QUELQUES NOTES (2024 – 2025 ET ANTIDATÉS)

GARNIER MATHIAS

Résumé. On compile diverses questions (parfois en ayant trouvé une réponse). À terme, il se peut que certaines de ces notes servent à l'écriture d'un petit livre. Rached, ne t'en fais pas ! Je ne fais pas (que) des bêtises ! (Et, je sais ! il y a plein de problèmes de mise en page.)

Table des matières

Première partie 1. **Foutoir**

1. Vague bestiaire de distances et normes

Aucune prétention à un quelconque sérieux dans la présentation. On recense simplement quelques distances et normes vues par-ci par-là (qui proviennent essentiellement de mes TD).

1.1. **Les archi classiques.**
Commençons par quelques exemples dans $\mathbf{R}^2$.

(d_0) (Distance discrète) Pour X, Y dans $\mathbf{R}^2$, $d_0(X,Y) = 1$ si $X \neq Y$, 0 sinon.

(d_1) (Distance associée à la norme 1) Pour $X = (x_1, x_2)$, $Y = (y_1, y_2)$ dans $\mathbf{R}^2$, on pose :
$$d_1(X,Y) = |x_1 - y_1| + |x_2 - y_2|.$$

(d_2) (Distance euclidienne) Pour $X = (x_1, x_2)$, $Y = (y_1, y_2)$ dans $\mathbf{R}^2$, on pose :
$$d_2(X,Y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.$$

(d_3) (Distance associée à la norme max) Pour $X = (x_1, x_2)$, $Y = (y_1, y_2)$ dans $\mathbf{R}^2$, on pose :
$$d_3(X,Y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}.$$

(d_4) (Distance SNCF) Pour $X = (x_1, x_2)$, $Y = (y_1, y_2)$ dans $\mathbf{R}^2$, on pose :
$$d_4(X,Y) = \begin{cases} d_2(X,Y) & \text{si } X \text{ et } Y \text{ sont alignés avec } P_0 = (0,0), \\ d_2(X,P_0) + d_2(P_0,Y) & \text{sinon.} \end{cases}$$

(d_5) Pour $X = (x_1, x_2)$, $Y = (y_1, y_2)$ dans $\mathbf{R}^2$, on pose :
$$d_5(X,Y) = |x_1 - y_1| + \delta(x_2, y_2).$$

Plus généralement, sur $\mathbf{R}^n$ par exemple, on peut trouver des familles de normes (et on reconnaîtra évidemment la proximité avec certaines des distances ci-dessus). La plus fameuse est celle des normes $p \in [1, +\infty]$:
$$N_p(x) = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$
si $p \neq +\infty$. Si $p = +\infty$, on définit $N_\infty(x) = \max\{|x_i|\}$.

L'équivalent sur les espaces $L^p(\mathbf{R}^d)$ est alors sans surprise :
$$||f||_p = \left(\int_{\mathbf{R}^d} |f(\mathbf{x})| \mathrm{d}\mathbf{x} \right)^{1/p}.$$
La norme $p = +\infty$ étant également à considérer séparément. Ce cas est toujours un peu particulier à traiter. Par exemple, sur $\mathcal{C}^0([0,1], \mathbf{R})$, la norme infinie est $\sup\limits_{x \in [0,1]} |f(t)|$.

On ne dira pas grand chose des produits scalaires (alors que pourtant... !), pensons pour l'instant seulement à la norme de Frobenius dérivant du produit scalaire $\mathrm{Tr}({}^tAB)$ sur $\mathcal{M}_n(\mathbf{R})$.

1.2. **Un peu moins courantes.**

- Considérons l'espace des suites à valeurs dans $[0,1]$. On y définit la distance suivante entre deux suites $x = (x_n)_n$ et $y = (y_n)_n$:
$$d(x,y) = \sum_{n=0}^{\infty} 2^{-n} |x_n - y_n|.$$
- Soient $a_1, a_2, \dots, a_n$ des nombres réels strictement positifs, alors on dispose de la norme $N : \mathbf{R}^n \to \mathbf{R}$ suivante :
$$N(x_1, x_2, \dots, x_n) = a_1|x_1| + a_2|x_2| + \dots + a_n|x_n|. \tag{1}$$
- Soient $n+1$ nombres réels distincts $x_0, x_1, \dots, x_n$. On dispose des normes suivantes sur $\mathbf{R}_n[X]$:
$$||P|| = \sum_{i=0}^{n} |P(x_i)|, \quad ||P|| = \max_{k \leq n} |P^k(0)|, \quad ||P|| = |P(0)| + \int_0^1 |P'(t)| \mathrm{d}t.$$

- Sur $\mathcal{C}^0([-1,1])$, on dispose des normes suivantes :
$$N : f \mapsto \int_{-1}^{1} |f(x)|\cos(\pi x/2)\mathrm{d}x, \quad N : f \mapsto \int_{-1}^{1} |f(x)|x^2\mathrm{d}x.$$
- Sur $\mathcal{C}^1([0,1])$, on dispose de la norme suivante :
$$N : f \mapsto \int_{0}^{1} |f(x)| + |f'(x)|\mathrm{d}x.$$
- Sur $\mathcal{C}^2([0,1])$, on dispose des normes suivantes :
$$N : f \mapsto \|f\|_\infty + \|f'\|_\infty, \quad N : f \mapsto \|f\|_\infty + \|f'\|_\infty + \|f''\|_\infty.$$
À leur sujet, notamment leur équivalence, on se réfèrera à la note 37.

On voit aisément comment généraliser la définition de certaines de ces normes afin d'obtenir toujours plus de normes (penser aux bornes, à de la composition...).

2. Représentations modulaires de $\mathrm{GL}_2(\mathbf{F}_p)$

Soit p un nombre premier, on note $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$. Laurent Berger et Sandra Rozensztajn proposent d'étudier certaines des représentations du groupe $\mathrm{GL}_2(\mathbf{F}_p)$ tout au long d'une quarantaine d'exercices. On se propose ici de rédiger les réponses à ces questions tout en ajoutant divers commentaires et développements. Ce travail fera (peut-être, qui sait...) l'objet, *de manière détournée*, d'une conférence étudiante fin 2024.

2.1. **Le groupe $\mathrm{GL}_2(\mathbf{F}_p)$ et ses sous-groupes.** Notons $\mathrm{M}_2(\mathbf{F}_p)$ l'ensemble des matrices 2×2 à coefficients dans $\mathbf{F}_p$. Sauf mention expresse du contraire, on adopte la convention de réserver l'utilisation des lettres a, b, c, d à des éléments de $\mathbf{F}_p$. Ainsi, un élément de $\mathrm{M}_2(\mathbf{F}_p)$ est représenté par la matrice $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Par exemple, l'identité Id est sans surprise $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Le déterminant de M se définit comme suit : $\det(M) = ad - bc$. Une matrice M est dite inversible s'il existe un élément $N \in \mathrm{M}_2(\mathbf{F}_p)$ tel que $MN = \mathrm{Id}$. L'ensemble des matrices inversibles de taille 2×2 à coefficients dans $\mathbf{F}_p$ est noté $\mathrm{GL}_2(\mathbf{F}_p)$

2.1. **Exercice.** Vérifier que $\det(MN) = \det(M)\det(N)$. Calculer $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$. Prouver que M est inversible si, et seulement si, $\det(M) \neq 0$.

Notons le résultat suivant :
$$MN = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix} \tag{2}$$
Ainsi :
$$\det(MN) = (aa' + bc')(cb' + dd') - (ab' + bd')(ca' + dc').$$
En comparaison, on a :
$$\det(M)\det(N) = (ad - bc)(a'd' - b'c')$$
On conclut alors en développant directement chacun des termes. En spécialisant l'équation 2 ($a' = d$, $b' = -b$, $c' = -c$ et $d' = a$), on trouve (en utilisant la commutativité de $\mathbf{F}_p$) :
$$MN = \begin{pmatrix} ad - bc & 0 \\ 0 & -cb + da \end{pmatrix} = \det(M)\mathrm{Id}.$$

On a donc exhibé, à un facteur multiplicatif près, l'inverse d'un élément M de $\mathrm{M}_2(\mathbf{F}_p)$. On conclut alors directement que M est inversible si, et seulement si, $\det(M) \neq 0$ et l'on a :
$$M^{-1} = \frac{1}{\det(M)}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$
On reconnaît la transposée de la comatrice de M.

On en profite pour remarquer que l'application $\det : (\mathrm{GL}_n(\mathbf{F}_p), \times) \to (\mathbf{F}_p^\times, \times)$ est un morphisme de groupes, avec $n \in \mathbf{N}$. Plus généralement, on peut remplacer $\mathbf{F}_p$ par un corps k arbitraire. On prouve sans soucis que ce morphisme de groupes est surjectif. Pour cela, considérer la matrice avec un élément quelconque de $k^\times$ sur le premier élément de la diagonale et des 1 autrement sur la diagonale (on impose des 0 hors diagonale). On peut construire certains des sous-groupes de $\mathrm{GL}_n(\mathbf{F}_p)$ par ruissellement, par exemple le groupe $\mathrm{SL}_n(\mathbf{F}_p)$ se définit comme le noyau de l'application déterminant. Donc, un élément de $\mathrm{SL}_2(\mathbf{F}_p)$ est toujours de déterminant 1.

Attelons-nous à quelques premières considérations de dénombrement.

2.2. **Exercice.** Prouver que $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{M}_2(\mathbf{F}_p)$ appartient à $\mathrm{GL}_2(\mathbf{F}_p)$ si, et seulement si, $\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right) \neq \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$ et $\left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right)$ n'est pas multiple de $\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)$. En déduire $\left|\mathrm{GL}_2(\mathbf{F}_p)\right| = (p^2-1)(p^2-p)$ et $\left|\mathrm{SL}_2(\mathbf{F}_p)\right| = p(p^2-1)$.

Procédons par équivalence. Conséquemment à l'exercice 1, l'élément $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ est dans $\mathrm{GL}_2(\mathbf{F}_p)$ si, et seulement si, $ad-bc \neq 0$. Si $\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right) \neq \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$ et $\left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right)$ n'est pas multiple de $\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)$, alors le déterminant est forcément non nul. On conclut réciproquement par contraposée. Il suffit alors de compter le nombre d'éléments de $\mathrm{M}_2(\mathbf{F}_p)$ vérifiant la caractérisation précédente. Sur la première colonne, on a p^2-1 choix possibles (on a seulement à exclure le cas $\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$). Sur la deuxième colonne, il faut ensuite exclure tous les multiples de la première colonne, ce qui donne p^2-p choix possibles. En tout, on obtient donc bien le résultat suivant :

$$\left|\mathrm{GL}_2(\mathbf{F}_p)\right| = (p^2-1)(p^2-p).$$

Plus généralement, en opérant le même raisonnement, on trouve :

$$\left|\mathrm{GL}_n(\mathbf{F}_p)\right| = (p^n-1)(p^n-p)\dots(p^n-p^{n-1}).$$

Un argument savant (pour nous pour l'instant) permet alors de déduire le cardinal de $\mathrm{SL}_2(\mathbf{F}_p)$. On dispose de la suite exacte suivante :

$$1 \to \mathrm{SL}_2(\mathbf{F}_p) \to \mathrm{GL}_2(\mathbf{F}_p) \xrightarrow{\det} (\mathbf{F}_p)^\times \to 1.$$

Un raisonnement équivalent permet de traiter le cas n quelconque. Focalisons nous sur le cas $n=2$. Il s'en déduit alors que :

$$\left|\mathrm{SL}_2(\mathbf{F}_p)\right| = \frac{\left|\mathrm{GL}_2(\mathbf{F}_p)\right|}{p-1} = \frac{p(p^2-1)(p-1)}{p-1} = p(p^2-1).$$

Remarquons qu'une écriture équivalente est la suivante : $p(p^2-1) = p^3(1-1/p^2)$. Plus généralement, et cela ravira les théoriciens des nombres, on a pour N entier plus grand que un :

$$\left|\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})\right| = N^3 \prod_{p|N}(1-1/p^2).$$

Colmez, dans ses *Éléments d'analyse, d'algèbre (et de théorie des nombres)*, Annexe B, donne un résultat plus général.

2.3. **Proposition.** *Si* K *est un corps fini de cardinal q, et si $n \geq 2$, alors :*

$$\left|\mathrm{SL}_n(\mathrm{K})\right| = q^{n-1}\prod_{k=2}^{n}(q^n-q^{n-k}) = q^{n^2-1}\prod_{k=2}^{n}(1-1/q^k).$$

2.4. **Exercice.** Soit $\mathfrak{S}_n$ le groupe des permutations de l'ensemble $\{1,2,\dots,n\}$.

(1) L'exercice 2.2 implique que $\left|\mathrm{GL}_2(\mathbf{F}_2)\right| = 6$. Combien de groupes de cardinal 6 connaissez vous ?

On pense évidemment d'abord à nos classiques : $\mathbf{Z}/6\mathbf{Z} \cong \mathbf{F}_2 \times \mathbf{F}_3$, $\mathfrak{S}_3$. On a désormais $\mathrm{GL}_2(\mathbf{F}_2)$ mais on verra à la question suivante qu'on le connaissait en fait déjà. Pour montrer qu'il n'y a en réalité, à isomorphisme près, que deux groupes d'ordre 6, on se reportera utilement aux notes de K. Conrad, *Groups of order* 4 *and* 6.

(2) Il y a 3 points non nuls dans $\mathbf{F}_2^2$. L'action de $\mathrm{GL}_2(\mathbf{F}_2)$ sur ces points induit une application $\mathrm{GL}_2(\mathbf{F}_2) \to \mathfrak{S}_3$. Prouver que cette application est un isomorphisme.

Énumérons les points de $\mathbf{F}_2^2$:

$$(0,0) \quad (0,1) \quad (1,0) \quad (1,1).$$

Il y a donc en effet 3 points non nuls.

(3) Identiquement, on a $\left|\mathrm{GL}_2(\mathbf{F}_3)\right| = 48$. Il y a 4 droites dans $\mathbf{F}_3^2$. Utiliser cela pour montrer qu'il existe une application surjective $\mathrm{GL}_2(\mathbf{F}_3) \to \mathfrak{S}_4$ telle que son noyau soit réduit à $\{\pm \mathrm{Id}\}$.

(4) (Plus difficile) Trouver une application surjective $\mathrm{GL}_2(\mathbf{F}_5) \to \mathfrak{S}_5$.

(5) Dans chacun des cas précédents, quelle est l'image de $\mathrm{SL}_2(\mathbf{F}_p)$?

Notons que la conjecture naturelle : il existe une application surjective de $\mathrm{GL}_2(\mathrm{F}_7)$ dans $\mathfrak{S}_6$ (et ainsi de suite) ne semble pas tenir (à vérifier !).

Introduisons $\mathcal{Z} = \{\left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right) \text{ avec } a \in \mathbf{F}_p^\times\}$.

2.5. **Exercice.** Montrer que $\mathcal{Z}$ est le centre de $\mathrm{GL}_2(\mathbf{F}_p)$.

Comme tout élément de $\mathcal{Z}$ s'écrit $a\mathrm{Id}$ pour un certain $a \in \mathbf{F}_p^\times$, il est évident que $\mathcal{Z}$ est inclus dans $\mathrm{GL}_2(\mathbf{F}_p)$. Réciproquement, montrons que $\mathcal{Z}(\mathrm{GL}_2(\mathbf{F}_p)) \subset \mathcal{Z}$. Un élément du centre commute avec tous les éléments de $\mathrm{GL}_2(\mathbf{F}_p)$ donc a fortiori avec les matrices $e_{i,j}$ constituées d'un 1 en position (i,j) et des 0 partout autrement. En comparant $Ae_{i,j}$ avec $e_{i,j}A$ pour A dans le centre, on obtient que A est forcément de la même forme qu'un élément de $\mathcal{Z}$. Donc $\mathcal{Z} = \mathcal{Z}(\mathrm{GL}_2(\mathbf{F}_p))$.

Remarquons que comme le sous-groupe $\mathcal{Z}$ est distingué dans $\mathrm{GL}_2(\mathbf{F}_p)$ (car c'est le centre), on peut former le groupe quotient $\mathrm{GL}_2(\mathbf{F}_p)/\mathcal{Z}$ et ainsi définir le groupe projectif linéaire sur $\mathbf{F}_p$.

N.B. J'espère revenir sur ce point plus tard (cf. 2.2) : comprendre les petits (i.e. les sous-groupes) de $\mathrm{GL}_2(\mathbf{F}_p)$. Ce n'est vraiment pas chose facile semblerait-il. Moi qui voulait comprendre la situation sur GL_2 pour ensuite attaquer GL_3... Une discussion avec Stéphane Lamy s'impose.

Je me suis cassé la tête un certain temps sur les quatre exercices suivants. Il m'aura fallu du temps pour me rendre compte d'une, je crois, typographie et de l'utilisation d'une notation du *Graduate Algebra* de Lang sans le dire.

2.6. **Exercice.** Prouver qu'il existe $g \in \mathbf{F}_p^\times$ tel que $\mathbf{F}_p^\times = \{1, g, \ldots, g^{p-2}\}$. Montrer que toute fonction multiplicative $f : \mathbf{F}_p^\times \to \mathbf{F}_p^\times$ est de la forme $x \mapsto x^r$, avec $0 \leq r \leq p-2$.

On définit l'ensemble des matrices triangulaires supérieures B inclues dans $\mathrm{GL}_2(\mathbf{F}_p)$. On ne parlera pas de sous-groupes de Borel, mais ce serait une voie possible. On pose ensuite U comme étant l'ensemble des matrices de la forme $\left(\begin{smallmatrix} 1 & a \\ 0 & 1 \end{smallmatrix}\right)$ avec $a \in \mathbf{F}_p$. De plus, on introduit $w = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. (C'est là que l'on suppose qu'il y a une typographie dans le document de Berger et Rozensztajn.)

2.7. **Exercice.** Prouver que toute fonction multiplicative $f : B \to \mathbf{F}_p^\times$ est de la forme suivante pour $0 \leq r, s \leq p-2$:

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \mapsto a^r \cdot d^s.$$

2.8. **Exercice.** (Décomposition de Bruhat) Montrer que $\mathrm{GL}_2(\mathbf{F}_p) = B \cup UwB$ et plus précisément que $\mathrm{GL}_2(\mathbf{F}_p)$ est l'union disjointe de ces deux ensembles.

2.9. **Exercice.** Prouver que toute fonction multiplicative $f : \mathrm{GL}_2(\mathbf{F}_p) \to \mathbf{F}_p^\times$ est de la forme $M \mapsto \det(M)^r$ pour r vérifiant $0 \leq r \leq p-2$.

2.2. **Les petits de $\mathrm{GL}_2(\mathbf{F}_p)$.** C'est un petit aparté à titre purement culturel. On dédiera *peut-être* une note à la justification de ces résultats. On se reporte essentiellement au Landesman, *Subgroups of* $\mathrm{GL}_2(\mathbf{F}_\ell)$ et à *Finite p-Irregular Subgroups of* $\mathrm{PGL}(2,k)$ de X. Faber.

2.3. **Représentations de groupes.** On part pour un cours éclair de théorie des représentations !

3. Injectivité de la transformée de Fourier

Dès que l'on sait deux ou trois choses de la transformée de Fourier, c'est (presque) évident, immédiat et gratuit. Le problème étant que la transformée de Fourier n'est plus étudiée dans le parcours recherche niveau Licence à Paul Sabatier. Laure Coutin ne mentait donc pas quand elle a dit que ce n'était pas un exercice compliqué mais plus fait ! À vrai dire, on avait déjà vu avec Julien Royer que la transformée de Fourier était injective (encore fallait-il lire entre les lignes !).

N'étant pas spécialement à l'aise avec l'énoncé probabiliste de la formule d'inversion de Fourier suivant, je donne une version dans $\mathrm{L}^1(\mathbf{R}^d)$.

3.1. **Proposition.** *Soit X une variable aléatoire. Si la fonction caractéristique ϕ_X est intégrable, alors X admet une densité donnée par :*

$$f(x) = \frac{1}{2\pi} \int_{\mathbf{R}} e^{-itx} \phi_X(t) \mathrm{d}t. \tag{3}$$

3.2. **Proposition.** *Soit $f \in \mathrm{L}^1(\mathbf{R}^d)$ tel que $\widehat{f} \in \mathrm{L}^1(\mathbf{R}^d)$. Alors, pour presque tout $x \in \mathbf{R}^d$, on a :*

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} e^{ix\cdot\xi} \widehat{f}(\xi) \mathrm{d}\xi. \tag{4}$$

L'injectivité de la transformée de Fourier dans L^1 est alors une conséquence directe de la linéarité de la transformée et de la formule d'inversion. En effet, comme $\mathscr{F} : f \mapsto \widehat{f}$ est une application linéaire, il suffit de montrer que son noyau est réduit à l'élément nul. Soit f dans le noyau, par la formule d'inversion (4) il vient automatiquement $f \equiv 0$. Donc la transformée de Fourier est injective.

Tout le problème repose donc (comme souvent) sur ce que l'on s'autorise à utiliser, ce que l'on considère connu. Suivons au moins la preuve donnée par Royer de la formule d'inversion. Ce petit travail est nécessaire ! (Je vais tenter de ne négliger aucun détail.)

4. Fourier et une équation différentielle

Cela provient d'un exercice donné par Royer et d'une tentative de généralisation naïve.

