix} \in \mathbf{Z}^2$  tel que la famille  $\left\{ \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} \right\}$  forme une base de  $\mathbf{Z}^2$ .
  - Énoncer le théorème des zéros isolés.

## Commentaires

### Développement asymptotique de l'intégrale $\int \frac{\sin t}{t} dt$

Ce résultat peut être présenté comme développement dans les leçons relatives aux intégrales, aux développements asymptotiques, et à la convergence des séries. *Leçons concernées :*

### Formule d'inversion de Fourier

Ce résultat peut être présenté comme développement dans les leçons relatives à l'intégration. Il peut être mentionné dans les leçons portant sur les séries de Fourier. *Leçons concernées :*

### Somme des inverses des carrés des entiers

Cette somme est pour la première fois calculée par Euler en 1735. La preuve classique passe par les séries de Fourier. Ici, on a renversé la présentation, afin d'obtenir des informations sur ces séries. Ces considérations ont leur place dans les leçons traitant d'intégration, de séries, mais aussi d'espace de Hilbert. *Leçons concernées :*

### Formule de Stirling

On peut proposer en développement le calcul de l'asymptotique par la méthode de Laplace ; ou bien on peut présenter l'encadrement de la factorielle, sans se préoccuper du calcul de la constante  $\sqrt{2\pi}$ . *Leçons concernées :*

### Sur les séries semi-convergentes

Ce résultat présente une approche relativement nouvelle sur des points très classiques du programme. Il est donc conseillé de le mentionner dans les leçons ayant trait aux séries et à l'intégration. Il est sans doute un peu court pour fournir un développement. *Leçons concernées :*

### Quaternions et rotations

Les quaternions sont un thème très classique du programme d'intégration. L'approche proposée ici cherche à faire un parallèle avec les nombres complexes ; un quaternion est présenté comme un nombre complexe généralisé dont la partie imaginaire est un vecteur de  $\mathbf{R}^3$ . On peut parler des quaternions dans les leçons relatives au groupe orthogonal, et proposer le calcul de l'angle de la composée de deux rotations. *Leçons concernées :*

### Un exemple de marche aléatoire

Un certain nombre de leçons sont consacrées à la théorie des séries ; il est important de mentionner quelques applications "extra-mathématiques" de cette théorie. En voici une, qui peut se recycler dans les leçons de probabilité. *Leçons concernées :*

### La loi forte des grands nombres $L^2$

Ce résultat peut être présenté en développement dans les leçons de probabilité, mais aussi dans les leçons en lien avec le concept de convergence ou de suite de fonctions. *Leçons concernées :*

### Le théorème de la limite centrée

Ce théorème est un résultat important de la théorie des probabilités. Il s'agit de bien comprendre son énoncé, ainsi que l'application qui suit sa démonstration. *Leçons concernées :*

### **Construction du pentagone régulier**

Cette construction peut figurer dans les leçons de géométrie, ainsi que dans les leçons relatives aux racines de l'unité. Ne pas hésiter à faire la figure au tableau ; le jury met à la disposition des candidats règle et compas. *Leçons concernées :*

### **Polyèdres platoniciens**

On peut en parler dans les leçons qui portent sur la combinatoire, les groupes orthogonaux ou la géométrie dans l'espace. La construction des solides platoniciens est le sujet du livre XIII des éléments d'Euclide. L'adresse suivante : <http://www.ics.uci.edu/~eppstein/junkyard/euler/> présente dix-neuf preuves de la formule d'Euler. *Leçons concernées :*

### **Sous-groupes de $S_4$**

Il s'agit d'un petit diagramme destiné à faire prendre conscience du fait que le groupe  $S_4$  est moins innocent qu'il en a l'air. On peut tester ses connaissances en théorie des groupes, en cherchant les sous-groupes distingués de  $S_4$ , les sous-groupes abéliens, classer les sous-groupes à conjugaison près, déterminer les groupes quotients etc.

### **Trace, formes quadratiques et extensions de corps**

On peut présenter en développement la première preuve de la non-dégénérescence de la trace, et donner en application l'indépendance sur  $\mathbf{Q}$  des racines des nombres premiers. *Leçons concernées :*

### **Sur la diagonalisation des matrices $2 \times 2$**

Les leçons d'algèbre linéaires proposées par les candidats à l'agrégation manquent souvent d'originalité. Le problème de la diagonalisation continue des matrices  $2 \times 2$  est très naturel, mais n'est mentionné nul part dans la littérature. Le résultat pour les matrices à coefficients réels peut être proposé en développement ; on peut dire un mot sur le cas complexe. *Leçons concernées :*

### **Théorie de Galois**

La théorie de Galois n'est pas au programme de l'agrégation. Le théorème de l'élément primitif peut par contre être proposé en développement. Ce petit texte montre comment déduire facilement le théorème "fondamental" de la théorie de Galois à partir de cet élément primitif et de la décomposition des polynômes symétriques.

### **Calculs en maple**

Dans les leçons d'analyse, il est important de mentionner ce qu'un logiciel est en mesure de calculer.

### **Approximation de la loi gaussienne**

Le développement divergent de la gaussienne s'obtient en faisant des intégrations par partie ; il mérite de figurer dans les leçons ayant trait aux séries et à l'intégration, mais aussi dans certaines leçons de probabilités. *Leçons concernées :*

### **Exercices d'oral**

Ces exercices classiques sont souvent posés à l'oral de l'agrégation. Il faut donc savoir les faire vite et bien.