4.1. Problème initial.

4.1. **Défi.** Soit $f : \mathbf{R} \to \mathbf{R}$ une fonction 2π-périodique continûment différentiable et soit α un réel non nul. On considère l'équation différentielle

$$x'(t) + \alpha x(t) = f(t).$$

Trouver une solution 2π-périodique de cette équation en écrivant $x(t)$ et $f(t)$ sous forme de séries de Fourier trigonométriques. Appliquer ce résultat au cas où $\alpha = 1$ et où

$$f(t) = \begin{cases} \left(t - \dfrac{\pi}{2}\right)^2 & \text{si } t \in [0, \pi[, \\ -\left(t - \dfrac{3\pi}{2}\right)^2 + \dfrac{\pi^2}{2} & \text{si } t \in [\pi, 2\pi[. \end{cases}$$

Cela se résout bien, il faut être scrupuleux, méticuleux et un peu attentionné tout de même.

5. En avant la dualité :

6. Brouille de modules !

7. Un langage diagrammatique et séquençoïdal

8. Recueil de suites exactes

9. Éducation façon catégories

10. Formule des traces pour les groupes finis

11. Topologie p-adique

12. Un peu d'algèbre différentielle

Au début d'un de leur cours, Peter Scholze et Dustin Clausen citent David Mumford : «*[Algebraic geometry] seems to have acquired the reputation of being esoteric, exclusive, and very abstract, with adherents who are secretly plotting to take over all the rest of mathematics. In one respect this last point is accurate*». On ne va pas faire ici de la géométrie algébrique, quoiqu'il soit possible de parvenir jusqu'à quelques unes de ses ramifications... Néanmoins, l'entrée et la mise en branle de l'algèbre et de sa mécanique va nous permettre de conquérir des paysages que l'on pouvait penser déraisonnablement loin et inaccessibles à l'algèbre. Un exemple de première mesure est le théorème de Liouville où l'algèbre différentielle explique l'obstruction à ce que certaines fonctions n'admettent pas de primitives dites élémentaires. Notre but sera celui ci : faire accoster l'algèbre en des rivages qui peuvent lui paraître étrangers. Une attention toute particulière sera portée à ce que tous les détails soient précautionneusement traités. Néanmoins, des erreurs plus ou moins flagrantes et graves pourront subsister. Elles sont le fait de l'auteur et grand joie lui ferait d'apprendre leur existence. Ne pas hésiter tout signalement d'erreur, d'imprécision ou encore toute remarque, développement. Signalons enfin qu'une inspiration nette provient des écrits de Kaplansky, Ritt, Kolchin ainsi que de nombreux cours (Bardavid, Tremblay, Hecky et Bartlett, Conrad, Hardouin...). Néanmoins, un point d'honneur a été mis à composer des exercices, imaginer des développements, envisager diverses applications... bref, avoir une première expérience de recherche mais également et surtout d'enseignement.

Ce petit texte se veut être une formalisation de *Primitive où te caches-tu ? Une introduction à l'algèbre différentielle* précédemment écrit.

12.1. **Définitions et propriétés élémentaires.** Dans la suite, k désignera invariablement un corps (plus de précisions proviendront du contexte). Nous travaillerons en caractéristique zéro, à moins qu'une mention expresse du contraire soit faite.

12.1. **Définition.** Un **anneau différentiel** (A, ∂) est un anneau commutatif unitaire A muni d'une dérivation $\partial : A \to A$ vérifiant, pour tous éléments f et g de A :

$$\partial(f+g) = \partial(f) + \partial(g), \qquad \partial(fg) = \partial(f)g + f\partial(g). \tag{5}$$

Lorsque A est un corps, on parlera alors du **corps différentiel** (A, ∂). De manière identique si A est un k-espace vectoriel, on parlera de l'**espace vectoriel différentiel** (A_k, ∂).

Par défaut, on supposera l'anneau A commutatif. On pourra au besoin ne pas s'en accommoder, notamment lorsque l'on travaillera avec des espaces matriciels (dont, un exemple de dérivation utilise le commutateur, nous y reviendrons).

Plus généralement, on comprend alors comment construire une **structure algébrique différentielle**. Profitons encore un petit peu de notre naïveté et de notre incrédulité avant d'introduire des notions similaires pour les modules. On cherche pour l'instant seulement à comprendre la base de la théorie. Assurons nous tout d'abord que la notion de dérivation fasse sens et qu'elle existe. Pour cela arrêtons nous sur différents exemples.

12.2. **Exemples.**

12.2.1. *Dérivation triviale.* La moins intéressante, bien évidemment ! Elle est définie pour tout anneau (commutatif et unitaire) A et vaut 0 en toutes circonstances. Néanmoins, son existence nous rassure au plus haut point. Nous aurons l'occasion d'y revenir plus tard mais, par exemple, sans la dérivation triviale, l'ensemble de toutes les dérivations sur un corps commutatif k ne formerait pas un k-espace vectoriel (nous le vérifierons plus tard).

Comme on peut munir tout anneau de la dérivation triviale, on peut alors voir la théorie des anneaux comme un cas particulier de l'algèbre différentielle.

12.2. **Défi.** Montrer que la seule dérivation possible sur $\mathbf{Z}$ ou $\mathbf{Q}$ est la dérivation triviale. Qu'en est-il pour les corps $\mathbf{R}$ et $\mathbf{C}$?

Soit ∂ une dérivation sur $\mathbf{Z}$ ou $\mathbf{Q}$. On souhaite montrer que pour tout élément x de $\mathbf{Z}$ ou $\mathbf{Q}$, on a $\partial(x) = 0$. En d'autres termes, on souhaite montrer que le noyau de ∂ est $\mathbf{Z}$ ou $\mathbf{Q}$. Plus généralement, soit $\partial : A \to A$ une dérivation, peut-on caractériser les anneaux A tels que $\ker(\partial) = A$? Nous reviendrons sur cette question lorsque l'on introduira l'anneau ou le corps des constantes d'une dérivation.

Commençons par remarquer que $\partial(0+0) = \partial(0) = \partial(0) + \partial(0)$. Donc $\partial(0) = 0$. De plus, $0 = \partial(0) = \partial(x - x) = \partial(x) + \partial(-x)$. De fait, restreindre le problème à un élément de $\mathbf{Z}_+$ ou $\mathbf{Q}_+$ ne fait perdre aucune généralité. On constate que $\partial(1) = \partial(1 \cdot 1) = \partial(1)1 + 1\partial(1)$. Donc $\partial(1) = 0$. Finalement, pour $n \in \mathbf{Z}_+$, comme $\partial(n) = \partial(n-1+1) = \partial(n-1) + \partial(1)$, c'est-à-dire $\partial(n-1) = \partial(n)$, on peut en déduire que la seule dérivation possible sur $\mathbf{Z}$ est la dérivation triviale. C'est à peu près le même combat que l'on doit mener sur $\mathbf{Q}$. Restreignons de nouveau à $\mathbf{Q}_+$ sans perte de généralité et soit $x = p/q$ un élément de cet ensemble (avec $p \in \mathbf{Z}_+$ et $q \in \mathbf{N} \setminus \{0\}$). Alors $\partial(x) = \partial(p/q) = \partial(p)1/q + p\partial(1/q)$. Or, on vient d'obtenir que $\partial(p) = 0$. Pour conclure, il suffit d'écrire que $0 = \partial(1) = \partial(q/q) = \partial(q)1/q + q\partial(1/q)$ (avec q différent de zéro par définition). Autrement dit, on obtient que $\partial(1/q) = 0$. Donc, la seule dérivation possible sur $\mathbf{Q}$ est la dérivation triviale.

Les cas de $\mathbf{R}$ et $\mathbf{C}$ relèvent moins de l'évidence. On se retrouve dans un premier temps face au problème de l'expression des nombres irrationnels (qui se définissent par négation de ce qu'est un nombre rationnel plutôt qu'intrinsèquement) ; à moins que... On conjecture que le raisonnement suivant est probant. Il suffit d'écrire un nombre rationnel sous la forme de son développement décimal (par exemple). À cet effet, on a pour un élément x réel (avec I et les a_i correctement choisis) :

$$\partial(x) = \partial\Big(\sum_{i \in I} a_i 10^i\Big) = \sum_{i \in I} \Big(\partial(a_i)10^i + a_i\partial(10^i)\Big).$$

Les a_i et 10^i étant entiers, on en conclut par le raisonnement précédent que $\partial(x) = 0$ pour tout réel x. En ce qui concerne $\mathbf{C}$, nous avons uniquement besoin de déterminer l'image de i par une dérivation quelconque. On remarque que $0 = \partial(-1) = \partial(i \cdot i) = \partial(i)i + i\partial(i)$. Donc $\partial(i)i = -i\partial(i)$. On divise par i. Ceci nous force à constater que $\partial(i) = 0$.

12.3. **Défi.** Existe-t-il une dérivation ∂ n'étant pas la dérivation triviale telle que $\partial \circ \partial = 0$? $\partial \circ \partial \circ \partial = 0$? etc... ? La composition alors obtenue est-elle une dérivation ?

Pour éviter toute confusion, convenons de préciser que $\partial\circ\partial = 0$ signifie que, pour tout élément x de la structure algébrique considérée, $\partial\circ\partial(x) = 0$. Par commodité, notons ∂^n la composition de n dérivations. Remarquons que si $\partial^2 = 0$, alors $\partial^k = 0$ pour tout $k \geq 2$. La réciproque n'est en revanche pas vraie.
Il n'est pas compliqué de construire une telle dérivation pour toute valeur de n entière. Considérons $\mathbf{R}_n[X]$ l'espace vectoriel des fonctions polynômiales de degré au plus n. Muni de la dérivation naturelle ∂, sur laquelle nous allons revenir ci-dessous, $(\mathbf{R}_n[X], \partial)$ est un $\mathbf{R}$-espace vectoriel différentiel (on le vérifiera ci-après). Alors, dans $\mathbf{R}_1[X]$ composer deux fois ∂ donne zéro, dans $\mathbf{R}_2[X]$ composer trois fois ∂ donne zéro... Peut-on obtenir une réponse similaire pour un corps ou un anneau différentiel ?

Nous entrons dans la deuxième partie de la question. Se peut-il qu'un tel ∂^n soit une dérivation ? Nous allons d'abord conduire les calculs dans le cas général.

Le défi précédent était motivé par une considération première provenant de l'homologie mais également par une discussion sur les dérivations localement nilpotentes sur lesquelles nous nous attarderons plus tard.

Dérivation naturelle. ss

Relèvement des coefficients

Dérivation matricielle

Dérivation galoisienne

Dérivation de Pincherle

Une dérivation en deux variables

Mélanges et combinaisons de dérivations

Nous avons laissé de côté des objets ou exotiques ou bien n'entrant pas exactement, à première vue, dans le cadre de l'algèbre différentielle. On pensera en premier lieu à la dérivation arithmétique introduite par Joyal. Pour l'instant, nous ne faisons que nous introduire à ces diverses dérivations, un peu de concret mais rien encore de totalement sérieux.

12.2.2. *Premières propriétés.* **Constantes d'une dérivation**

Petits calculs

Règle de Leibniz

Tordre la règle de Leibniz

Dérivation logarithmique

12.2.3. *Caractéristique p.*

12.3. Analyse classique.

12.3.1. *Analyse réelle à une variable.*

12.3.2. *Densité et passage à la limite.*

12.3.3. *Analyse réelle à plusieurs variables.*

12.3.4. *Un peu d'analyse complexe.*

12.4. Opération inverse : primitivation.

12.5. Extension et prolongement de dérivations.

12.6. Polynômes différentiels et Opérateurs différentiels.

13. Le terrible escalier des lois de probabilité sans densité

14. Tribu borélienne pour $X \neq \mathbf{R}^d$

En cours, on a uniquement considéré sérieusement le cas de la tribu engendrée par les ouverts de $\mathbf{R}^d$, notée $\mathcal{B}(\mathbf{R}^d)$. Rien n'empêche de remplacer $\mathbf{R}^d$ par un espace topologique X (relativement) quelconque.

15. Analyse complexe et arithmétique : au revoir les déboires avec Dirichlet

16. Fan club de Florian Thiry

Cette section ne concerne en rien des mathématiques. On inventorie seulement (certains) des hauts faits d'armes de Florian Thiry. Pendant un temps, le doute était permis, nous pensions que les faits étaient de Laurence Fourier (nous avons été induits en erreurs par ses t-shirts : Los Pollos Hermanos, Harakis Surf, Dunder Mifflin, Faker SKT...).

TD 0. *Qu'elle vous soit une lumière dans les endroits ténébreux, quand toutes les autres s'éteindront.*

TD 1. *Il n'est pas prudent d'écarter de ses calculs un dragon vivant, quand on est près de lui.*

TD 2. *Le vaste monde est tout autour de vous : vous pouvez vous enfermer, mais vous ne pouvez pas le clôturer pour toujours.*

TD 3. *Trop longtemps tu m'as traqué, trop longtemps je t'ai évité... Plus maintenant.*

TD 4. *La connaissance tue l'action, pour agir il faut être obnubilé par l'illusion.*

TD 5. *Tu apprendras que beaucoup de vérités auxquelles nous tenons dépendent avant tout de notre propre point de vue.*

17. La course aux 24, les résultats démontrés

Les 24 ($= (42)^{-1}$) problèmes confectionnés par Alain Troesch recèlent de résultats non triviaux. On répertorie ici ceux démontrés, c'est une vraie course ! À vrai dire, il y en a bien plus de 24 mais 24 est un beau nombre. Ce qui justifie son utilisation.

18. Une vieille intégrale d'un certain 26 mai 2021

19. L'ogre Green-Riemann et son oncle Stokes

20. Opérateurs elliptiques sur un graphe, quelques mots

21. Des sommes directes aux intégrales directes

22. Théorème spectral façon Royer

En parlant simplement, Royer m'a conduit et fait entrevoir ce que pouvait être le théorème spectral version plus forcément pour les enfants.

23. Un Léon Marchand topologique

Il faut le voir pour le croire : Ignat Radu nageant dans le vide face au tableau nous expliquant ce qu'est un plongement topologique.

24. Le couac des quaternions

25. Incursion des suites exactes en analyse

26. Complétons, complétons !

27. Transport de structure

28. The discrete Laplacian (a first encounter)

Everyone is well-aware of the definition of the Laplacian (the author is not able to make any link with differential geometry, Riemannian manifolds and the so-called *Laplace-Beltrami operator*). Let f be a twice differentiable multi-variable function, we then define the Laplace operator :

$$\Delta f = \sum_{i=1}^{n} \partial_i^2 f, \text{ for instance, if } n = 3 : \Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}. \tag{6}$$

This definition can be motivated from different perspectives. For instance, the Laplace operator in all respects is equal to $\mathrm{Tr}(\mathrm{Hess}(f))$ or even $\mathrm{div}(\overrightarrow{\mathrm{grad}} f)$. Using the appropriate definition may sometime be fruitful. That is why, even if we'll be inclined to use a « canonical » discrete Laplacian, we'll construct in many ways the discrete Laplacian operator such that we mimic the definition and properties of the Laplacian in a graph-theoretic context.

How to derive a discrete version of the Laplace operator ? More generally, is there limits to expectations from going to global to local (or local to global) ? We have some evidences that graphs are limit objects of continuous ones. Thus, we expect to proceed the following way : starting from the definition of the derivative operator, one discretizes it and combines adequately to get a discretized Laplacian operator. We won't discuss in which manner a graph can serve as a mesh for a manifold and scaling limits.

Recall the usual real derivative definition. Let f be one-variable differentiable function. We then have for some real number x :

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}. \tag{7}$$

To compute a one-variable Laplace operator, we need the second derivative. It expresses as (for a twice-differentiable function) :

$$\begin{aligned} f''(x) &= \lim_{h\to 0} \frac{f'(x+h) - f'(x)}{h} \\ &= \lim_{h\to 0} \frac{f(x+2h) - f(x+h) - (f(x+h) - f(x))}{h^2} \\ &= \lim_{h\to 0} \frac{f(x+2h) - 2f(x+h) + f(x)}{h^2}. \end{aligned}$$

Moreover, if f is twice continuously differentiable, we find that the second derivative is more or less a *mean* between its neighbouring points :

$$f''(x) = \lim_{h\to 0} \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} = \lim_{h\to 0} \frac{2}{h^2}\left(\frac{f(x+h) + f(x-h)}{2} - f(x)\right). \tag{8}$$

On a discrete graph, we would like to reproduce this situation : the discrete Laplacian should be a function of its neighbouring points. For instance, the set $\mathbf{Z}$ can be endowed by a graph structure where each vertex is an integer coordinate points of the real line. The action of the discrete Laplacian on 0 shall depend of the point itself, 1 and -1 :

$$\cdots \quad \bullet_{-3} \quad \bullet_{-2} \quad \bullet_{-1} \quad \bullet_{0} \quad \bullet_{1} \quad \bullet_{2} \quad \bullet_{3} \quad \cdots$$

A natural question arise : at the time, the Laplacian varies only with neighbouring points, can one asks more ? For instance, (a modified discrete) Laplacian at 0 could be a function of $0, 1, -1, 2, -2$ or $0, 1, -1, 2, -2, 3, -3$... We will call such operator an *unscreened Laplacian operator or order* r, where r is the size of the neighbourhood. A study will be made in the next part. The name has been chosen having in mind the Slater screening effect on electrons.

For the sake of completeness, we need to derive a similar expression for a multi-variable function. The partial derivative is then, for some $\mathbf{x}$ in $\mathbf{R}^d$ and $i \in [\![1, d]\!]$:

$$\partial_i f(\mathbf{x}) = \lim_{h\to 0} \frac{f(\mathbf{x}+h) - f(\mathbf{x})}{h} = \lim_{h\to 0} \frac{f(x_1, \ldots, x_i + h, \ldots, x_d) - f(x_1, \ldots, x_d)}{h}. \tag{9}$$

The second partial derivative is :

$$\begin{aligned} \partial_i^2 f(\mathbf{x}) &= \lim_{h\to 0} \frac{\partial_i f(\mathbf{x}+h) - \partial_i f(\mathbf{x})}{h} \\ &= \lim_{h\to 0} \frac{f(\mathbf{x}+2h) - f(\mathbf{x}+h) - \big(f(\mathbf{x}+h) - f(\mathbf{x})\big)}{h^2} \\ &= \lim_{h\to 0} \frac{f(\mathbf{x}+2h) - 2f(\mathbf{x}+h) + f(\mathbf{x})}{h^2} \end{aligned}$$

Nothing changes (so, we expect the dimension changes only computations not the way the Laplacian behaves).

For the same reason as before, if f is a twice continuously differentiable function, we find :

$$\partial_i^2 f(\mathbf{x}) = \lim_{h\to 0} \frac{f(\mathbf{x}+h) - 2f(\mathbf{x}) + f(\mathbf{x}-h)}{h^2} = \lim_{h\to 0} \frac{2}{h^2}\left(\frac{f(\mathbf{x}+h) + f(\mathbf{x}-h)}{2} - f(\mathbf{x})\right). \tag{10}$$

The Laplacian now writes as :

$$\Delta f = \sum_{i=1}^{d} \partial_i^2 f(\mathbf{x}) = \lim_{h\to 0} \frac{2}{h^2} \sum_{i=1}^{d} \left(\frac{f(\mathbf{x}+h) + f(\mathbf{x}-h)}{2} - f(\mathbf{x})\right) \tag{11}$$

(the inversion of the sum and the limit is possible because the sum is finite). Writing the coordinates, we obtain :

$$\Delta f = \lim_{h\to 0} \frac{2}{h^2} \sum_{i=1}^{d} \left(\frac{f(x_1,\dots,x_i+h,\dots,x_d) + f(x_1,\dots,x_i-h,\dots,x_d)}{2} - f(x_1,\dots,x_d) \right). \tag{12}$$

For instance, for $d = 3$, we get :

$$\begin{aligned}\Delta f = \lim_{h\to 0} \frac{2}{h^2} \Bigg(& \frac{f(x_1+h,x_2,x_3) + f(x_1-h,x_2,x_3)}{2} \\ & + \frac{f(x_1,x_2+h,x_3) + f(x_1,x_2-h,x_3)}{2} \\ & + \frac{f(x_1,x_2,x_3+h) + f(x_1,x_2,x_3-h)}{2} - 3f(x_1,x_2,x_3) \Bigg)\end{aligned}$$

$$\begin{aligned}\Delta f = \lim_{h\to 0} \frac{6}{h^2} \Bigg(& \frac{f(x_1+h,x_2,x_3) + f(x_1-h,x_2,x_3)}{6} \\ & + \frac{f(x_1,x_2+h,x_3) + f(x_1,x_2-h,x_3)}{6} \\ & + \frac{f(x_1,x_2,x_3+h) + f(x_1,x_2,x_3-h)}{6} - f(x_1,x_2,x_3) \Bigg).\end{aligned}$$

Up to a scalar constant, the definition of the discrete Laplacian operator can be modelled on equation 12 with a parameter h equal to 1. For an arbitrary d, one would obtain :

$$\Delta_{\text{discretized}} f = \sum_{i=1}^{d} \left(\frac{f(x_1,\dots,x_i+1,\dots,x_d) + f(x_1,\dots,x_i-1,\dots,x_d)}{2d} \right) - f(x_1,\dots,x_d).$$

A major problem on a graph is the following : a vertex can have more (or less) than two neighbours. That is why we will care about this aspect in next section by considering the size of the neighbourhood of a vertex. (Need to be done.)

29. Exponentielle d'une matrice

Ce petit texte avait été écrit pour un physicien (Raphaël Decan de Chatouville) de ma promotion.

L'exponentielle (réelle) possède bon nombre de caractéristiques qui la rendent intéressante (bijection réciproque du logarithme népérien, équation fonctionnelle, dérivation, solution d'équations différentielles ordinaires, théorie de Fourier, trigonométrie...). Rien n'empêche alors de supputer qu'elle puisse être également particulièrement utile en algèbre (linéaire). Mais alors, comment la définir ? Parallèlement et de manière symétrique à ce que l'on peut faire en analyse réelle, on est particulièrement tenté de définir l'exponentielle d'une matrice $A \in \mathcal{M}_n(\mathbf{k})$, avec $\mathbf{k}$ égal à $\mathbf{R}$ ou $\mathbf{C}$, comme une série entière :

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}. \tag{13}$$

Sans surprise cette définition ne contredit pas le cas où A est un scalaire (cad., $A \in \mathcal{M}_1(\mathbf{k}) = \mathbf{k}$). On s'attend de plus à ce que nombre de résultats vrais sur k possèdent un équivalent sur $\mathbb{M}_n(\mathbf{k})$.

L'un des résultats les plus évidents tient à la convergence : quand peut-on calculer l'exponentielle d'une matrice (sous-entendu obtenir une matrice à coefficients tous finis) ? Remarquons que se ramener à une étude de la convergence normale de l'exponentielle d'une matrice revient à se battre avec une fonction que l'on connaît bien : l'exponentielle usuelle. En effet, si l'on dispose d'une norme $|||\cdot|||$ pour laquelle l'inégalité suivante est vraie :

$$\left|\left|\left| \frac{A^k}{k!} \right|\right|\right| \le \frac{1}{k!} |||A|||^k \tag{14}$$

alors la convergence normale serait assurée (notons que l'équivalence des normes joue en notre faveur en dimension finie). En effet, en sommant ensuite sur l'ensemble des valeurs entières k, on obtient la convergence normale (uniforme et simple) de l'exponentielle d'une matrice et même la continuité de l'application étudiée. On a d'ailleurs montré que $||\exp(A)|| \leq \exp(||A||)$:

$$||\exp(A)|| = \left|\left|\sum_{k=0}^{\infty} \frac{A^k}{k!}\right|\right| \leq \sum_{k=0}^{\infty} \left|\left|\frac{A^k}{k!}\right|\right| = \sum_{k=0}^{\infty} \frac{||A||^k}{k!} = \exp(||A||). \tag{15}$$

Dans cette note explicative, on ne s'attarde pas trop sur l'existence et le pourquoi du comment de la norme sous-multiplicative $|||\cdot|||$. On renvoie au Gourdon, *Algèbre – Topologie sur les endomorphismes* pour quelques précisions mais l'on réserve une note future à l'étude de ces objets.

Bien que nous sachions possible le calcul de l'exponentielle d'une matrice, il n'est absolument pas évident de le mener en pratique. Dans certains cas particuliers, on peut se débrouiller sans faire preuve de beaucoup d'astuce. Par exemple dans le cas d'une matrice nilpotente A, l'exponentielle est une somme finie et *a fortiori* un polynôme en A. Plus généralement un résultat similaire existe pour une matrice $A \in \mathcal{M}_n(\mathbf{k})$.

29.1. **Théorème.** *Pour toute matrice A de $\mathcal{M}_n(\mathbf{k})$, il existe un polynôme P_A à coefficients dans $\mathbf{k}$ tel que $\exp(A) = P_A(A)$.*

Ce théorème nous permet d'affirmer que A et $\exp(A)$ commutent ensemble :

$$A\exp(A) = \exp(A)A. \tag{16}$$

De manière plus générale, lorsque l'on étudie la conservation (ou non) de certaines propriétés on se rend compte que la commutativité peut entrer un jeu. Un exemple flagrant se trouve dans l'équation fonctionnelle. Nous sommes, chez nos amis complexes et réels, bien habitués au résultat suivant $\exp(z_1 + z_2) = \exp(z_1)\exp(z_2)$, avec $(z_1, z_2) \in \mathbf{C}$.

On peut, par exemple, prouver un résultat similaire sur $\mathcal{M}_n(\mathbf{k})$.

29.2. **Théorème.** *Soient A et B deux matrices de $\mathcal{M}_n(\mathbf{k})$ qui commutent. Alors,*

$$\exp(A + B) = \exp(A)\exp(B) = \exp(B)\exp(A). \tag{17}$$

Démonstration. On commencer en établissant une preuve classique. Ensuite, on introduira un schéma de preuve qui tient trait pour trait des idées que l'on retrouvera en théorie de Lie (on se réfère à la preuve du théorème I.4).

On ne traite que l'égalité $\exp(A+B) = \exp(A)\exp(B)$, la seconde égalité se traite identiquement. Commençons par rappeler les forces en présence :

$$\exp(A)\exp(B) = \left(\sum_{k=0}^{\infty} \frac{A^k}{k!}\right)\left(\sum_{k=0}^{\infty} \frac{B^k}{k!}\right). \tag{18}$$

Comme les séries sont normalement convergentes, on en prend le produit de Cauchy :

$$\left(\sum_{k=0}^{\infty} \frac{A^k}{k!}\right)\left(\sum_{k=0}^{\infty} \frac{B^k}{k!}\right) = \sum_{k=0}^{\infty} c_k, \text{ avec } c_k = \sum_{p=0}^{k} \frac{A^p}{p!}\frac{B^{k-p}}{(k-p)!}. \tag{19}$$

On remarque que $\frac{1}{k!}\binom{k}{p} = \frac{1}{p!(k-p)!}$. De ce fait :

$$c_k = \frac{1}{k!}\sum_{p=0}^{k} \binom{k}{p} A^p B^{k-p}. \tag{20}$$

Comme A et B commutent, on reconnaît la formule du binôme de Newton. On conclut que :

$$\exp(A)\exp(B) = \sum_{k=0}^{\infty} \frac{1}{k!}(A + B)^k = \exp(A + B). \tag{21}$$

Une autre preuve, aux idées plus profondes peut être proposée. On va suivre le document précédemment cité. □

Cette seconde preuve nous amènera à donner une nouvelle définition de l'exponentielle. Nous définirons, dans une autre note explicative, ce que peut être l'application exponentielle entre une algèbre et un groupe de Lie. Le cas des groupes de Lie matriciels est particulièrement satisfaisant puisque la définition coïncide avec l'exponentielle de matrice que nous considérons ici.

Revenons à nos problèmes actuels, peut-on espérer une formule dans le cas où A et B ne commutent pas forcément entre eux ? On pointe seulement l'existence de la formule BCH (Baker-Campbell-Hausdorff).

29.3. **Proposition.** (cf. Neeb-Hilgert) *Soient X et Y deux matrices de $\mathcal{M}_n(\mathbf{k})$ telles que leur norme soit strictement inférieure à $\frac{1}{2}\log\left(2-\frac{\sqrt{2}}{2}\right)$. On pose $\Psi(z)=\frac{1-e^{-z}}{z}$ avec z un nombre complexe. Alors :*

$$\log(\exp(X)\exp(Y)) = X + \int_0^1 \Psi(\exp(\mathrm{ad}X)\exp(\mathrm{ad}Y))Y\mathrm{d}t. \tag{22}$$

avec l'opérateur adjoint à définir, de même pour le logarithme.

Concrètement, on trouve la formule suivante :

$$\exp(X)\exp(Y) = \exp\Big(X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}[X,[X,Y]]+\frac{1}{12}[Y,[Y,X]]+\dots\Big) \tag{23}$$

où l'on utilise le commutateur $[X,Y]=XY-YX$.

Nous aurons tout le temps d'éclairer ces diverses notions dans un futur proche. Pour le moment, rappelons nous seulement qu'il convient de faire attention à l'ordre des termes !

Grâce au théorème 29.2, on peut prouver que l'élément $\exp(A)$ est inversible. En effet, il suffit de prendre $B=-A$ pour obtenir :

$$\mathrm{Id} = \exp(0) = \exp(A+(-A)) = \exp(A)\exp(-A) = \exp(-A)\exp(A). \tag{24}$$

On a donc sans grande surprise (et nous en sommes contents puisque l'on retombe sur un résultat connu) : $\exp(A)^{-1}=\exp(-A)$. Continuons sur notre lancée et voyons un petit résultat amusant (qui ne nous surprendra guère).

Introduisons la fonction $\mathbf{R}\to\mathcal{M}_n(\mathbf{R})$ qui à t associe $\exp(tA)$. On peut montrer que c'est une application de classe C^∞ et en calculer la dérivée, qui vaut $A\exp(tA)$. Outre ses diverses propriétés "analytiques", l'exponentielle complexe se révèle bien utile dès que l'on pense réduction, diagonalisation et tout le tintouin ! On sent que l'exponentielle va s'avérer utile en regardant les deux résultats suivants.

29.4. **Proposition.** *Soit A un élément de $\mathcal{M}_n(\mathbf{k})$ et P un élément de $\mathrm{GL}_n(\mathbf{k})$. Alors :*

$$\exp(PAP^{-1}) = P\exp(A)P^{-1}. \tag{25}$$

La preuve ne pose aucun problème. On peut alors s'intéresser au cas particulier où l'exponentielle prend une matrice diagonale en argument.

29.5. **Proposition.** *Soit $D=\mathrm{Diag}(a_1,\dots,a_n)$ une matrice diagonale de taille n, alors*

$$\exp(D) = \mathrm{Diag}(e^{a_1},\dots,e^{a_n}). \tag{26}$$

Dans la mesure où l'on sait calculer sans problème aucun la puissance d'une matrice diagonale, la preuve ne pose de nouveau aucun problème.

Peut-être est-il temps de voir nos premiers calculs d'exponentielle de matrices ? Nous allons utiliser SageMath (*a free open-source mathematics software system*).

```
A_1 = Matrix([[1, 1], [1, 1]]); exp(A_1)
A_2 = Matrix([[0, -2*pi], [2*pi, 0]]); exp(A_2)
A_3 = Matrix([[1/e, pi], [-1, 1]]); exp(A_3)
```

On trouve :

$$\exp(A_1) = \begin{pmatrix} 1/2e^2+1/2 & 1/2e^2-1/2 \\ 1/2e^2-1/2 & 1/2e^2+1/2 \end{pmatrix},\ \exp(A_2)=\mathrm{Id},\ \exp(A_3)=\text{horreur} \tag{27}$$

En règle générale, il faut s'attendre à tomber sur des horreurs... à moins que la vie soit (trop) bien faite et que l'on ait été mis sur une voie particulièrement plaisante ! Par exemple, dans le livre de Hall on peut trouver le résultat suivant :

$$A = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix},\ \exp(A) = \begin{pmatrix} \cos(a) & -\sin(a) \\ \sin(a) & \cos(a) \end{pmatrix} \tag{28}$$

30. Un problème de L1 mettant sur les traces de l'analyse complexe

On se propose de résoudre un exercice de la plaquette de TD (j'ai écrit ça quand j'étais en première année, ça remonte !).

30.1. **Défi.** Soit $c \in \mathbb{C}$ avec $|c| < 1$ et soit $z \in \mathbb{C}$.

1. Montrer que $|z+c| \leq |1+\overline{c}z|$ si et seulement si $|z| \leq 1$. Quand a-t-on égalité ?
2. Soient $D = \{z \in \mathbb{C}, |z| \leq 1\}$ le disque unité et $C = \{z \in \mathbb{C}, |z| = 1\}$ le cercle unité. Montrer que l'application $f : D \to D$ qui a z fait correspondre $\dfrac{z+c}{1+\overline{c}z}$ est bien définie et que c'est une bijection vérifiant $f(C) = C$.

Question 1 :
Calculons le carré des deux quantités suivantes $|z+c| et |1+\overline{c}z|$:

$$\begin{aligned}|z+c|^2 &= (z+c)\overline{(z+c)}\\ &= (z+c)(\overline{z}+\overline{c})\\ &= z\overline{z} + z\overline{c} + c\overline{z} + c\overline{c}\\ &= |z|^2 + z\overline{c} + c\overline{z} + |c|^2.\end{aligned}$$

$$\begin{aligned}|1+\overline{c}z|^2 &= (1+\overline{c}z)\overline{(1+\overline{c}z)}\\ &= (1+\overline{c}z)(1+c\overline{z})\\ &= 1 + c\overline{z} + \overline{c}z + \overline{c}zc\overline{z}\\ &= 1 + c\overline{z} + \overline{c}z + |c|^2|z|^2.\end{aligned}$$

Soustrayons les résultats obtenus :

$$\begin{aligned}|1+\overline{c}z|^2 - |z+c|^2 &= 1 + |c|^2|z|^2 - |z|^2 - |c|^2\\ &= (|z|^2-1)(|c|^2-1).\end{aligned}$$

Par hypothèse, l'on a $|c| < 1$. On en déduit que $|c|^2 - 1$ est négatif. Ainsi, pour que la différence $|1+\overline{c}z|^2 - |z+c|^2$ soit positive ou nulle, il est nécessaire et suffisant que $|z|^2 - 1$ soit également négatif ou nul, ce qui est équivalent à avoir $|z| \leq 1$, comme désiré.

Il reste néanmoins à conclure en la forme voulue : l'on a démontré l'équivalence suivante :

$$|1+\overline{c}z|^2 - |z+c|^2 \geq 0 \iff |z| \leq 1.$$

Une ultime manipulation achève la démonstration :

$$|1+\overline{c}z|^2 - |z+c|^2 \geq 0 \iff |1+\overline{c}z|^2 \geq |z+c|^2.$$

En raison de la valeur absolue, prendre la racine des deux côtés de l'inéquation ne brise pas la chaîne d'équivalences.

L'on remarque enfin que l'inégalité devient une égalité dès lors que c est nul et que z est un point quelconque du cercle unité.

Question 2 :
Montrons tout d'abord que l'application $f : D \to D$ est bien définie.

Pour ce faire, l'on doit montrer que le dénominateur ne s'annule jamais. Montrons donc que : $1 + \overline{c}z \neq 0$. C'est strictement équivalent à montrer que : $\overline{c}z \neq -1$. Raisonnons sur le module de $\overline{c}z$ (étant différent de 1). Par hypothèse, l'on a $|c| < 1$ et $|z| \leq 1$ car z est un point de D. Ainsi, $|c||z| = |\overline{c}||z| < 1$, donc, *a fortiori* différent de 1. On en déduit donc que l'application est bien définie.

Pour montrer que l'application f est une bijection de D dans D, on doit montrer que f est injective et surjective.

Montrons tout d'abord l'injectivité de f. Pour ce faire, prenons z et z' deux éléments de D, et, supposons que $f(z) = f(z')$. Ainsi, $\dfrac{z+c}{1+\overline{c}z} = \dfrac{z'+c}{1+\overline{c}z'}$. Simplifions l'équation : $(z+c)(1+\overline{c}z') = (z'+c)(1+\overline{c}z)$. L'on a :

$$(z+c)(1+\overline{c}z') = z + z\overline{c}z' + c + |c|^2z'.$$

$$(z'+c)(1+\overline{c}z) = z' + z'\overline{c}z + c + |c|^2 z.$$

On réinjecte alors pour obtenir, après simplification : $z+|c|^2z' = z'+|c|^2z$. On obtient donc : $z'-z = |c|^2(z'-z)$. Raisonnons par l'absurde et supposons que z' est différent de z. On divise des deux côtés de l'équation par $z'-z$. Il en résulte que $|c|^2 = 1$, ce qui est impossible. Donc $z = z'$. L'application f est ainsi injective.

Montrons enfin que f est une application surjective. Pour ce faire, prenons un point Z du disque unité, on veut montrer qu'il existe toujours un point z du disque unité tel que $f(z) = Z$. On a cherche donc à montrer qu'il existe toujours z tel que : $\dfrac{z+c}{1+\overline{c}z} = Z$. Il suffit alors d'isoler z. On obtient donc : $z = \dfrac{Z-c}{1-\overline{c}Z}$. (Il convient de ne pas oublier de vérifier que $1-\overline{c}Z$ ne s'annule jamais.)

On a donc démontré que l'application $f : D \to D$ est une bijection. Il ne reste alors plus qu'à établir que $f(C) = C$. On a :

$$\begin{aligned} f(C) &= \left\{ \frac{z+c}{1+\overline{c}z} \middle| z \in \mathbb{C} \right\} \\ &= \left\{ \frac{e^{i\theta}+c}{1+\overline{c}e^{i\theta}} \middle| \theta \in [0, 2\pi[\right\}. \end{aligned}$$

Calculons le module d'un élément quelconque de $f(C)$:

$$\begin{aligned} \left| \frac{e^{i\theta}+c}{1+\overline{c}e^{i\theta}} \right|^2 &= \left(\frac{e^{i\theta}+c}{1+\overline{c}e^{i\theta}} \right) \overline{\left(\frac{e^{i\theta}+c}{1+\overline{c}e^{i\theta}} \right)} \\ &= \left(\frac{e^{i\theta}+c}{1+\overline{c}e^{i\theta}} \right) \left(\frac{e^{-i\theta}+\overline{c}}{1+ce^{-i\theta}} \right) \\ &= \frac{1+e^{i\theta}\overline{c}+ce^{-i\theta}+c\overline{c}}{1+ce^{-i\theta}+e^{i\theta}\overline{c}+c\overline{c}} \\ &= 1. \end{aligned}$$

De ce fait, le module d'un élément quelconque de $f(C)$ est 1. Ceci permet de conclure que $f(C) = C$.

On peut désormais être tenté de "tracer" une telle application (on utilise un module python spécialement adapté (*cplot*)). Voyons ce que nous obtenons pour différentes valeurs de c :

On discerne une forme de *dynamique (complexe)*.

31. Noyau, Image, Pseudo-inverse et une propriété universelle !

Soit E un espace vectoriel de dimension quelconque et $f \in \mathcal{L}(E)$.

31.1. **Définition.** On dit que $g \in \mathcal{L}(E)$ est un **pseudo-inverse** de f si :

$$f \circ g = g \circ f,\ f \circ g \circ f = f, \text{ et } g \circ f \circ g = g.$$

L'endomorphisme f est alors dit **pseudo-inversible** s'il admet un pseudo-inverse.

Par la suite, s'il n'y a pas d'ambiguïté, on se permet de simplifier la notation $f \circ g$ en fg.

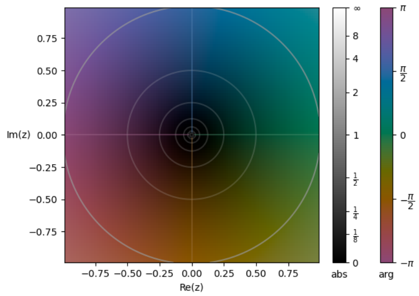
On se propose de résoudre un exercice sur le **pseudo-inverse d'Azumaya-Drazin**. Nous allons prouver le théorème 31.2. On en profitera pour prouver des résultats intermédiaires ou auxiliaires qui permettent d'éclairer la situation.

31.2. **Théorème.** *L'endomorphisme f est pseudo-inversible si et seulement si* $\ker(f) = \ker(f^2)$ *et* $\mathrm{Im}(f) = \mathrm{Im}(f^2)$.

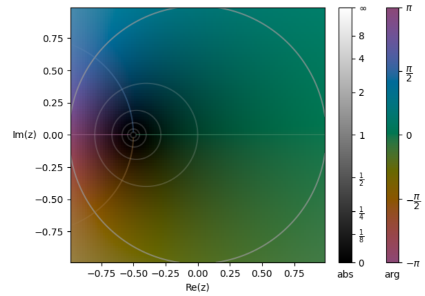
1. *Montrer que si f est pseudo-inversible, il admet un unique pseudo-inverse que l'on notera $f^\sharp$.* Procédons par l'absurde et supposons que f admette deux pseudo-inverses différents g et g'. La propriété de commutativité de f et g' nous assure que $fg' = g'f$. Composons par g' puis utilisons la deuxième propriété des pseudo-inverses : $g'fg' = g'g'f = g'g'fgf$. Développons et réarrangeons : $g'g'fgf = g'fg'fg = g'fg$. Ainsi $g' = g'fg$. On conclut alors à une absurdité en remarquant que :

$$g' = g'fg = g'fgfg = fg'fgg = fgg = gfg = g.$$

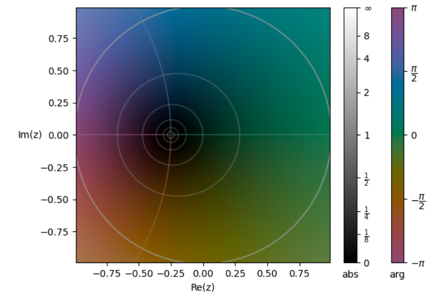
Contradiction. Un pseudo-inverse est unique (s'il existe).



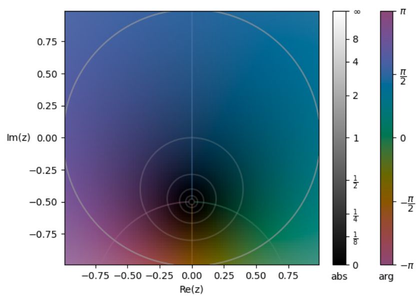
(A) $c = 0$



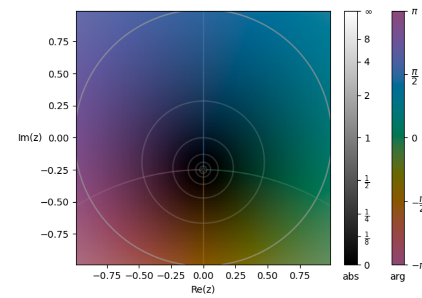
(B) $c = 1/2$



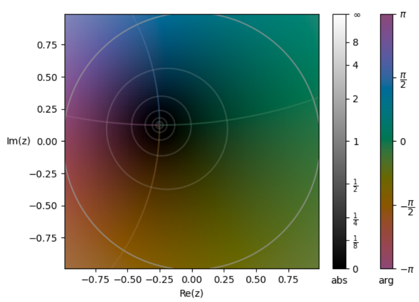
(C) $c = 1/4$



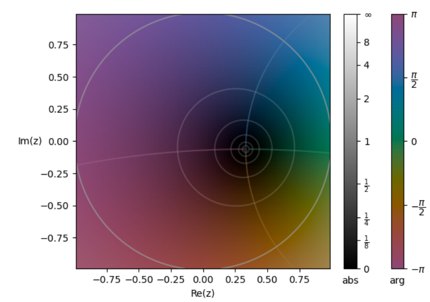
(D) $c = i/2$



(E) $c = i/4$



(F) $c = 1/4 - i/8$



(G) $c = -1/3 + i/16$

2.a) ***Monter que si f est inversible, alors il est pseudo-inversible. Déterminer son pseudo-inverse $f^\sharp$.*** On rappelle que f est inversible si et seulement si il existe un élément g de $\mathcal{L}(E)$ tel que $gf = \mathrm{Id}$ et $fg = \mathrm{Id}$. Nous n'allons montrer qu'une des deux égalités, l'autre se montre de manière similaire. Supposons que le pseudo-inverse de f soit l'inverse au sens usuel (un tel inverse existe car f est supposée inversible). Il ne reste qu'à vérifier qu'un tel pseudo-inverse convient :

— $ff^{-1} = \mathrm{Id} = f^{-1}f$,
— $ff^{-1}f = f \circ \mathrm{Id} = f$, et
— $f^{-1}ff^{-1} = f^{-1} \circ \mathrm{Id} = f^{-1}$.

2.b) ***Même question si désormais f est un projecteur.*** Rappelons qu'un projecteur f vérifie, par définition, $f^2 = f$. Nous allons montrer que le pseudo-inverse d'un tel f est lui-même. Vérifions que les inégalités soient bien satisfaites avec un tel choix :

— $ff = ff$, et
— $fff = ff = f$ ce qui suffit à remplir les deux dernières conditions de la définition 31.1.

3. ***Plus généralement, on suppose que*** $\ker(f) \oplus \mathrm{Im}(f) = E$***. Montrer que f est pseudo-inversible.*** Une vraie horreur cette question. Impossible sans cette fameuse indication : *on pourra utiliser la propriété universelle de la somme directe pour définir* $f^\sharp$ *séparément sur* $\ker(f)$ *et* $\mathrm{Im}(f)$*, après avoir remarqué que* f *induisait un automorphisme de* $\mathrm{Im}(f)$. Bon, essentiellement, on découvre... Suivons les indications !

31.3. **Lemme.** *Dans les conditions de l'énoncé, la restriction de* f *à* $\mathrm{Im}(f)$ *induit un automorphisme de* $\mathrm{Im}(f)$*.*

Démonstration. Jusqu'alors, nous savons simplement que la restriction de f, notée par la suite $\tilde{f}$, est un endomorphisme. Nous devons montrer que $\tilde{f}$ est un isomorphisme. La surjectivité est tout d'abord évidente car l'image de l'application $\tilde{f}$ est égale à elle-même. Reste à montrer l'injectivité. Soit $x \in \ker(\tilde{f})$. Nous savons que x est également un élément un $\mathrm{Im}(\tilde{f})$ (rappelons sommairement que $\mathrm{Im}(\tilde{f})$ est également le domaine de définition de $\tilde{f}$). N'oublions pas que nous avions supposé $\ker(f) \oplus \mathrm{Im}(f) = E$ donc *a fortiori* que $\ker(\tilde{f}) \cap \mathrm{Im}(\tilde{f}) = \{0\}$. On en conclut alors que $\ker(\tilde{f}) \subset \{0\}$. En conclusion $\ker(\tilde{f}) = \{0\}$, donc $\tilde{f}$ est injective et par là même bijective (comme désiré). □

Nous disposons donc d'un automorphisme de $\mathrm{Im}(f)$. Étant un automorphisme, il est inversible donc *a fortiori* pseudo-inversible en vertu de la question *2.a)*. Le problème est "réglé" sur l'image de f.

Le problème se poursuit... Il faut comprendre comment se servir de cette "propriété universelle de la somme directe" (chose toute nouvelle pour moi). Un document semblant présenter cela d'une manière fort agréable est venu m'éclairer. (Un petit coup d'oeil au livre d'Ibrahim Assem chez Calvage & Mounet ne semble rien donner.)

L'idée va être de travailler au corps et séparément le noyau ainsi que l'image. Il faudra ensuite réussir à "recoller l'information" pour obtenir une information non plus simplement sur $\ker(f)$ et $\mathrm{Im}(f)$ pris séparément mais bien sur $\ker(f) \oplus \mathrm{Im}(f)$.

Expliquons en quoi consiste la propriété universelle de la somme directe (il semblerait que ces histoires de propriétés universelles, problèmes universels... aient un langage commun dans lequel ils s'expriment agréablement : la théorie des catégories).

31.4. **Proposition.** *(La proposition n'est pas énoncée dans toute sa généralité, on se restreint au cas présent.) Soient* $h_1 : \ker(f) \to E$ *et* $h_2 : \mathrm{Im}(f) \to E$*. Il existe une unique application linéaire* $u : \ker(f) \oplus \mathrm{Im}(f) \to E$ *telle que* $h_1 = u\iota_{\ker(f)}$ *et* $h_2 = u\iota_{\mathrm{Im}(f)}$*. La situation est résumée dans le diagramme ci-dessous.*

$$\begin{tikzcd}
\ker(f) \arrow[dr, "\iota_{\ker(f)}"] \arrow[ddr, bend right, "h_1"'] & & \mathrm{Im}(f) \arrow[dl, "\iota_{\mathrm{Im}(f)}"'] \arrow[ddl, bend left, "h_2"] \\
& \ker(f) \oplus \mathrm{Im}(f) \arrow[d, dashed, "u"] & \\
& E &
\end{tikzcd}$$

On remarque(rait si l'on faisait bien les choses) que $\ker(f) \oplus \mathrm{Im}(f)$ *est l'unique espace vectoriel avec une telle propriété.*

Précisons peut-être un peu qui sont ces applications ι_X $(X = \{\ker(f), \mathrm{Im}(f)\})$! Ce ne sont rien d'autre que les projections de l'espace de départ sur l'espace d'arrivée (un foncteur d'oubli ?). De ce fait, $\iota_{\ker(f)}$ est égal à $(x, 0)$ pour tout $x \in \ker(f)$ et $\iota_{\mathrm{Im}(f)}$ est égal à $(0, x)$ pour tout $x \in \mathrm{Im}(f)$.

Dans la mesure où $E = \ker(f) \oplus \mathrm{Im}(f)$, on peut refaire le diagramme sous la forme qui suit.

$$\begin{tikzcd}
\ker(f) \arrow[r, bend left, "\iota_{\ker(f)}"] \arrow[r, bend right, "h_1"'] & \ker(f) \oplus \mathrm{Im}(f) = E \arrow[loop below, dashed, "u"] & \mathrm{Im}(f) \arrow[l, bend right, "\iota_{\mathrm{Im}(f)}"'] \arrow[l, bend left, "h_2"]
\end{tikzcd}$$

4. ***Le but de cette question est de montrer la réciproque de la question précédente. On suppose désormais*** f ***pseudo-inversible.***

4.a) ***Monter que*** $ff^\sharp$ ***est un projecteur, de noyau*** $\ker(f^\sharp)$ ***et d'image*** $\mathrm{Im}(f)$***. Qu'en déduit-on ?*** Remarquons tout d'abord que si f est inversible, en vertu de la question 2, la réponse est évidente. Dans notre cas, f n'est pas supposée inversible mais pseudo-inversible.

L'application $ff^\sharp$ est effectivement un projecteur, comme on le vérifie :

$$(ff^\sharp)^2 = f\underbrace{f^\sharp f f^\sharp}_{f^\sharp} = ff^\sharp.$$

Reste à déterminer le noyau et l'image d'une telle application.

- Noyau de $ff^\sharp$: Montrons que $\ker(ff^\sharp) = \ker(f^\sharp)$. Il est tout d'abord évident que $\ker(f^\sharp)$ est contenu dans $\ker(ff^\sharp)$. Il reste alors à montrer que $\ker(ff^\sharp) \subset \ker(f^\sharp)$. Soit $x \in \ker(ff^\sharp)$, alors $ff^\sharp(x) = 0$. En composant par $f^\sharp$, on obtient que $f^\sharp(x) = f^\sharp ff^\sharp(x) = f^\sharp(0) = 0$. Donc x est un élément de $\ker(f^\sharp)$, ce qui permet de conclure.
- Image de $ff^\sharp$: On souhaite désormais prouver que $\mathrm{Im}(ff^\sharp) = \mathrm{Im}(f)$. Il est (de nouveau) évident que $\mathrm{Im}(ff^\sharp)$ est contenu dans $\mathrm{Im}(f)$. Montrons que $\mathrm{Im}(f) \subset \mathrm{Im}(ff^\sharp)$. Soit $x \in \mathrm{Im}(f)$, alors il existe $y \in E$ tel que $x = f(y)$. En composant par $f^\sharp$, on obtient que $f^\sharp(x) = f^\sharp f(y)$. On compose ensuite par f et finalement :
$$f^\sharp f(x) = ff^\sharp f(y) = f(y) = x.$$
Ainsi, x est un élément de $\mathrm{Im}(ff^\sharp)$, ce qui achève la preuve.

Grâce à la question précédente, nous sommes armés pour traiter avec des projecteurs ! On sait que son pseudo-inverse est lui même. Par voie de conséquence, l'image et le noyau de $ff^\sharp$ et de son pseudo-inverse sont les mêmes.

4.b) ***Montrer que*** $E = \ker(f) \oplus \mathrm{Im}(f)$***.*** Commençons par montrer que $\ker(f) \cap \mathrm{Im}(f) = \{0\}$. Il est (de nouveau) évident que $\{0\}$ est contenu dans $\ker(f)$ et $\mathrm{Im}(f)$. Démontrons alors que $\ker(f) \cap \mathrm{Im}(f) \subset \{0\}$. Soit un élément x de $\ker(f) \cap \mathrm{Im}(f)$. Si l'on compose $f(x) = 0$ par $f^\sharp$, on obtient : $f^\sharp f(x) = f^\sharp(0) = 0$. De plus, on sait qu'il existe un élément de E nommé y tel que $x = f(y)$, composons par $f^\sharp$ une nouvelle fois : $f^\sharp(x) = f^\sharp f(y)$. Si l'on réunit les informations, on remarque qu'en composant par f :

$$0 = f^\sharp f = ff^\sharp(x) = ff^\sharp f(y) = f(y).$$

On a évidemment utilisé les propriétés des pseudo-inverses. En conclusion, on a démontré que y est un élément de $\ker(f)$. Or, $x = f(y)$, ce qui impose la valeur de 0 à x. En définitive, le sous-espace vectoriel $\ker(f) \cap \mathrm{Im}(f)$ est réduit à $\{0\}$.

Il faut maintenant établir que $E = \ker(f) + \mathrm{Im}(f)$. Soit $x \in E$, on veut montrer que $x = x_1 + x_2$ avec x_1 un élément de $\ker(f)$ et x_2 un élément de $\mathrm{Im}(f)$. On remarque que $ff^\sharp(x) \in \mathrm{Im}(ff^\sharp) = \mathrm{Im}((ff^\sharp)^2)$ (l'égalité est due au fait que $ff^\sharp$ soit un projecteur). De plus, il existe y dans E tel que $ff^\sharp(x) = ff^\sharp(y)$. Montrons que l'on peut écrire x sous la forme $x_1 + x_2 = (x - ff^\sharp(y)) + (ff^\sharp(y))$. On remarque tout d'abord que $x_2 = ff^\sharp(y)$ est un élément de $\mathrm{Im}(ff^\sharp)$ donc *a fortiori* de $\mathrm{Im}(f)$ (en raison de la réponse à la question précédente). Il reste à montrer que $x_1 = x - ff^\sharp(y)$ est un élément de $\ker(f)$. Pour ce faire, composons par $ff^\sharp$, on remarque alors que :

$$ff^\sharp(x_1) = ff^\sharp(x - ff^\sharp(y)) = ff^\sharp(x) - (ff^\sharp)^2(y) = 0.$$

Ainsi, $x_1 \in \ker(ff^\sharp) = \ker(f^\sharp) = \ker(f)$. Ce qui conclut.

5. ***Démontrer le théorème 31.2.*** Drôle de coïncidence (ou pas), j'avais traité une question relativement similaire en partiel la veille... Nous allons avoir besoin des lemmes suivants :

31.5. **Lemme.** *Soit u un endomorphisme, alors* $\ker(u) = \ker(u^2)$ *si et seulement si* $\ker(u) \cap \mathrm{Im}(u) = \{0\}$.

Démonstration. La preuve de ce lemme et du suivant émane de la correction d'un partiel (L1 S2). On procède par double implication :

- $\ker(u) = \ker(u^2) \implies \ker(u) \cap \mathrm{Im}(u) = \{0\}$. Supposons que $\ker(u) = \ker(u^2)$. Soit x un élément de $\ker(u) \cap \mathrm{Im}(u)$, montrons que $x = 0$. Comme $x \in \mathrm{Im}(u)$, il existe y dans E tel que $x = u(y)$. Comme $x \in \ker(u)$, $0 = u(x) = u^2(y)$. Donc $y \in \ker(u^2)$. Or, par hypothèse, $\ker(u^2) = \ker(u)$, alors $y \in \ker(u)$. En conséquence, $0 = u(y) = x$. Ainsi $\ker(u) \cap \mathrm{Im}(u) = \{0\}$.
- $\ker(u) \cap \mathrm{Im}(u) = \{0\} \implies \ker(u) = \ker(u^2)$. Supposons désormais que $\ker(u) \cap \mathrm{Im}(u) = \{0\}$. Il est toujours vrai que $\ker(u) \subset \ker(u^2)$, il reste donc à montrer que $\ker(u^2)$ est contenu dans $\ker(u)$. Soit $x \in \ker(u^2)$, alors $0 = u^2(x) = u(u(x))$. Donc $u(x) \in \ker(u)$. Mais $u(x) \in \mathrm{Im}(u)$, donc $u(x)$ est un élément de $\ker(u) \cap \mathrm{Im}(u) = \{0\}$. Ce qui implique que $u(x) = 0$, donc que $x \in \ker(u)$, comme désiré.

□

31.6. **Lemme.** *L'égalité* $\mathrm{Im}(u) = \mathrm{Im}(u^2)$ *est satisfaite si et seulement si* $\ker(u) + \mathrm{Im}(u) = E$.

Démonstration. On procède de nouveau par double implication :

— $\mathrm{Im}(u) = \mathrm{Im}(u^2) \Longrightarrow \ker(u) + \mathrm{Im}(u) = E$. Supposons que $\mathrm{Im}(u) = \mathrm{Im}(u^2)$. Soit $x \in E$, alors $u(x) \in \mathrm{Im}(u) = \mathrm{Im}(u^2)$. Donc, il existe $y \in E$ tel que $u(x) = u^2(y)$. Posons $x_1 = u(y)$ et $x_2 = x - u(y)$, alors : $x = x_1 + x_2$ avec $x_1 \in \mathrm{Im}(u)$ par définition et $u(x_2) = u(x - u(y)) = u(x) - u^2(y) = 0$ donc $x_2 \in \ker(u)$. Ainsi, $E = \ker(u) + \mathrm{Im}(u)$.

— $\ker(u) + \mathrm{Im}(u) = E \Longrightarrow \mathrm{Im}(u) = \mathrm{Im}(u^2)$. Supposons que $E = \ker(u) + \mathrm{Im}(u)$. Soit $y \in \mathrm{Im}(u)$, alors il existe y dans E tel que $x = u(y)$. Comme E est la somme de $\ker(u)$ et $\mathrm{Im}(u)$, alors il existe y_1 dans $\mathrm{Im}(u)$ et y_2 dans $\ker(u)$ tels que $y = y_1 + y_2$. Comme $y_1 \in \mathrm{Im}(u)$, il existe z dans E tel que $y_1 = u(z)$. Donc $y = u(z) + y_2$. Ainsi :
$$x = u(y) = u(u(z) + y_2) = u^2(z) + u(y_2) = u^2(z) + 0 = u^2(z).$$
Donc x appartient à $\mathrm{Im}(u^2)$. Ainsi, $\mathrm{Im}(u) \subset \mathrm{Im}(u^2)$. Comme par ailleurs on a toujours $\mathrm{Im}(u^2) \subset \mathrm{Im}(u)$, le résultat suit.

□

En combinant les deux lemmes précédents, on remarque immédiatement qu'avoir $E = \ker(f) \oplus \mathrm{Im}(f)$ est équivalent à avoir $\ker(f) = \ker(f^2)$ et $\mathrm{Im}(f) = \mathrm{Im}(f^2)$. Il ne reste plus qu'à utiliser la question 3 pour prouver que :
$$\ker(f) = \ker(f^2) \text{ et } \mathrm{Im}(f) = \mathrm{Im}(f^2) \Longrightarrow f \text{ est pseudo inversible,}$$
et à utiliser le résultat de la question 4 pour prouver que :
$$f \text{ est pseudo inversible } \Longrightarrow \ker(f) = \ker(f^2) \text{ et } \mathrm{Im}(f) = \mathrm{Im}(f^2).$$
Ce qui achève la démonstration du théorème 31.2.

Profitons-en pour faire quelques remarques. Il semblerait que le sujet ENS Ulm-Sèvres 1988 soit sur le sujet (la question 8 m'interpelle en ce qu'elle semble donner une condition un peu moins restrictive). J'ai également remarqué que le Grifone semble traiter le sujet en appendice (A.5) dont voici un court extrait motivant :

> Le concept d'inverse généralisée a été introduit pour la première fois par Fredholm dans l'étude des opérateurs intégraux. Pour les matrices les premiers résultats importants ont été obtenus par E. H. Moore en 1920 qui a pu définir une notion d'inverse généralisée *unique* pour toute matrice à coefficients complexes. Ces résultats, qui n'avaient pas été publiés et n'avaient fait l'objet que d'une conférence, ont été retrouvés plus tard, sous des formes différentes par divers mathématiciens. En particulier, Penrose en 1955 a pu caractériser l'inverse généralisée de Moore par un système d'axiomes, ce qui en a facilité l'application à divers domaines des mathématiques.

32. Entre la Terminale et la Licence, un peu de calcul matriciel

Un vieil écrit sur lequel j'avais jeté pas mal de forces.

Malheureusement, on ne fait pas d'algèbre linéaire ici... Par exemple, on ne va pas pouvoir introduire une matrice comme le fait Godement par exemple :

32.1. **Définition.** (Godement, § 12) Soit
$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{p1} \\ \vdots & \ddots & \vdots \\ \alpha_{1q} & \cdots & \alpha_{pq} \end{pmatrix} \tag{29}$$
Un tableau de la forme (29) s'appelle une **matrice à** p **colonnes et** q **lignes à coefficients dans l'anneau** K (les α_{ij} s'appellent aussi les **termes** de la matrice en question), et on dit que (29) est la **matrice de l'homomorphisme** f **par rapport à la base** $(a_i)_{1 \leq i \leq p}$ **de** L **et à la base** $(b_i)_{1 \leq j \leq q}$ **de** M. La notion de matrice joue pour les homomorphismes un rôle analogue à celui que joue, pour les vecteurs, la notion de coordonnées.

Pour nous les matrices seront essentiellement des **tableaux de nombres remplis par des coefficients et possédant certaines opérations (addition, multiplication...)** [1]. En enlevant toutes les références à l'algèbre linéaire (à proprement dit), d'un côté on se décharge d'un poids et d'une (fausse) lourdeur mais de l'autre on perd une réelle puissance de frappe et

1. Aussi naturel qu'une telle idée puisse paraître, je me demande bien combien de personnes y avaient pensé avant qu'on leur introduise. Et, pis encore, combien de gens (en particulier chez les spécialités) ont tilté qu'un nombre était une matrice 1×1.

de conception de l'objet (à vrai dire, même si on a rapidement parlé de la notion de groupe précédemment, ce n'est pas en Terminale (ni même en spécialité) que l'on se voit introduire les $\mathrm{GL}_n(\mathbb{K})$, $\mathrm{SL}_n(\mathbb{K})$... Il faudra proprement revenir sur cela plus tard !)

À dessein (pour plus tard), je vais donner la définition et la petite (petite) introduction (rapide) aux matrices chez Colmez :

32.2. **Définition.** (Colmez, Matrices à coefficients dans un corps) Soit $\mathbf{K}$ un corps commutatif. Si n, m sont des entiers ≥ 1, on note $\mathbf{M}_{n\times m}(\mathbf{K})$ l'ensemble des *matrices* $A = (a_{i,j})_{i\leq n, j\leq m}$ à n lignes et m colonnes à coefficients dans $\mathbf{K}$ (i.e. $a_{i,j} \in \mathbf{K}$ pour tous i, j). Pour les calculs, il est souvent commode de représenter $A = (a_{i,j})_{i\leq n, j\leq m}$ (notée simplement $(a_{i,j})$, si n et m sont clairs) sous la forme d'un tableau $n \times m$

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} \tag{30}$$

L'ensemble $\mathbf{M}_{n\times m}(\mathbf{K})$ des matrices $n \times m$ est, de manière naturelle, un espace vectoriel avec l'addition et la multiplication par un scalaire définies composante par composante :

$$(a_{i,j})_{i\leq n, j\leq m} + (b_{i,j})_{i\leq n, j\leq m} = (a_{i,j} + b_{i,j})_{i\leq n, j\leq m} \text{ et } \lambda(a_{i,j})_{i\leq n, j\leq m} = (\lambda a_{i,j})_{i\leq n, j\leq m} \tag{31}$$

Et ensuite : quelques mots sur la dimension de $\boldsymbol{M}_{n\times m}(\boldsymbol{K})$ ainsi que sur un isomorphisme d'espaces vectoriels entre $\boldsymbol{M}_{n\times m}(\boldsymbol{K})$ et Hom$(\boldsymbol{K}^m, \boldsymbol{K}^n)$. Malheureusement, pas de tout ça pour nous pour l'instant !

Pour la structure globale de cette sous-partie, vu que ce n'est qu'une introduction "à la volée" (sans algèbre linéaire ; sans algèbre "du tout" même), je vais essentiellement reprendre mon cours de spécialité mathématiques. Je vais l'agrémenter de quelques parties qui sortent un peu du cadre de spécialité (mais très peu). Je suis par exemple allé zieuter dans le Gourdon d'Algèbre pour dénicher deux trois idées (rien de folichon ; le niveau est pour l'instant trop relevé pour l'utiliser).

32.1. **Généralités.** On ne va pas faire une introduction aux matrices mais plus une introduction au calcul matriciel. Sans algèbre linéaire, mis à part faire de la modélisation pour les enfants ainsi que quelques calculs, on ne va pas aller bien loin ! Malheureusement, conceptuellement, on ne tire pas grand chose d'un tel module. Il n'y a bien que la partie sur les résolutions de systèmes linéaires qui peut étonner... et encore, une fois que l'on a compris la chose, on applique bêtement mais on n'a guère plus de frisson que lorsque l'on résolvait nos premières équations linéaires à l'école primaire et au collège.

Perdu pour perdu, on va essayer de faire du pur calcul. Il n'y a pas grand chose à comprendre. Néanmoins, je vais essayer de parsemer le contenu de mon cours de terminale de quelques remarques qui pourraient donner du fond à la forme.

32.1.1. *Définition.*

32.3. **Définition.** (Matrice, spécialité Math) Une **matrice de taille** (n, p) est un tableau de nombres à n lignes et p colonnes.

Les nombres qui composent une matrice sont appelés les **coefficients de la matrice**, on les note a_{ij} avec $1 \leq i \leq n$ et $1 \leq j \leq p$.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{np} \end{pmatrix} \tag{32}$$

C'est bon, je commence à maîtriser cette histoire de ligne et colonne. Même si encore aujourd'hui j'ai tendance à confondre ligne / colonne ; horizontal / vertical ; droite / gauche... Je suis bien dans le beau drap si Godement et ma prof font les choses à l'envers l'un de l'autre ! Bref...

Avant de manipuler concrètement des matrices, une opération fondamentale s'impose à nous : celle de l'égalité entre deux matrices :

32.4. **Définition.** (Égalité entre deux matrices) Deux matrices A et B sont dites égales **si et seulement si** elles ont la même taille et tous les coefficients sont égaux.

Autrement dit : $A = B \iff \forall (i,j)_{\substack{1\leq i\leq n \\ l\leq j\leq p}}$ on a $a_{ij} = b_{ij}$, où a_{ij} désignent les coefficients de la matrice A et b_{ij} ceux de la matrice B. Les matrices A et B ayant une taille identique.

Lorsque l'on manipule des matrices, il existe des "types" bien spécifiques qui apparaissent fréquemment :

32.5. **Définition.** (Matrice carrée) On appelle **matrice carrée** une matrice qui a autant de lignes que de colonnes, c'est-à-dire que $n = p$.

Si on ne fait pas de la "géométrie déguisée" (produit scalaire, vecteur... ; et encore...), on a guère peu de chance de trouver autre chose que des matrices carrées. De plus, généralement, les opérations (un peu spécifiques) dont on va parler plus tard sont typiquement définies pour des matrices carrées. On conçoit alors parfaitement bien l'importance d'un tel type de matrice.

32.6. **Définition.** (Matrice ligne, vecteur ligne) Une **matrice ligne** (également appelée vecteur ligne) est une matrice ne contenant qu'une seule ligne, c'est-à-dire que $n = 1$.

32.7. **Exemple.**

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix} \tag{33}$$

Aux vues de la manière dont est définie la multiplication (entre matrices), en l'espèce, si l'on ne possédait que la notion de matrice ligne... on serait vite limité. C'est pourquoi, tout naturellement, on définit ce qu'est une matrice colonne.

32.8. **Définition.** (Matrice colonne, vecteur colonne) Une **matrice colonne** (également appelée vecteur colonne) est une matrice ne contenant qu'une seule colonne, c'est-à-dire que $p = 1$.

32.9. **Exemple.**

$$\begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \end{pmatrix} \tag{34}$$

Wooow, nous retrouvons l'écriture (adoptée "dès" la seconde) des vecteurs. Effectivement, un vecteur est un type bien particulier de matrice. Donc, première remarque, est-ce que ce que l'on pouvait faire avec les vecteurs (notamment en géométrie) se généralise au cas des matrices ? Oh que oui ! Même si les matrices sont introduites si tard, sans le dire (voire même de manière carrément explicite dans des sous-parties précédentes que je n'ai pas encore écrite à l'heure actuelle) on a déjà manipulé plein de matrices ! Mais pour l'instant on est que dans du calcul matriciel : un mode bien commode de calculer des choses. On est encore loin d'avoir utilisé rien qu'un bout de la puissance d'un tel objet (qui peut lui même se généraliser fort bien et fort loin) !

32.10. **Définition.** (Matrice nulle) La **matrice nulle** est une matrice dont tous les coefficients sont nuls (égaux à 0).

32.11. **Remarque.** On ne requiert aucune dimension sur la taille (la dimension) d'une matrice nulle. Elle peut être carrée ou pas ! Toutefois, à l'avenir, on notera les matrices nulles carrées de taille n : 0_n.

32.12. **Définition.** (Matrice diagonale) Une **matrice diagonale** est une matrice carrée dont tous ses coefficients sont nuls exceptés ceux situés sur la diagonale.

32.13. **Remarque.** Par commodité, que la matrice soit carrée ou pas, on pourra noter $\text{diag}(a_{11}, a_{22}, \ldots, a_{nn})$ la matrice diagonale de l'exemple suivant (où n est fixé à 4).

32.14. **Exemple.**

$$\begin{pmatrix} a_{11} & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 \\ 0 & 0 & a_{33} & 0 \\ 0 & 0 & 0 & a_{44} \end{pmatrix} \tag{35}$$

Une telle matrice paraît bien inoffensive ! On s'en servira plus qu'on ne le pense. Il existe par ailleurs un cas bien particulier de matrice diagonale : la matrice identité. De plus, il existe d'autres cas de matrices définies par le rapport entre ses éléments et la diagonale.

32.15. **Définition.** (Matrice identité) La **matrice identité d'ordre** n, notée I_n voire Id_n, est la matrice diagonale dont tous ses coefficients sont égaux à 1.

32.16. **Exemple.**

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathrm{diag}(1,1,1,1) \tag{36}$$

32.17. **Définition.** (Matrice triangulaire supérieure et inférieure) Une **matrice triangulaire supérieure** (respectivement **inférieure**) est une matrice dont tous les coefficients en dessous (resp. au-dessus) de la diagonale sont nuls. Formellement, les coefficients d'une matrice triangulaire supérieure (resp. inférieure) vérifient : pour tous (i,j) tels que $i > j, a_{ij} = 0$ (resp. pour tous (i,j) tels que $i < j, a_{ij} = 0$).

32.18. **Exemple.** (Matrice triangulaire supérieure)

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{pmatrix} \tag{37}$$

32.19. **Exemple.** (Matrice triangulaire inférieure)

$$\begin{pmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \tag{38}$$

Étant donné que l'on a pour l'instant défini aucune opération sur les matrices, on ne peut guère voir beaucoup d'autres "types" de matrice (symétrique, hermitienne, orthogonale, unitaire...) [2].

32.1.2. *Opérations.* **Addition**. Additionner est une chose courante. Une fois que l'on a saisi le "truc", on peut se débrouiller à l'infini. Toutefois, comment "généraliser" cela ? En particulier dans le cas d'une matrice ! On peut se dire que l'on va faire naturellement les choses : prenant deux matrices de tailles identiques, y'a plus qu'à ! On additionne bêtement coefficient à coefficient ! Y a-t'il ne serait-ce qu'une raison ? En effet, si l'on appliquait ce raisonnement pour la multiplication, on se trouverait bien embêté à multiplier "terme à terme " [3]... (comme on le verra, le produit de matrice est défini d'une manière quelque peu déroutante aux premiers égards).

Y a-t'il un lien entre addition et multiplication (*produit*) ? Un lien relativement similaire à celui que l'on est habitué à trouver (dans $\mathbb{N}$ par exemple) (cf. équation ci-dessous) ?

$$\underbrace{1 + 1 + \cdots + 1}_{\mathbf{n} \text{ fois}} = 1 \cdot \mathbf{n} \tag{39}$$

Malheureusement, il n'est pas vraiment le temps de se poser mille et une questions. Acceptons *pour l'instant* que les choses soient ainsi ! (Même si c'est triste à dire...)

32.20. **Définition.** (Somme de deux matrices) Soient A et B deux matrices de tailles identiques (n,p). La matrice $C = A + B$ est la matrice constituée des coefficients définis ci-après : pour tous $1 \leq i \leq n$ et $1 \leq j \leq p$, on a

$$c_{ij} = a_{ij} + b_{ij} \tag{40}$$

Qu'est-ce que cela nous apprend ? Pas franchement grand chose ! L'addition matricielle se comporte "componentwise" essentiellement comme l'addition usuelle, classique. Donc, assez logiquement (bien que l'on ait pas "créé" quelque chose de tout nouveau), on récupère les propriétés de l'addition classique : à savoir, associativité, commutativité...

2. Voir de telles matrices n'aurait d'ailleurs que fort peu de sens à ce stade.
3. Il existe néanmoins ce que l'on appelle le *produit matriciel d'Hadamard*, on ne détaille néanmoins pas (ici ; beaucoup trop tôt ; néanmoins, par rapport au produit matriciel que nous verrons tout à l'heure, en utilisant le produit matriciel d'Hadamard : on conserve la commutativité !). Il faudrait faire de l'algèbre linéaire pour y voir plus clair !

32.21. **Proposition.** (Principales propriétés) *Sachant que l'addition usuellement définies sur un "ensemble de nombre"*[4] *est* ***commutative****,* ***associative*** *et "possède" un* ***élément absorbant*** *: il en est de même pour l'addition matricielle.*

Démonstration. On ne donne qu'une idée de preuve, les détails étant surtout formels. Il convient simplement de se rendre compte que l'addition est définie terme à terme (cf. 40). De fait, la propriété se transporte naturellement aux matrices. (On pourrait raisonner dans l'autre sens en voyant une matrice (somme de deux autres matrices) de taille (n,p) comme np additions à effectuer. Et, logiquement (dans ce cas là), ce qui est vrai localement demeure vrai globalement : l'addition terme à terme étant commutative, associative [...], il en sera de même pour l'addition matricielle. Elle sera effectivement commutative, associative [...].) Oh et puis merde, faisons les détails (au moins pour la commutative).

Soit A et B deux matrices compatibles pour l'addition (dénotons leur taille par (n,p)). On va chercher à montrer que la somme matricielle est commutative (c'est-à-dire que $A+B=B+A$). Appliquons simplement la définition 32.20 à l'addition de la matrice B à la matrice A :

$$A+B=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{np} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ b_{n1} & \cdots & \cdots & b_{np} \end{pmatrix} \tag{41}$$

On constate, sans beaucoup de surprise, que c'est égal à :

$$\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1p}+b_{1p} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2p}+b_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1}+b_{n1} & \cdots & \cdots & a_{np}+b_{np} \end{pmatrix} \tag{42}$$

Ce qui est identique (par définition de l'addition usuelle) à :

$$\begin{pmatrix} b_{11}+a_{11} & b_{12}+a_{12} & \cdots & b_{1p}+a_{1p} \\ b_{21}+a_{21} & b_{22}+a_{22} & \cdots & b_{2p}+a_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ b_{n1}+a_{n1} & \cdots & \cdots & b_{np}+a_{np} \end{pmatrix} \tag{43}$$

En fin de compte, cela se ramène à :

$$\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ b_{n1} & \cdots & \cdots & b_{np} \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{np} \end{pmatrix} = B+A \tag{44}$$

Donc : $A+B=B+A$. Wow. □

32.22. **Remarque.** Soient A, B et C trois matrices (compatibles pour l'addition) et 0_{np} la matrice nulle de taille (n,p). On a donc :

$$A+B=B+A \tag{45}$$

$$A+(B+C)=(A+B)+C=A+B+C \tag{46}$$

$$0_{np}+A=A \tag{47}$$

Comme pour l'addition usuelle, la somme a son pendant : la différence.

32.23. **Définition.** (Différence de deux matrices) La différence de deux matrices A et B (compatibles pour l'addition), notée $A-B$, est définie par $A-B:=A+(-B)$.

32.24. **Remarque.** Néanmoins, à l'heure actuelle, nous ne pouvons pas décemment utiliser un tel résultat ! On trouve un signe moins devant le B, ce qui, implicitement, revient à une multiplication par moins un ! Or, nous n'avons pas défini le produit d'une matrice par un réel ! C'est ce que nous allons faire tout de suite.

4. Il conviendrait de mieux cerner cette notion, quoique vue précédemment, elle reste ouverte à débat pour nous.

Produit par un réel. On commence tout doux doucement à entrer dans des choses un poil plus intéressantes ! Mais, comme pour l'addition, pour l'instant pas franchement quoi que ce soit de nouveau sous le soleil (**et c'est sans doute à partir de là que l'on commence à comprendre que la manière "naturelle" de définir la multiplication pour les matrices (c'est-à-dire *componentwise*) risque de ne pas marcher**).

À vrai dire, on se retrouve un peu contraint dans notre démarche !

32.25. **Remarque.** N'oublions pas qu'une multiplication est une addition dissimulée ! À cet égard, $A + A = 2 \cdot A$. Si l'on regarde "le comportement des coefficients" lorsque l'on fait une telle opération, évidemment, on remarque que tout se fait "componentwise". Plus généralement, pour k un entier naturel et A une matrice, on a :

$$\underbrace{A + A + \cdots + A}_{\text{k fois}} = k \cdot A \tag{48}$$

La question qui reste à se poser étant : et si k est un réel, obtient-on un résultat similaire dans l'esprit ? (Aller voir ce que fait Tao dans son livre d'analyse (avec les suites de Cauchy). (En théorie, c'est une question qui a été traitée précédemment (mais pour les scalaires).))

Je reconnais que je suis allé un peu vite en besogne dans ma remarque ("$A + A = 2 \cdot A$"). J'ai court-circuité la chaîne de raisonnement. Néanmoins, sur l'intuition, c'est ça. Et, franchement, dans cette partie introductive, on ne cherche pas franchement tellement plus que de l'intuition (sentir l'objet, la manière de l'amener, la manière dont il est construit et ce que l'on pourrait potentiellement en faire).

32.26. **Définition.** Le produit d'une matrice A par un réel k (noté $k \cdot A$ et abrégé kA) est la matrice dont les coefficients sont obtenus en multipliant ceux de A par k. En d'autres termes, en notant a_{ij} les coefficients de A : pour tous i et j tels que $1 \leq i \leq n$ et $1 \leq j \leq p$, on a :

$$(kA)_{ij} = ka_{ij} \tag{49}$$

32.27. **Exemple.**

$$-\pi \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} -\pi a_{11} & -\pi a_{12} \\ -\pi a_{21} & -\pi a_{22} \end{pmatrix} \tag{50}$$

Immédiatement, un certain nombre de propriétés basiques en découlent :

32.28. **Proposition.** *Soient A et B deux matrices de même taille, 0 la matrice nulle et deux réels k et k'. Alors :*

— $k(A + B) = kA + kB$;

— $(k + k')A = kA + k'A$;

— $k(k'A) = kk'A = k'kA = (kk')A$;

— $1A = A$;

— $0A = 0$; *et,*

— $A + (-A) = 0$.

Produit d'une matrice ligne par une matrice colonne. Jusque là, on a été relativement épargné par les problèmes de "compatibilité" entre matrices (sous-entendu quand on tente de faire une opération entre elles). Jusqu'alors, on a jamais franchement requis plus que la condition suivante : les deux matrices doivent être de même taille. Enfin si ! À vrai dire, on a déjà demandé un petit peu plus sans vraiment le dire ! Dans la dernière section, on fait le produit d'une matrice de taille quelconque par un réel. Or, qu'est-ce qu'un réel ? Rien de moins qu'une matrice de taille $(1, 1)$. En somme, on a multiplié une matrice $(1, 1)$ avec une matrice (n, p). Or, on le verra plus tard, mais : pour parler, en "toute généralité", de compatibilité du produit entre deux matrices : on requiert que leurs dimensions soient compatibles, c'est-à-dire de taille (n, p) pour l'une et (p, q) pour l'autre. Il y a ce p en commun !

Donc, on a une petite subtilité (?) en ce qui concerne les scalaires. Désormais, raisonnons sur les vecteurs : les **matrices lignes** et les **matrices colonnes**.

32.29. **Définition.** Soient $(a_i)_{1\leq i\leq n}$ les éléments d'un vecteur ligne A de taille $(1,n)$ et $(b_i)_{1\leq i\leq n}$ les éléments d'un vecteur colonne B de taille $(n,1)$. Alors, on définit le produit entre A et B, noté AB, tel que :

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + a_2b_2 + \cdots + a_nb_n = \sum_{i=1}^{n} a_ib_i \tag{51}$$

32.30. **Remarque.** Attention, il ne faut pas faire la bêtise que je viens de faire ! Il faut veiller précautionneusement à respecter les hypothèses. J'ai cru que cette formule n'était pas cohérente avec celle précédemment vu concernant le produit d'une matrice par un réel. Je me suis dit, eh bien, supposons désormais A de taille $(1,1)$. Et j'ai bêtement appliqué la formule pour obtenir :

$$a \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \overset{?}{=} {}^{?}\, ab_1 + ab_2 + \cdots + ab_n = \sum_{i=1}^{n} ab_i = a\sum_{i=1}^{n} b_i \tag{52}$$

Or ! Faire cela revient à oublier la nature de n. En effet, n est une variable de l'énoncé. Si l'on fixe sa valeur en un endroit, il faut la fixer partout ailleurs (avec une valeur strictement identique ou équivalente). Attention à cet écueil donc ! Mais d'un autre côté, il ne faut pas en déduire que la formule de multiplication d'une matrice par un réel est fausse. Ce ne sont simplement pas les mêmes conditions qui s'appliquent. Les deux énoncés ne nous parlent pas exactement des mêmes choses, des mêmes cas.

Notons que la matrice obtenue est de taille $(1,1)$, c'est donc un scalaire. Voici un petit indice terminologique qui peut nous faire penser au **produit scalaire** ! Évidemment qu'il y a un lien ! Regardons les cas des dimensions 2 et 3 :

32.31. **Exemple.** (Dimension 2)

$$\begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1b_1 + a_2b_2 = \sum_{i=1}^{2} a_ib_i \tag{53}$$

32.32. **Exemple.** (Dimension 3)

$$\begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1b_1 + a_2b_2 + a_3b_3 = \sum_{i=1}^{3} a_ib_i \tag{54}$$

32.33. **Remarque.** Attention à l'ordre ! La nécessité de la compatibilité dimensionnelle se ressent sur l'ordre des opérations (et plus globalement sur l'ordre de multiplication). Ainsi, le produit suivant est mal défini (tout du moins, en ce qui concerne la définition 32.29) :

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \tag{55}$$

En effet, on multiplie une matrice $(n,1)$ par une matrice $(1,n)$. Formellement, cela ne respecte pas la définition 32.29. Mais est-ce véritablement un problème ? Oui et non. Non car on peut se débrouiller pour définir cela. Oui car on vient de découvrir la **non commutativité** et que ça va tout de suite compliquer le bins.

Attendons encore un peu pour pouvoir définir véritablement ledit produit. (Surprise : ça va donner une matrice (n,n).)

Produit d'une matrice par une matrice colonne. C'est notre dernier pas avant de pouvoir définir en toute généralité le produit entre deux matrices (compatibles).

32.34. **Définition.** Soient $(a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ les éléments d'une matrice A de taille (n,p) et $(b_i)_{1\leq i\leq p}$ les éléments d'un vecteur colonne B de taille $(p,1)$. Alors, on définit le produit entre A et B, noté AB, tel que :

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{np} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix} = \begin{pmatrix} a_{11}b_1 + a_{12}b_2 + \cdots + a_{1p}b_p \\ a_{21}b_1 + a_{22}b_2 + \cdots + a_{2p}b_p \\ \vdots \\ a_{n1}b_1 + a_{n2}b_2 + \cdots + a_{np}b_p \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{p} a_{1k}b_k \\ \sum_{k=1}^{p} a_{2k}b_k \\ \vdots \\ \sum_{k=1}^{p} a_{nk}b_k \end{pmatrix} \tag{56}$$

Intuitivement, la manière de procéder pour multiplier deux telles matrices est dans la "droite lignée" de ce que l'on a fait au paragraphe précédent (c'est "comme" si pour chaque ligne on avait calculé un produit scalaire (en dimension n)).

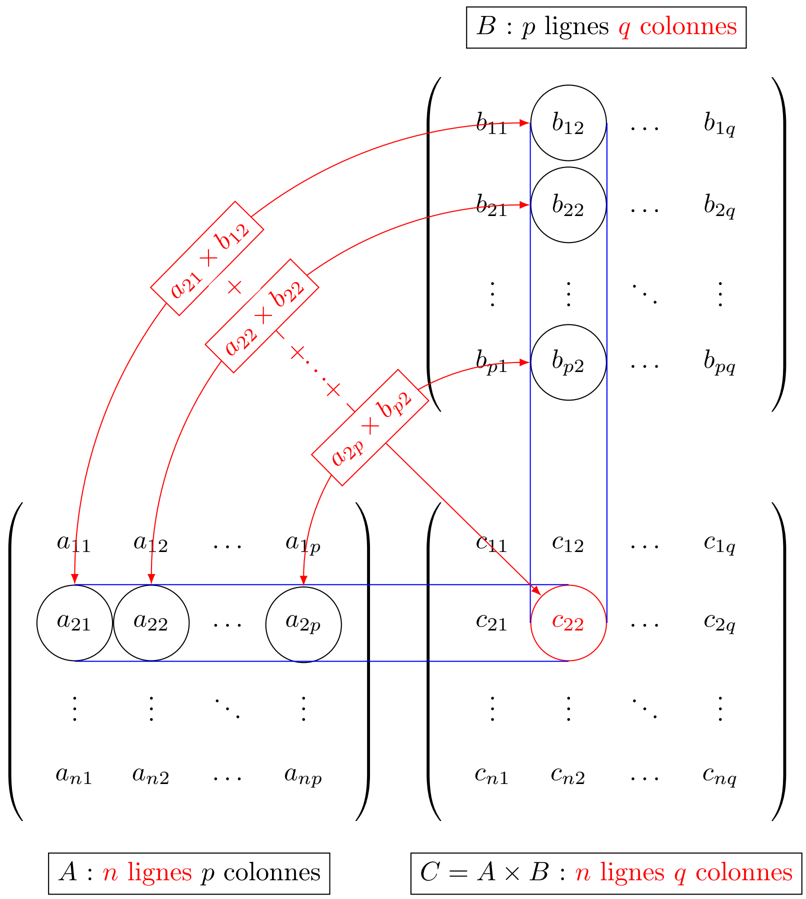
Multiplication, produit de deux matrices. C'est plus franchement la joie ! Comment concevoir globalement les choses désormais ? ! Plusieurs questions se posent à nous. Même si... franchement, en soulevant deux trois idées, les choses devraient apparaître sans trop d'encombre.

Dans la première partie sur l'addition, on était un peu embêté dans la simple mesure où (à vue de nez) la manière de généraliser aux matrices addition et produit semblait quelque peu nous dépasser. Pour mieux comprendre et approcher le concept de la manière la moins "étrange" possible, on a simplement eu à décomposer la manière de multiplier deux matrices en fonction des tailles des matrices. Un argument de compatibilité (sur la dimension) a particulièrement retenu notre attention.

32.35. **Définition.** Soient n, p et q trois entiers strictement positifs. Soient $(a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ les éléments d'une matrice A de taille (n,p) et $(b_{ij})_{\substack{1\leq i\leq p\\1\leq j\leq q}}$ les éléments d'une matrice B de taille (p,q). Alors, on définit le produit C entre A et B, noté $C = AB$, tel que C soit de dimension (n,q) et chacun de ses coefficients c_{ij} est défini tel que :

$$c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj} \tag{57}$$

Schématiquement, voici comment on obtient chacun des coefficients c_{ij} (ici, c_{22}) :



Dans l'idée, lorsque l'on regardait le produit d'une matrice par une matrice colonne, on avait émis l'hypothèse que c'était *"comme" si pour chaque ligne on avait calculé un produit scalaire*. Désormais, lorsque l'on multiples entre elles deux matrices quelconques compatibles, "on va" *juste* calculer plein de produits scalaires dans tous les sens. (Est-ce à dire qu'un coefficient de la matrice produit ainsi obtenue correspond (par exemple) à la mesure d'un certain angle entre deux vecteurs ? Oui, entre la i-ème ligne et la j-ème colonne obtenues.)

C'est la libération ! ! ! On peut enfin multiplier des choses entre elles, ouf ! Mais c'est maintenant que les choses se corsent. On a esquissé le sujet plus tôt : des subtilités viennent se frayer un passage et un petit peu déstabiliser nos sens solidement acquis et habitués aux règles traditionnelles de calculs avec des réels ou complexes.

Non commutativité. La multiplication des matrices n'est (en général) **pas commutative**. Cela signifie que AB n'a pas vraiment de raison d'être égal à BA, en général. Mais est-ce souvent le cas ? On peut être en droit de se poser la question : mais diable, quand est-ce que deux matrices commutent ? (On pourrait sans doute raisonner de manière pas trop bourine avec de l'algèbre linéaire.) On pourrait se lancer dans des calculs qui ne nous apporteraient pas franchement grand chose d'autre que : bah, faut faire le calcul en fin de compte, il faut vérifier si $AB = BA$ ou pas. Le problème, face à ce genre de question est que : si l'on fait des calculs, on aura du mal à se rattacher à des résultats pré-existants (et permettant de donner de la perspective) pour la simple et bonne raison que l'on en a pas ! Par exemple, chercher du côté d'histoire de diagonalisation, étudier avec de bons objets le centre[5] $\mathcal{Z}(G)$ de certains groupes G de matrice... Bref, ça nous emmerde.

De manière dégénéré, l'aspect non commutatif peut survenir de manière un peu étonnante : n'oublions pas les problèmes de dimension ! En toute généralité, on a vu dans la définition 32.35 que la dimension des matrices A et B devait être respectivement (n, p) et (p, q). Ainsi, l'on peut très bien faire le produit AB. Mais peut-on nécessairement faire le produit BA ? Si l'on regarde tout d'abord, la dimension : on ferait le produit d'une matrice (p, q) avec une matrice (n, p). Aïe. Problème. À un jeu de réécriture près, du simple sucre syntaxique, en appliquant la condition sur la dimension de la définition 32.35 on requerrait que : $q = n$.

5. Le centre d'un groupe n'est autre que "l'ensemble" des éléments du groupe commutant. (Attention au vocabulaire, c'est "plus" qu'un ensemble.)

Or, en général, ça n'a pas franchement de raisons de l'être. Donc, d'un côté : AB est bien défini mais BA n'a aucune raison d'être.

Peut-être qu'un dernier exemple montrant l'importance de la commutativité peut s'avérer utile. Dans $\mathbb{R}$ ou $\mathbb{C}$, on a pu s'habituer à la formule du **binôme de Newton**. Prenons en désormais garde ! Fondamentalement, la formule du binôme de Newton encode la manière d'élever à une puissance la somme de deux "nombres". (Il existe des généralisations, mais passons. Nous ne verrons que celle portant sur les matrices.) Par exemple, pour a et b deux nombres complexes :

32.36. **Exemple.**

$$(a+b)^2 = a^2 + 2ab + b^2 \tag{58}$$

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3 \tag{59}$$

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4 \tag{60}$$

Or, en regroupant certains termes, on "commet" implicitement une chose (généralement) impensable chez les matrices : on postule que le produit commute. En effet, prenons l'exemple d'une élévation à la puissance 2 (il se passe exactement les mêmes choses pour des puissances supérieures) : $(a+b)^2 = a^2 + ab + ba + b^2$. Regrouper les ab et ba en $2ab$ revient évidemment à supposer que le produit commute. Aucun problème pour les complexes ! En revanche, c'est une autre paire de manche pour les matrices.

Il faut en comprendre par là que : soient A et B deux matrices, alors : $(A+B)^2 = AA + AB + BA + BB$ (par commodité, de la même manière que pour les scalaires, on notera les puissances A^k). Ce qui revient à $(A+B)^2 = A^2 + AB + BA + B^2$. Peut-on aller plus loin ? En règle générale non. Pour continuer le calcul et trouver que $(A+B)^2 = A^2 + 2AB + B^2$, on requiert que A et B commutent. Plus généralement :

32.37. **Proposition.** (Binôme de Newton matriciel) *Soient A et B deux matrices commutatives compatibles pour le produit. Alors :*

$$(A+B)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k} \tag{61}$$

Démonstration. La démonstration n'a franchement rien à envier à celle dans le cas des "nombres" / des scalaires. Une récurrence et ça dégage. (Il serait intéressant de chercher d'autres démonstrations (typiquement une combinatoire ou d'autres plus "exotiques".) □

32.38. **Corollaire.** (sous les hypothèses de la proposition 32.37) *Étant donné que la somme de deux matrices est toujours commutative, on a que : $(A+B)^n = (B+A)^n$, ce qui revient à :*

$$\sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k} = (A+B)^n = (B+A)^n = \sum_{k=0}^{n} \binom{n}{k} B^k A^{n-k} \tag{62}$$

Donc :

$$\sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k} = \sum_{k=0}^{n} \binom{n}{k} A^{n-k} B^k \tag{63}$$

Ce corollaire 32.38, qui est en fait simplement un jeu d'écriture, peut constituer un petit tour de passe passe si on s'y prend bien.

Toutefois, on met un détail sous le tapis : est-ce que si A et B commutent, alors des puissances quelconques de A et B vont-elles également commuter ? Eh hop :

32.39. **Proposition.** *Si A et B sont deux matrices (compatibles pour le produit) commutant, alors A^m et B^n commutent également pour des entiers naturels m et n quelconques.*

Démonstration.

$$A^m B^n = A^{m-1}\mathbf{A}\ \mathbf{B}B^{n-1} = A^{m-1}\mathbf{B}\ \mathbf{A}B^{n-1} = A^{m-2}\mathbf{A}\ \mathbf{B}\ \mathbf{A}\ \mathbf{B}B^{n-2} = \cdots = B^n A^m \tag{64}$$

(Peut-être faire une double récurrence pour que ce soit plus "propre".) □

Propriétés du produit. On va évidemment retrouve l'associativité ou la distributivité ou la présence d'un élément absorbant et d'un élément neutre. Néanmoins, comme le présente le paragraphe précédent : en toute généralité, deux matrices ne commutent pas. On va également trouver quelques petites subtilités, quelques méfiances que l'on devra avoir.

32.40. **Proposition.** *Soient A, B et C des matrices (compatibles entre elles pour le produit) et k un nombre réel (voire même complexe). Alors :*

— $A(BC) = (AB)C = ABC$ *;*

— $A(B+C) = AB + AC$ *; et,*

— $A(kB) = k(AB) = kAB$*.*

Démonstration. Encore une fois, rien de fondamentalement très intéressant dans la démonstration. C'est juste un bête jeu de vérification sur les coefficients. Par exemple, pour la deuxième proposition :

Soient n, p et q trois entiers strictement positifs. Soient $(a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ les éléments d'une matrice A de taille (n,p) et $(b_{ij})_{\substack{1\leq i\leq p\\1\leq j\leq q}}$ et $(c_{ij})_{\substack{1\leq i\leq p\\1\leq j\leq q}}$ les éléments des matrice B et C de taille (p,q).

On remarque tout d'abord que, conformément à la définition de l'addition entre deux matrices 32.20, l'addition de B et C est possible (car ayant toutes deux la même dimension). Et maintenant, comme des bourrins, on calcule juste :

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{np} \end{pmatrix} \left[\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \ddots & \ddots & \vdots \\ b_{p1} & \cdots & \cdots & b_{pq} \end{pmatrix} + \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1q} \\ c_{21} & c_{22} & \cdots & c_{2q} \\ \vdots & \ddots & \ddots & \vdots \\ c_{p1} & \cdots & \cdots & c_{pq} \end{pmatrix} \right] \tag{65}$$

Étant donné que l'on ne peut pas utiliser la distributivité (car c'est ce que l'on veut démontrer), on ne peut faire qu'une seule et unique chose : calculer la somme et ensuite faire le produit de la matrice A avec la matrice conséquente de la somme entre B et C.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{np} \end{pmatrix} \begin{pmatrix} b_{11}+c_{11} & b_{12}+c_{12} & \cdots & b_{1q}+c_{1q} \\ b_{21}+c_{21} & b_{22}+c_{22} & \cdots & b_{2q}+c_{2q} \\ \vdots & \ddots & \ddots & \vdots \\ b_{p1}+c_{p1} & \cdots & \cdots & b_{pq}+c_{pq} \end{pmatrix} \tag{66}$$

Grâce à la définition 32.35, on connaît la forme de chacun des coefficients de $A(B+C)$. Ces derniers sont :

$$\mu_{ij} = \sum_{k=1}^{p} a_{ik}\Big[(b+c)_{kj}\Big] = \sum_{k=1}^{p} a_{ik}\Big[b_{kj} + c_{kj}\Big] = \sum_{k=1}^{p} a_{ik}b_{kj} + a_{ik}c_{kj} \tag{67}$$

La dernière égalité est vraie pour la simple et bonne raison que la propriété de distributivité est vérifiée pour les nombres usuels / les scalaires. (Tiens tiens tiens... c'est étonnant héhé.)

$$\mu_{ij} = \underbrace{\sum_{k=1}^{p} a_{ik}b_{kj}}_{\text{coefficient } (ab)_{ij}} + \underbrace{\sum_{k=1}^{p} a_{ik}c_{kj}}_{\text{coefficient } (ac)_{ij}} \tag{68}$$

dû à l'associativité de l'addition (définie sur les scalaires).

On remarque alors, conformément à la définition 32.35, que les coefficients μ_{ij} résultent de la somme de deux produits matriciels ; et, enfin que :

$$A(B+C) = AB + AC \tag{69}$$

Wooow. Cool. □

(*Pour la petite histoire, c'est la première fois de ma vie que j'atteins une page de démonstration "un minimum propre" en jouant avec un système de référence et de pointage inter-document de divers résultats. C'est petit mais mignon.*)

Élément absorbant et produit matriciel. Dans la lignée des propriétés sur le produit matriciel, une peut décemment retenir notre attention : il faut y faire attention ! On va raisonner en deux temps : montrer une propriété évidente puis ensuite montrer qu'une (fausse) évidence n'est pas une propriété véridique.

32.41. **Proposition.** *Soient A et B deux matrices. Si A* **ou** *B est la matrice nulle alors le produit*[6] *de A et B est également la matrice nulle.*

Démonstration. Faisons une disjonction de cas (motivée par le "ou"). Trois cas possibles : **ou** A et B sont égales à la matrice nulle, **ou** A est la matrice nulle et B une matrice quelconque, **ou** B est la matrice nulle et A une matrice quelconque. Sans perte de généralité, on remarque que les deux derniers cas sont équivalents (il n'y a qu'à faire attention à la dimension à la rigueur, mais par hypothèse on s'en fiche). Donc deux cas :

— si A **et** B sont égaux à la matrice nulle. Un rapide calcul montre que le produit de A et B équivaut également à la matrice nulle.

— si A est la matrice nulle et que B est une matrice quelconque de taille (p, q). On se débrouille pour que la multiplication entre les deux soit compatible. Choisissons alors pour la matrice nulle A une taille de (n, p). Alors AB est égal à :

$$\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \ddots & \ddots & \vdots \\ b_{p1} & \cdots & \cdots & b_{pq} \end{pmatrix} \tag{70}$$

Encore une fois, grâce à la définition 32.35, on connaît la forme de chacun des coefficients de AB. Ces derniers sont donc :

$$(ab)_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} = \sum_{k=1}^{p} 0 b_{kj} = 0 \tag{71}$$

Donc chacun des coefficients de la matrice obtenue est 0. Donc $AB = 0_{nq}$.

□

La réciproque d'un tel résultat est-elle vraie ? Si $AB = 0$, a-t'on nécessairement A ou B égal à la matrice nulle ? C'est faux, à cet égard, on peut exhiber des exemples pour le prouver :

32.42. **Exemple.** Si $A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$ et $B = \begin{pmatrix} 5 & 9 \\ 5 & 9 \end{pmatrix}$, alors $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Mais manifestement, A et B ne sont ni l'une ni l'autre la matrice nulle.

Plus généralement que pour la matrice nulle, si on obtient l'égalité suivante $AB = AC$ entre les trois matrices A, B et C, alors on ne peut pas simplifier par A et déduire que $B = C$. On va le voir plus tard mais cette opération de "simplification" a un nom, elle s'appelle "prendre l'inverse" mais n'est pas toujours possible ! (Donc ce n'est pas vraiment une simplification comme pourrait l'être la division dans $\mathbb{R}$ vu que dans $\mathbb{R}$ on peut toujours diviser (sauf par zéro, bien entendu).)

Élément neutre et produit matriciel. À côté de l'élément neutre (le "zéro"), on trouve l'élément neutre (le "un"). De la même manière que chez les réels et complexes, on attend de lui qu'il préserve, ne change pas ce sur quoi on le fait agir. Assez logiquement, on en déduit que :

32.43. **Proposition.** *Soit A une matrice carrée compatible pour la multiplication avec la matrice identité I_n. Alors $AI_n = I_nA = A$.*

Démonstration. La matrice A est de taille (n, n) (pour être compatible pour la multiplication avec I_n et car elle est carrée). Pour prouver le résultat escompté, il n'y a qu'à dérouler les calculs en raisonnant sur les coefficients (on introduit à la volée la notation $\text{coeff}(X)_{ij}$ donnant les coefficients de la matrice X) :

$$\text{coeff}(AI_n)_{ij} = \sum_{k=1}^{n} a_{ik} \text{coeff}(I_n)_{kj} = \sum_{k=1}^{n} a_{ik} \delta_{kj} \tag{72}$$

$$\text{coeff}(I_nA)_{ij} = \sum_{k=1}^{n} \text{coeff}(I_n)_{ik} a_{kj} = \sum_{k=1}^{n} \delta_{ik} a_{kj} \tag{73}$$

6. En se débrouillant pour qu'il soit compatible, c'est-à-dire en choisissant des dimensions adéquates pour les matrices nulles.

où δ_{ij} est le symbole de Kronecker défini tel que $\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$. Dans une autre version de la démonstration, j'ai fait une connerie et je suis allé trop vite en besogne. Regardons patiemment, ce que l'on obtient pour quelques valeurs de i et j :

$$\text{pour } j = 1,\ \sum_{k=1}^{n} a_{ik}\delta_{k1} = a_{i1} \cdot 1 + a_{i2} \cdot 0 + \cdots + a_{in} \cdot 0 = a_{i1} \tag{74}$$

$$\text{pour } j = 2,\ \sum_{k=1}^{n} a_{ik}\delta_{k2} = a_{i1} \cdot 0 + a_{i2} \cdot 1 + \cdots + a_{in} \cdot 0 = a_{i2} \tag{75}$$

$$\text{pour } i = 1,\ \sum_{k=1}^{n} \delta_{1k}a_{kj} = 1 \cdot a_{1j} + 0 \cdot a_{2j} + \cdots + 0 \cdot a_{nj} = a_{1j} \tag{76}$$

$$\text{pour } i = 2,\ \sum_{k=1}^{n} \delta_{2k}a_{kj} = 0 \cdot a_{1j} + 1 \cdot a_{2j} + \cdots + 0 \cdot a_{nj} = a_{2j} \tag{77}$$

On remarque (et démontre) plus généralement que pour i et j "quelconques" (dans la limite du possible, c'est à dire compris entre 1 et n) :

$$\text{coeff}(AI_n)_{ij} = \sum_{k=1}^{n} a_{ik}\delta_{kj} = a_{ij} \tag{78}$$

$$\text{coeff}(I_nA)_{ij} = \sum_{k=1}^{n} \delta_{ik}a_{kj} = a_{ij} \tag{79}$$

Ce qui permet de conclure. Il reste néanmoins à montrer que AI_n ou I_nA est égal à A. Ce qui est automatique. □

32.44. **Remarque.** Le seul point un petit peu délicat dans la démonstration précédente intervient lors de l'introduction du symbole de Kronecker δ_{ij}. Donnons juste un élément pour justifier son utilisation : on peut définir les coefficients de la matrice identité I_n par une condition : si l'on est sur la diagonale principale, alors on met un un ; sinon, un zéro. Ce qui revient à utiliser le symbole de Kronecker (en fonction de si $i = j$ ou pas).

Petite chose délicate quand on est étourdi également : ne pas chercher à trouver les coefficients de A avec la formule 57. Dans une telle situation, ce n'est pas seulement con (on n'a aucune information sur A et ses coefficients, donc on cherche quelque chose d'inatteignable). Pire, c'est même faux : on ne vérifie pas les conditions de compatibilité sur la dimension. Certes, on a une matrice carrée (n, n). Dans ce cas, selon la définition 32.35, il faudrait considérer une matrice de dimension (n, qqch). Or, bah non : $a_{ij} \neq \sum_{k=1}^{n} a_{ik} \cdot 1$. Gné, non. En plus on mélange tout... Non. À ne pas faire comme connerie !

Puissance d'une matrice. On ne considère que les matrices carrées (autrement, on aurait de petit problème de compatibilité sur la multiplication : en effet, on multiplierai la même matrice par elle même, donc une matrice (n, p) par une matrice (n, p), ce qui n'est pas compatible).

32.45. **Définition.** (Convention) On note A^2 le produit de la matrice carrée A par elle-même. Ainsi, $A^2 := AA$. On définit, par récurrence, de manière identique l'élévation de la matrice A à une puissance k (e.g. $A^3 = A^2A = AAA$).

32.46. **Remarque.** Par convention $A^0 = I_n$.

32.47. **Remarque.** $A^{n+1} = A^nA = AA^n$

Doit-on se limiter à un exposant positif ou nul ? À vrai dire "non". On le verra plus tard avec la notion d'inverse, le fameux A^{-1}.

La dynamique avec les matrices semble nettement plus riche que celle pour les scalaires. Sans chercher à entrer trop dans les détails (la question pourrait être intéressante en développement et en étude), l'on pourrait se poser des problèmes analogues à ceux de la recherche des racines de l'unité dans $\mathbb{C}$. Pour les matrices, c'est déjà un peu plus trépidant, ça va un peu plus loin que les (non pas moins intéressants) ζ_n : par exemple concernant les racines deuxièmes de l'unité : on trouve évidemment que $I_2^2 = I_2$ ou $(-I_2)^2 = I_2$. Mais ce n'est pas tout [7] :

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^2 = I_2, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2 = I_2, \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}^2 = I_2 \tag{80}$$

ou des trucs un peu plus exotiques comme

$$\begin{pmatrix} i & 73+\sqrt{5327} \\ 73-\sqrt{5327} & -i \end{pmatrix}^2 = I_2 \tag{81}$$

. Et en plus, y'a moyen que ça fasse de beaux *petits* problèmes d'arithmétiques (avec une orientation algorithmique) car le carré de $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est $\begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & cb+d^2 \end{pmatrix}$. Donc le problème se ramènerait au système suivant :

$$\begin{cases} a^2+bc & = 1 \\ ab+bd & = 0 \\ ac+cd & = 0 \\ cb+d^2 & = 1 \end{cases} \tag{82}$$

Mais ce serait bien qu'enfin on fasse de l'algèbre linéaire, bon dieu ! Malheureusement, on verra dans quelques longs mois. En attendant, développons quelques éléments de calcul matriciel. On les aura juste en tête, histoire d'avoir des repères plutôt que de réelles connaissances approfondies.

32.2. **Quelques éléments de calcul matriciel.** Pour la suite des hostilités, je me suis basé sur différents documents que voici : les polycopiés de Ycart, Bodin, Rouget et Schwartz.

On va voir quelques opérations qui, pour l'instant, ne relèveront de quasiment aucune utilité. À part calculer, on ne risque pas d'aller très loin. On va simplement essayer de se familiariser un peu avec ces derniers (en démontrant notamment des propriétés très basiques et rudimentaires).

32.2.1. *Trace.* On se place dans le cadre des matrices carrées.

32.48. **Définition.** Soit A une matrice carrée de dimension n. La **trace** de la matrice A, notée $\mathrm{Tr}(A)$, est la somme des coefficients diagonaux de A :

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii} \tag{83}$$

où les a_{ij} sont les coefficients de la matrice A.

32.49. **Remarque.** La trace d'une matrice est donc la somme de ses éléments diagonaux (ceux présents sur la diagonale principale).

32.50. **Exemple.**

$$\mathrm{Tr}\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{pmatrix} = a_{11} + a_{22} + \cdots + a_{nn} \tag{84}$$

32.51. **Proposition.** *Soient A et B deux matrices carrées de dimension n. Alors :*

7. Apparemment, il y aurait une histoire avec les matrices de Pauli (?).

1. $\mathrm{Tr}(A+B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$ *;*
2. $\mathrm{Tr}(\alpha A) = \alpha \mathrm{Tr}(A)$ *pour tout* α *appartenant à* $\mathbb{C}$ *;*
3. $\mathrm{Tr}(A^\intercal) = \mathrm{Tr}(A)$ *; et,*
4. $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$*.*

Démonstration. 1. appliquons simplement la définition 32.48 à $A+B$ et déroulons les calculs :

(85)
$$Tr(A+B) = Tr\left(\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ b_{n1} & \cdots & \cdots & b_{nn} \end{pmatrix}\right)$$

(86)
$$\mathrm{Tr}(A+B) = \mathrm{Tr}\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1}+b_{n1} & \cdots & \cdots & a_{nn}+b_{nn} \end{pmatrix} = (a_{11}+b_{11}) + (a_{22}+b_{22}) + \cdots + (a_{nn}+b_{nn})$$

(87)
$$\mathrm{Tr}(A+B) = \sum_{i=1}^{n}(a_{ii}+b_{ii}) = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \mathrm{Tr}(A) + \mathrm{Tr}(B)$$

2. même logique que pour la démonstration précédente : on "construit" / forme la matrice αA et lui applique la définition 32.48 afin de démontrer ladite propriété :

(88)
$$\mathrm{Tr}(\alpha A) = \mathrm{Tr}\begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ \alpha a_{n1} & \cdots & \cdots & \alpha a_{nn} \end{pmatrix}$$

(89)
$$\mathrm{Tr}(\alpha A) = \sum_{i=1}^{n} \alpha a_{ii} = \alpha \sum_{i=1}^{n} a_{ii} = \alpha \mathrm{Tr}(A)$$

3. on n'a pas encore défini la notion de transposée, anticipons un tout petit peu. *Intuitivement*, la transposée va "faire tourner la matrice sur elle-même" en échangent les coefficients entre eux : les a_{ij} vont devenir les a_{ji} et inversement (les a_{ji} deviennent les a_{ij}). Ainsi :

(90)
$$\mathrm{Tr}(A^\intercal) = \mathrm{Tr}\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \ddots & \ddots & \vdots \\ a_{1n} & \cdots & \cdots & a_{nn} \end{pmatrix}$$

On constate que les coefficients sur la diagonale principale n'ont aucunement changé (seuls ceux en dessous et ceux au-dessus ont changé). De ce fait :

(91)
$$\mathrm{Tr}(A^\intercal) = \mathrm{Tr}(A)$$

Une telle formule est tout à fait logique : la diagonale principale "agit comme" / "se comporte comme" un axe de symétrie, il est donc naturel qu'elle soit invariante (sous sa propre action).

4. tout d'abord, on sait que l'on aura aucun problème de compatibilité du produit dans la mesure où les matrices A et B sont toutes deux carrées d'ordre n. Les produits AB et BA sont donc bien définis. Comment démontrer la propriété 4 ? Un petit retour au source à la formule 57 donnant la "tête" des coefficients de AB et BA puis une simple application de la définition 32.48 et le tour est joué !

En utilisant la formule 57, on obtient les coefficients de AB et de BA :

$$(ab)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \tag{92}$$

$$(ba)_{ij} = \sum_{k=1}^{n} b_{ik} a_{kj} \tag{93}$$

On a précédemment remarqué (ou sinon on le fait maintenant) que la trace *ne s'intéresse qu'aux éléments sur la diagonale principale*. De ce fait, on peut se concentrer uniquement sur les coefficients pour lesquels $i = j$ (c'est-à-dire $(ab)_{ii}$ et $(ba)_{ii}$). Et c'est maintenant que je peux recaser le travail que j'avais fait pour la démonstration de la proposition 32.71. C'est pile ce qu'il faut ici. On va alors simplement chercher à démontrer que

$$\mathrm{Tr}(AB) = \sum_{k=1}^{n} (ab)_{ii} = \sum_{k=1}^{n} (ba)_{ii} = \mathrm{Tr}(BA) \tag{94}$$

Pour peu que l'on ordonne et présente suffisamment bien les choses, ça devient une évidence. Voyons donc cela :

$$\begin{aligned}
(ab)_{11} &= a_{11}b_{11} + a_{12}b_{21} + \dots + a_{1n}b_{n1} && (95)\\
(ab)_{22} &= a_{21}b_{12} + a_{22}b_{22} + \dots + a_{2n}b_{n2} && (96)\\
\dots &= \dots + \dots + \dots + \dots && (97)\\
(ab)_{nn} &= a_{n1}b_{1n} + a_{n2}b_{2n} + \dots + a_{nn}b_{nn} && (98)
\end{aligned}$$

$$\begin{aligned}
(ba)_{11} &= b_{11}a_{11} + b_{12}a_{21} + \dots + b_{1n}a_{n1} && (99)\\
(ba)_{22} &= b_{21}a_{12} + b_{22}a_{22} + \dots + b_{2n}a_{n2} && (100)\\
\dots &= \dots + \dots + \dots + \dots && (101)\\
(ba)_{nn} &= b_{n1}a_{1n} + b_{n2}a_{2n} + \dots + b_{nn}a_{nn} && (102)
\end{aligned}$$

On se rend visuellement compte qu'il existe une égalité algébrique (exactement celle désirée). Afin que ce soit plus clair et plus visible, trois exemples ont été mis en couleur (ceux des premiers, seconds et derniers coefficients de la diagonale principale). Avec un petit jeu de ré-arrangement des coefficients lorsqu'on les somme, on tombe évidemment sur l'égalité 94 désirée.

Une telle égalité peut surprendre dans la mesure où l'on s'est habitué à la non commutativité de A et B. Toutefois, il ne convient de pas s'arrêter à cette intuition dans la mesure où l'on raisonne uniquement sur la diagonale principale avec la trace (diagonale principale "utilisée comme un axe de symétrie" par la trace, comme vu plus tôt).

□

32.52. **Corollaire.** (sous les hypothèses de la proposition 32.51) *Si A et B sont des* ***matrices semblables****, c'est-à-dire qu'elles vérifient l'égalité $B = P^{-1}AP$ (où P est une matrice carrée inversible d'ordre n), alors* $\mathrm{Tr}(A) = \mathrm{Tr}(B)$ *;*

Démonstration. Conséquemment à la proposition 32.51, on sait que $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. Injectons simplement la valeur $B = P^{-1}AP$, et nous obtenons :

$$\mathrm{Tr}(B) = \mathrm{Tr}(P^{-1}(AP)) = \mathrm{Tr}((AP)P^{-1}) = \mathrm{Tr}(A(PP^{-1})) = \mathrm{Tr}(AI_n) = \mathrm{Tr}(A) \tag{103}$$

La démonstration pourra être parfaitement comprise dès que l'on aura fait la partie sur l'inverse d'une matrice, dans très peu de temps (cf. la définition 32.68). □

Plus précisément que les propriétés une et deux de la proposition 32.51, la trace est une **forme linéaire** sur "l'espace des matrices carrées". On reste volontairement flou pour l'instant, il est encore trop tôt. Néanmoins, en pratique, ça signifie simplement que pour deux matrices carrées A et B et deux scalaires λ et μ (d'une certaine structure algébrique), on a l'égalité suivante :

$$\mathrm{Tr}(\lambda A + \mu B) = \lambda \mathrm{Tr}(A) + \mu \mathrm{Tr}(B) \tag{104}$$

32.53. **Remarque.** La trace peut être intéressante à divers égards. Par exemple, elle peut être utilisée pour démontrer qu'il n'existe pas de couple de matrices carrées (A, B) tel que $AB - BA = I_n$. En effet : en raisonnant sur la trace, on constate que $\mathrm{Tr}(AB - BA) = \mathrm{Tr}(AB) - \mathrm{Tr}(BA) = \mathrm{Tr}(AB) - \mathrm{Tr}(AB) = 0$. Or, la trace de la matrice identité d'ordre n vaut précisément n. Dans la mesure où 0 est différent de n (n étant un entier strictement positif), il n'existe alors pas de couple (A, B) tel que $AB - BA = I_n$.

En somme, la trace nous a ici permis de vérifier si une égalité est probable ou pas du tout. Si les deux traces ne correspondent pas, il est impossible que les deux matrices soient égales (en revanche, veillons à faire attention : la contraposée n'est pas vraie : deux matrices différentes peuvent avoir des traces identiques).

32.2.2. *Transposée.* Contrairement à précédemment (pour la trace ; ou plus tard pour le déterminant), on n'est plus obligé de se restreindre à des matrices carrées. En effet, on peut très bien calculer la **transposée** d'une matrice (n, p).

32.54. **Définition.** Soit A une matrice de taille (n, p). La **transposée** de A, notée $A^\intercal$, est la matrice obtenue en interchangeant les colonnes avec les lignes (et inversement) de A. Formellement, en dénotant a_{ij} les np coefficients de la matrice A, on définit les coefficients de la matrice transposés tel que :

$$\mathrm{coeff}(A^\intercal)_{ji} = a_{ij} \tag{105}$$

32.55. **Exemple.**

$$\begin{pmatrix} a_{11} & a_{\mathbf{12}} & \cdots & a_{\mathbf{1\ p}} \\ a_{\mathbf{21}} & a_{22} & \cdots & a_{\mathbf{2\ p}} \\ \vdots & \ddots & \ddots & \vdots \\ a_{\mathbf{n\ 1}} & \cdots & \cdots & a_{np} \end{pmatrix}^\intercal = \begin{pmatrix} a_{11} & a_{\mathbf{21}} & \cdots & a_{\mathbf{n\ 1}} \\ a_{\mathbf{12}} & a_{22} & \cdots & a_{\mathbf{n\ 2}} \\ \vdots & \ddots & \ddots & \vdots \\ a_{\mathbf{1\ p}} & \cdots & \cdots & a_{np} \end{pmatrix} \tag{106}$$

32.56. **Corollaire.** *En conséquence de l'équation 105 définissant les coefficients de la matrice $A^\intercal$, la matrice transposée de A est de taille (p, n).*

On avait déjà pu rapidement voir la notion de transposée à la proposition 32.51 (troisième propriété). À vrai dire, cette notion était déjà latente dans d'autres concepts. En effet, lorsque l'on a remarqué que le produit scalaire pouvait avoir une expression matricielle aux exemples 32.31 et 32.32, sans trop le dire, on a usé de la notion de transposition. Dans cette optique, lorsque l'on considère deux vecteurs, on va les prendre de même dimension (disons $(n, 1)$, c'est-à-dire une matrice colonne (cf. définition 32.8). Problème, problème ! Pour calculer un produit scalaire, on conviendra que l'on calcul un *produit*, gné ! Or, au sens de la condition de compatibilité sur la dimension de la définition 32.35, on est ici emmerdé ! En effet, on veut faire le produit d'une matrice $(n, 1)$ avec une matrice $(n, 1)$, gné. Donc comment faire ? ! On va, subrepticement, utiliser, vite fait en soum-soum, la notion de transposée pour obtenir une matrice $(1, n)$. Et désormais, le produit est compatible ! (Et faut faire néanmoins gaffe à l'ordre des multiplications ! En effet, on ne veut pas obtenir une matrice (n, n) mais bien un scalaire (donc une matrice $(1, 1)$).)

32.57. **Exemple.** Imaginons que l'on souhaite faire le produit scalaire entre les deux vecteurs colonnes suivants :

$$\overrightarrow{v} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \overrightarrow{w} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{107}$$

$\overrightarrow{v}$ scalaire $\overrightarrow{w}$, noté $\langle \overrightarrow{v}, \overrightarrow{w} \rangle$ ou encore $\overrightarrow{v} \cdot \overrightarrow{w}$, va s'écrire :

$$\overrightarrow{v} \cdot \overrightarrow{w} = (\overrightarrow{v}^\intercal)\overrightarrow{w} = \begin{pmatrix} a_1 & a_2 & \ldots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \sum_{i=1}^{n} a_i b_i \tag{108}$$

On comprend et voit alors aisément que pour n quelconque, on peut définir le produit scalaire entre deux vecteurs A et B tel que $A \cdot B = A^\intercal B$.

32.58. **Proposition.** *Soient A une matrice de dimension (n, p). Alors :*

1. $(A^\intercal)^\intercal = A$ *;*

2. $(\alpha A)^\intercal = \alpha(A^\intercal)$ *;*

3. $(A+B)^\intercal = A^\intercal + B^\intercal$*, où* B *est une matrice compatible pour l'addition (cf. 32.20) ; et,*

4. $(AB)^\intercal = B^\intercal A^\intercal$*, où* A *et* B *sont compatibles pour la multiplication (cf. 32.35) ;*

Démonstration. 1. Cette propriété est équivalente au fait de dire que la transposée est **involutive**. Prouvons la. Par définition, on applique simplement deux fois la relation 105 :

$$A^\intercal = \begin{pmatrix} a_{11} & a_{\mathbf{12}} & \cdots & a_{\mathbf{1\ p}} \\ a_{\mathbf{21}} & a_{22} & \cdots & a_{\mathbf{2\ p}} \\ \vdots & \ddots & \ddots & \vdots \\ a_{\mathbf{n\ 1}} & \cdots & \cdots & a_{np} \end{pmatrix}^\intercal = \begin{pmatrix} a_{11} & a_{\mathbf{21}} & \cdots & a_{\mathbf{n\ 1}} \\ a_{\mathbf{12}} & a_{22} & \cdots & a_{\mathbf{n\ 2}} \\ \vdots & \ddots & \ddots & \vdots \\ a_{\mathbf{1\ p}} & \cdots & \cdots & a_{np} \end{pmatrix} \tag{109}$$

$$(A^\intercal)^\intercal = \begin{pmatrix} a_{11} & a_{\mathbf{21}} & \cdots & a_{\mathbf{n\ 1}} \\ a_{\mathbf{12}} & a_{22} & \cdots & a_{\mathbf{n\ 2}} \\ \vdots & \ddots & \ddots & \vdots \\ a_{\mathbf{1\ p}} & \cdots & \cdots & a_{np} \end{pmatrix}^\intercal = \begin{pmatrix} a_{11} & a_{\mathbf{12}} & \cdots & a_{\mathbf{1\ p}} \\ a_{\mathbf{21}} & a_{22} & \cdots & a_{\mathbf{2\ p}} \\ \vdots & \ddots & \ddots & \vdots \\ a_{\mathbf{n\ 1}} & \cdots & \cdots & a_{np} \end{pmatrix} = A \tag{110}$$

Une telle relation n'a rien de franchement surprenant dans la mesure où "pivoter deux fois de 180 degrés revient à effectuer un tour complet" (c'est-à-dire à "ne pas bouger", dans la mesure où 360 degrés est congru à 0 degré, modulo 360).

2. Il n'y a qu'à appliquer les relations 105 et 49 :

$$(\alpha A)^\intercal = \begin{pmatrix} \alpha a_{11} & \alpha a_{\mathbf{12}} & \cdots & \alpha a_{\mathbf{1\ p}} \\ \alpha a_{\mathbf{21}} & \alpha a_{22} & \cdots & \alpha a_{\mathbf{2\ p}} \\ \vdots & \ddots & \ddots & \vdots \\ \alpha a_{\mathbf{n\ 1}} & \cdots & \cdots & \alpha a_{np} \end{pmatrix}^\intercal = \alpha \begin{pmatrix} a_{11} & a_{\mathbf{21}} & \cdots & a_{\mathbf{n\ 1}} \\ a_{\mathbf{12}} & a_{22} & \cdots & a_{\mathbf{n\ 2}} \\ \vdots & \ddots & \ddots & \vdots \\ a_{\mathbf{1\ p}} & \cdots & \cdots & a_{np} \end{pmatrix} = \alpha A^\intercal \tag{111}$$

3. Encore une fois, c'est de l'application pure et dure de relations fondamentales (105 et 32.20). On a de plus supposé les deux matrices compatibles pour l'addition, donc aucun problème !

$$(A+B)^\intercal = \left[\begin{pmatrix} a_{11} & a_{\mathbf{12}} & \cdots & a_{\mathbf{1\ p}} \\ a_{\mathbf{21}} & a_{22} & \cdots & a_{\mathbf{2\ p}} \\ \vdots & \ddots & \ddots & \vdots \\ a_{\mathbf{n\ 1}} & \cdots & \cdots & a_{np} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{\mathbf{12}} & \cdots & b_{\mathbf{1\ p}} \\ b_{\mathbf{21}} & b_{22} & \cdots & b_{\mathbf{2\ p}} \\ \vdots & \ddots & \ddots & \vdots \\ b_{\mathbf{n\ 1}} & \cdots & \cdots & b_{np} \end{pmatrix} \right]^\intercal \tag{112}$$

$$(A+B)^\intercal = \left[\begin{pmatrix} a_{11}+b_{11} & a_{\mathbf{12}}+b_{\mathbf{12}} & \cdots & a_{\mathbf{1\ p}}+b_{\mathbf{1\ p}} \\ a_{\mathbf{21}}+b_{\mathbf{21}} & a_{22}+b_{22} & \cdots & a_{\mathbf{2\ p}}+b_{\mathbf{2\ p}} \\ \vdots & \ddots & \ddots & \vdots \\ a_{\mathbf{n\ 1}}+b_{\mathbf{n\ 1}} & \cdots & \cdots & a_{np}+b_{np} \end{pmatrix} \right]^\intercal \tag{113}$$

$$(A+B)^\intercal = \begin{pmatrix} a_{11}+b_{11} & a_{\mathbf{21}}+b_{\mathbf{21}} & \cdots & a_{\mathbf{n\ 1}}+b_{\mathbf{n\ 1}} \\ a_{\mathbf{12}}+b_{\mathbf{12}} & a_{22}+b_{22} & \cdots & a_{\mathbf{n\ 2}}+b_{\mathbf{n\ 2}} \\ \vdots & \ddots & \ddots & \vdots \\ a_{\mathbf{1\ p}}+b_{\mathbf{1\ p}} & \cdots & \cdots & a_{np}+b_{np} \end{pmatrix} \tag{114}$$

Ce qui nous permet alors de conclure que :

$$(A+B)^\intercal = \begin{pmatrix} a_{11} & a_{\mathbf{21}} & \cdots & a_{\mathbf{n\ 1}} \\ a_{\mathbf{12}} & a_{22} & \cdots & a_{\mathbf{n\ 2}} \\ \vdots & \ddots & \ddots & \vdots \\ a_{\mathbf{1\ p}} & \cdots & \cdots & a_{np} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{\mathbf{21}} & \cdots & b_{\mathbf{n\ 1}} \\ b_{\mathbf{12}} & b_{22} & \cdots & b_{\mathbf{n\ 2}} \\ \vdots & \ddots & \ddots & \vdots \\ b_{\mathbf{1\ p}} & \cdots & \cdots & b_{np} \end{pmatrix} = A^\intercal + B^\intercal \tag{115}$$

4. Pas de nouveauté, on applique des relations (105 et 57) et tout roule. On va essayer d'être un peu plus précautionneux que pour les trois dernières propriétés. Tout d'abord, nous n'avons aucun problème de compatibilité pour l'addition. En effet, en tant que pré-requis, on requiert de B qu'elle soit compatible pour la multiplication avec A (c'est-à-dire que B soit de taille (p,q)). Nous allons chercher à exhiber les coefficients de la transposée de AB afin de se rendre compte qu'ils coïncident parfaitement avec ceux de la transposée de B multiplié par la transposée de A. En fait non, on va faire les choses dans l'autre sens. Donc :

$$(b^\intercal a^\intercal)_{\mathbf{i}\ \mathbf{j}} = \sum_{k=1}^{p} (b^\intercal)_{\mathbf{i}\ \mathbf{k}} (a^\intercal)_{\mathbf{k}\ \mathbf{j}} \tag{116}$$

C'est désormais que nous avons le seul point un tant soit peu *délicat*. En appliquant la définition 32.54 de la transposée, grâce à la formule 105, on constate que la transposée "échange les coefficients", elle "change leur ordre" (un (i,j) devient un (j,i)). Donc, tout logiquement, on trouve que :

$$(b^\intercal a^\intercal)_{\mathbf{i}\ \mathbf{j}} = \sum_{k=1}^{p} b_{\mathbf{k}\ \mathbf{i}} a_{\mathbf{j}\ \mathbf{k}} = \sum_{k=1}^{p} a_{\mathbf{j}\ \mathbf{k}} b_{\mathbf{k}\ \mathbf{i}} \tag{117}$$

Cette dernière équation nous permet de conclure :

$$(b^\intercal a^\intercal)_{\mathbf{i}\ \mathbf{j}} = \sum_{k=1}^{p} a_{\mathbf{j}\ \mathbf{k}} b_{\mathbf{k}\ \mathbf{i}} = (ab)_{\mathbf{j}\ \mathbf{i}} = ((ab)^\intercal)_{\mathbf{i}\ \mathbf{j}} \tag{118}$$

Wooow. Super cool.

□

Pour les grandes lignes de la transposée, on devrait être plus ou moins bon (il reste un petit résultat avec le déterminant, on le traitera dans le paragraphe sur le déterminant). On va désormais voir deux mini-résultats et ensuite enchaîner sur deux types de matrices assez importants : les matrices symétriques et les matrices antisymétriques.

On a déjà pu avoir un tout petit peu à faire avec des matrices semblables. Définissons un peu plus proprement que précédemment (cf. la proposition 32.52) une telle notion :

32.59. **Définition.** Deux matrices carrées A et B sont dites **semblables** s'il existe une matrice P inversible telle que :

$$A = PBP^{-1} \tag{119}$$

De telles matrices sont toutes mignonnes et peuvent être rudement intéressantes notamment si l'on souhaite calculer des exponentiation par exemple. Le but du jeu va être de trouver une matrice P (et son inverse) ainsi qu'une matrice B (plutôt ravissante, c'est-à-dire dont on calcule aisément ses puissances) liées à A par la relation 119. Dans la pratique, quand on va chercher à calculer, par exemple, A^k on va se simplifier le travail grâce à la formule suivante (qui peut se démontrer par récurrence) :

$$A^k = (PBP^{-1})^k = \underbrace{\underbrace{PBP^{-1}}\underbrace{PBP^{-1}}\dots\underbrace{PBP^{-1}}}_{\text{k fois}} = PB^kP^{-1} \tag{120}$$

Sans doute pas forcément très utile en pratique, on trouve, en théorie, un résultat liant une matrice A à sa transposée :

32.60. **Proposition.** *Une matrice A est semblable à sa transposée, c'est-à-dire qu'il existe une matrice P inversible telle que :*

$$A = PA^\intercal P^{-1} \tag{121}$$

Démonstration. Apparemment, en utilisant une chose qu'on appellerait la *réduction de Jordan*, ça deviendrait une trivialité. Autrement, ça semble être un exercice un petit peu "compliqué". (Par exemple, c'est un sujet de Centrale, 2003, Mathématiques 2 (TSI).) On verra ce que l'on peut en faire plus tard. □

Enfin, avant de passer aux matrices symétriques et antisymétriques, voyons un petit résultat qui va de lui-même :

32.61. **Proposition.** *Soit A une matrice triangulaire supérieure (resp. inférieure), la transposée de A est une matrice triangulaire inférieure (resp. supérieure).*

Démonstration. On rappelle la définition de matrice triangulaire supérieure 32.17, resp. inférieure : les coefficients a_{ij} de A vérifient la relation suivante : pour tous i et j tels que $i > j$, $a_{ij} = 0$, resp pour tous i et j tels que $i < j$, $a_{ij} = 0$.

Nous n'allons traiter que le cas d'une matrice triangulaire supérieure. Celui d'une matrice triangulaire inférieure se traite parallèlement.

Par hypothèse, nous savons que A est de la forme suivante :

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ 0 & a_{22} & \dots & a_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{np} \end{pmatrix} \tag{122}$$

Calculons simplement la transposée en appliquant la formule 105 :

$$A^{\intercal} = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{12} & a_{22} & \dots & \vdots \\ \vdots & \dots & \ddots & 0 \\ a_{1p} & a_{2p} & \dots & a_{np} \end{pmatrix} \tag{123}$$

On reconnaît immédiatement une matrice triangulaire inférieure, comme désiré. □

Matrices symétriques.

32.62. **Définition.** Soit une matrice carrée A d'ordre n. La matrice A est dite **symétrique** si elle est égale à sa transposée, c'est-à-dire si :

$$A = A^{\intercal} \tag{124}$$

32.63. **Remarque.** Les coefficients sont dits symétriques par rapport à la diagonale principale.

On pourrait, apparemment, s'amuser à compter le nombre de matrices symétriques pour un ordre n donné.

Matrices antisymétriques.

32.64. **Définition.** Soit une matrice carrée A d'ordre n. La matrice A est dite **antisymétrique** si elle est égale à sa l'opposé de sa transposée, c'est-à-dire si :

$$A = -A^{\intercal} \tag{125}$$

32.65. **Remarque.** On appelle, en anglais, ces matrices *skew-symmetric*.

Une fois les matrices symétriques et les matrices antisymétriques introduites, on peut montrer un petit résultat (qui a son pendant en analyse réelle, cf. la démonstration) :

32.66. **Proposition.** *Toute matrice est la somme d'une matrice symétrique et d'une matrice antisymétrique.*

Démonstration. Il va falloir raisonner un peu astucieusement. Pour ce faire, on va pouvoir utiliser la même idée que pour montrer que toute fonction peut être écrite comme la somme d'une fonction paire et d'une fonction impaire. On introduit deux quantités :

$$\delta_{\mathrm{S}} = \frac{A + A^{\intercal}}{2} \tag{126}$$

$$\delta_{\mathrm{AS}} = \frac{A - A^{\intercal}}{2} \tag{127}$$

On remarque alors que $A = \delta_{\mathrm{AS}} + \delta_{\mathrm{S}}$. Il ne reste qu'à montrer que δ_{AS} est antisymétrique et que δ_{S} est symétrique. Pour cela, on utilise les résultats de la proposition 32.58 :

$$(\delta_{\mathrm{S}})^{\intercal} = \left[\frac{A + A^{\intercal}}{2}\right]^{\intercal} = \frac{1}{2}\big(A^{\intercal} + (A^{\intercal})^{\intercal}\big) = \frac{1}{2}\big(A^{\intercal} + A\big) = \delta_{\mathrm{S}} \tag{128}$$

Donc δ_{S} est bel et bien une matrice symétrique.

$$(\delta_{\mathrm{AS}})^\intercal = \left[\frac{A - A^\intercal}{2}\right]^\intercal = \frac{1}{2}\big(A^\intercal - (A^\intercal)^\intercal\big) = \frac{1}{2}\big(A^\intercal - A\big) = -\delta_{\mathrm{AS}} \tag{129}$$

Donc δ_{AS} est bel et bien une matrice antisymétrique. □

Matrice adjointe (Transposée du conjugué).

On se donne une matrice A à coefficients complexes. On va chercher à voir s'il n'existerait pas un petit raffinement de ce que l'on connaît comme étant la matrice transposée de A à coefficient réel.

32.67. **Définition.** On appelle matrice adjointe, notée A^* ou A^{H} voire $A^\dagger$, d'une matrice A à coefficients complexes, la matrice transposée de la matrice conjuguée de A. La matrice conjuguée, notée $\overline{A}$, n'étant rien d'autre que que la matrice obtenue en passant au conjugué chacun des éléments de ladite matrice.

Ainsi :

$$A^* = (\overline{A})^\intercal = \overline{(A^\intercal)} \tag{130}$$

Pour l'instant, un tel objet ne nous est d'aucune quelconque utilité. Ce sera intéressant lorsque l'on s'intéressera aux espaces hermitiens (et, *a fortiori*, à plus encore) ! Je ne pense pas qu'on revoie la matrice adjointe de si tôt... donc on va un peu laisser de côté les propriétés afférentes.

32.2.3. *Déterminant.* On arrive désormais sur du lourd, mais malheureusement, on ne va pas vraiment voir en quoi ça peut être utile (changement de variable dans une intégrale multidimensionnelle, volume...). Encore une fois on se cantonne à quelques définitions rudimentaires et quelques propriétés venant d'elles-mêmes.

Au fond le déterminant (au moins en toute petite dimension) est quelque chose de très commun, on en trouve des traces dans les petites classes et notamment en seconde dans le chapitre sur les équations cartésiennes ou encore en spé Mathématiques concernant l'inverse d'une matrice.

À vrai dire, on peut donner quelques propriétés du déterminant (que l'on attend de lui) et montrer qu'il existe bien un objet (et un seul), **le déterminant**, les vérifiant. [8]

On est encore loin de pouvoir prétendre à des formules, exprimant le déterminant, type :

$$\sum_{\sigma \in \mathfrak{S}_n} \epsilon(\sigma) \prod_{j=1}^{n} a_{\sigma(j),j} \tag{131}$$

En réalité, on va voir cette formule mais pas avec tous les "artifices" provenant de la théorie des groupes et de l'algèbre générale.

Au fond, qu'est-ce que le déterminant ? Intuitivement ? On l'a vu les matrices peuvent permettre de considérer l'orthogonalité, des longueurs... et si c'était également possible d'avoir une idée du volume d'un polygone multidimensionnel (engendré par un certain nombre de points). Plusieurs problèmes arrivent : pourquoi un polygone ? qu'est-ce que cette histoire de points ? pourquoi un volume ? Pour la suite, on ne va considérer que des matrices carrées. Sur ces matrices (n, n), on peut se dire que reposent n vecteurs de n composantes : un pour chaque colonne. Au fond, chacun de ces vecteurs (pour peu qu'on le fasse partir de l'origine) va permettre de représenter un point. De manière plus générale, si l'on considère l'ensemble de ces n vecteurs, on obtient n points de l'espace. Il y a fort à penser (on ne le démontre pas) que l'ensemble de ces points forme une figure à n sommets. Et s'il existait une manière de calculer son volume, ce serait merveilleux.

On laisse carrément de côté l'aspect géométrique pour ne se concentrer plus que sur du calcul pur. À moins que... à moins qu'on laisse planer le doute jusqu'à l'année prochaine ! En effet, on ne va rien faire de plus. La grande majorité des notions fait appel à des éléments du supérieur. Disons simplement que l'on a fait une petite introduction à l'existence d'une chose nommée déterminant. Il est trop tôt pour aller sérieusement plus loin.

32.2.4. *Inverse d'une matrice carrée.* On a déjà un tout petit peu évoqué la notion d'**inverse d'une matrice** lorsque l'on définissait ce qu'était la puissance d'une matrice 32.45. Voyons plus précisément ce que c'est. On s'attardera sur l'exemple bien particulier des matrices carrées d'ordre 2 avant d'esquisser quelques considérations générales.

Généralités La notion d'inversibilité pour les matrices n'est rien d'autre que la spécification d'une notion déjà bien connue (par exemple chez les réels). L'on trouve aisément l'inverse de 7 dans $\mathbb{R}$, par exemple. On se retrouve un peu plus embêté dès

8. On verra plus tard des énoncés type : *l'application déterminant en base B est l'unique forme n-linéaire alternée sur un espace E vérifiant* $\det_B(e_1, e_2, \ldots, e_n) = 1$, *abrégé en* $\det_B(B) = 1$.

que l'on réfléchi dans "quelque chose" de "plus petit" que $\mathbb{Q}$ ($\mathbb{Z}$ par exemple). L'on va également pouvoir être "gêné" chez les matrices. Il existe des choses que l'on verra en détail plus tard($\mathrm{GL}_n(K)$, $\mathrm{SL}_n(K)$...) qui utilisent la notion d'inverse et permettent de ne pas trop se poser de question : est-ce que X ou Y est inversible ? C'est là d'ailleurs une question importante (à notre modeste échelle). Le calcul de l'inverse d'une matrice peut être extrêmement laborieux. Mais, restons sur de grandes généralités pour l'instant.

32.68. **Définition.** Soit A une matrice carrée d'ordre n. A est dite **inversible** s'il existe une matrice carrée B d'ordre n telle que $AB = BA = I_n$.

32.69. **Exemple.** Si l'on remarque que

$$A^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^2 = -I_2 \tag{132}$$

que peut-on en déduire ? Que A est inversible d'inverse $-A$. En effet : $A(-A) = I_2$.

32.70. **Remarque.** On jette à la trappe la matrice nulle. En effet, la matrice nulle étant un *élément absorbant*, toute produit d'une matrice quelconque avec la matrice nulle donne la matrice nulle : donc impossible que ce puisse être égal à la matrice identité. Ainsi, la matrice nulle n'est pas inversible (car il n'existe pas d'inverse adéquat).

On commence à s'apercevoir que toutes les matrices ne sont pas forcément inversibles (pour l'instant, on en connaît au moins une).

En effet, étudions plus précisément la matrice $A = \begin{pmatrix} 1 & 4 \\ -\frac{1}{2} & -2 \end{pmatrix}$. Considérons désormais une matrice B quelconque et supposons qu'elle soit l'inverse de la matrice A. Alors, on s'attend à ce que $AB = BA = I_2$. Calculons AB :

$$\begin{pmatrix} 1 & 4 \\ -\frac{1}{2} & -2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+4c & b+4d \\ -\frac{1}{2}a - 2c & -\frac{1}{2}b - 2d \end{pmatrix} = I_2 \tag{133}$$

Cela nous conduit au système de quatre équations suivant (on verra "en détail" cette histoire de systèmes et le lien avec les matrices sous très peu) :

$$\begin{cases} a + 4c & = 1 \\ b + 4d & = 0 \\ -\frac{1}{2}a - 2c & = 0 \\ -\frac{1}{2}b - 2d & = 1 \end{cases} \tag{134}$$

Or problème : si l'on regarde la première et la troisième équation, l'on trouve une contradiction. En effet, d'un côté $a + 4c = 1$ et de l'autre $-\frac{1}{2}a - 2c = 0$ (ce qui revient à, après multiplication par moins deux : $a + 4c = 0$).[9]

En somme, la matrice A n'est pas inversible.

Si l'on regarde la définition 32.68, il existe en réalité un élément superflus qui peut être déduit (et n'a donc par conséquent pas vraiment besoin d'être posé dans la définition, toutefois, c'était le choix de ma professeure).

32.71. **Proposition.** *Soit A une matrice carrée d'ordre n. Si B est une matrice carrée d'ordre n telle que $AB = I_n$, alors $BA = I_n$.*

Réciproquement, si B est une matrice carrée d'ordre n telle que $BA = I_n$, alors $AB = I_n$.

Démonstration. On laisse la réciproque de côté. On ne prouve que la première proposition. (Notons que, avec le "peu de moyen" que l'on a pour l'instant, il va falloir démontrer la propriété en la travaillant au corps. En appliquant les définitions fondamentales).

On va chercher à prouver la commutativité. L'idée est simple : si B est l'inverse de A, on s'attend à ce que A soit également l'inverse de B.

Tout d'abord, on sait que l'on aura aucun problème de compatibilité du produit dans la mesure où les matrices A et B sont toutes deux carrées d'ordre n. Les produits AB et BA sont donc bien définis.

9. Il conviendrait une meilleure justification. L'idée est qu'une somme de choses linéaires ne peut pas avoir deux valeurs possibles.

Je suis con, à la base je m'étais fait tout un bins à essayer de raisonner sur les coefficients etc... en cherchant à montrer des choses type $\sum_{k=1}^{n}(ba)_{ii} = \sum_{k=1}^{n}(ab)_{ii}$. (Le problème étant que le résultat n'est pas suffisamment fort, il ne caractérise pas suffisamment bien ce qui nous intéresse.)

En fait, on peut juste raisonner sur la transposée : $I_n = (I_n)^\intercal = (AB)^\intercal = BA$ et hop c'est plié. □

L'intérêt d'une telle proposition permet de ne retenir **que** le critère suivant : pour montrer qu'une matrice A est inversible, **il suffit** de montrer qu'il existe une matrice B telle que $AB = I_n$. (Sans avoir à se préoccuper de la réciproque.)

Montrons désormais un petit résultat important qui scelle définitivement le sort de l'inverse d'une matrice :

32.72. **Proposition.** *Soit A une matrice carrée d'ordre n. Si A est inversible alors l'**unique** matrice B d'ordre n telle que $AB = I_n$ est appelée **matrice inverse de** A et est notée A^{-1}.*

Démonstration. Raisonnons par l'absurde pour prouver que l'inverse d'une matrice est unique. À cet effet, supposons que A possède plus d'un inverse, disons 2 pour fixer les choses que l'on va noter B_1 et B_2. Par hypothèse, on sait que $B_1A = AB_1 = I_n = B_2A = AB_2$. On cherche alors à montrer que $B_1 = B_2$. En vertu de la réciproque de la proposition 32.41, on sait que l'on ne peut pas procéder en utilisant une équation produit nul (c'est-à-dire former le produit $A(B_1 - B_2) = 0_n$ et en déduire que $B_1 = B_2$).

On considère la quantité suivante : B_1AB_2, elle est égale à $B_1I_n = B_1$ ou encore à I_nB_2, ce qui permet de conclure que $B_1 = B_2$. En effet, plus visuellement, on a :

$$\underbrace{B_1A}_{=I_n}B_2 = B_1\underbrace{AB_2}_{=I_n} \tag{135}$$

□

Enfin, avant de s'attarder sur le cas particulier des matrices carrées d'ordre 2 et d'entrapercevoir ce qu'il se passe pour n quelconque, voyons quelques propriétés de l'inverse.

32.73. **Proposition.** *Soit A une matrice carrée d'ordre n. On suppose que A est inversible.*

1. $AA^{-1} = A^{-1}A = I_n$ *;*
2. $A^nA^{-1} = A^{n-1}$ *;*
3. $(A^{-1})^{-1} = A$ *;*
4. $(AB)^{-1} = B^{-1}A^{-1}$*, pour B une matrice carrée inversible d'ordre n ;*
5. $(\alpha A)^{-1} = \alpha^{-1}A^{-1}$*, pour un α scalaire non nul ;*
6. $(A^\intercal)^{-1} = (A^{-1})^\intercal$*, si $A^\intercal$ est inversible ; et,*
7. $AC = BC$ *implique $A = B$ si C est une matrice carrée inversible d'ordre n et B une matrice carrée d'ordre n.*

Démonstration.
1. par définition, c'est une évidence.
2. idem.
3. par application de la propriété numéro 1, on a que $(A^{-1})^{-1}A^{-1} = I_n$, d'où : $(A^{-1})^{-1}A^{-1}A = I_nA = A$.
4. dans la même logique que pour la démonstration de la propriété numéro 3, on a : $(AB)^{-1}AB = I_n$, d'où : $(AB)^{-1}ABB^{-1} = I_nB^{-1} = B^{-1}$ et enfin : $(AB)^{-1}AA^{-1} = B^{-1}A^{-1}$.
5. c'est une conséquence de la propriété numéro 4 en remarquant simplement que la multiplication par un scalaire est commutative (en notant que l'inverse d'un scalaire reste et est toujours un scalaire, pour peu que l'on considère le bon "ensemble", mais ici aucun problème avec les réels ou les complexes par exemple, tant que α est différent de zéro du moins). En clair, en appliquant la propriété 4, on trouve :

$$(\alpha A)^{-1} = A^{-1}\alpha^{-1} = \alpha^{-1}A^{-1} \tag{136}$$

6. sans plus de surprise que pour les démonstrations des propriétés 3 et 4, on raisonne de manière similaire. Néanmoins, nous allons utiliser une propriété supplémentaire : $(AB)^\intercal = B^\intercal A^\intercal$ (cf. proposition 32.58). Ainsi : $(A^{-1})^\intercal A^\intercal = (AA^{-1})^\intercal = I_n^\intercal = I_n$, d'où alors : $(A^{-1})^\intercal A^\intercal(A^\intercal)^{-1} = I_n(A^\intercal)^{-1}$. Ce qui donne le résultat escompté : $(A^{-1})^\intercal = (A^\intercal)^{-1}$.

7. on entraperçoit quelque peu la section suivante sur les systèmes linéaires. Démontrons cette propriété afin de commencer à se familiariser avec ces derniers. On a supposé C inversible donc multiplions à gauche par l'inverse de C :

$$ACC^{-1} = BCC^{-1} \text{ qui devient } A\underbrace{CC^{-1}}_{=I_n} = B\underbrace{CC^{-1}}_{=I_n} \tag{137}$$

d'où $A = B$.

□

Matrices carrées d'ordre 2

32.74. **Proposition.** *Soit A la matrice carrée d'ordre 2 suivante $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. On suppose A non nulle. Alors, si A est inversible, son inverse s'exprime tel que :*

$$A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{138}$$

Démonstration. Il n'y a qu'à faire le calcul :

$$AA^{-1} = \left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right]\left[\frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\right] = \frac{1}{ad - bc}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{139}$$

$$AA^{-1} = \frac{1}{ad - bc}\begin{pmatrix} ad - bc & -ab + ba \\ cd - dc & -cb + da \end{pmatrix} = \frac{1}{ad - bc}\begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = I_2 \tag{140}$$

□

32.75. **Remarque.** On remarque alors qu'une matrice carrée d'ordre 2 n'est inversible qu'à la condition suivante : que $ad - bc$ soit non nul (ce qui revient à postuler que le déterminant de ladite matrice soit non nul).

Matrices carrées d'ordre n

32.76. **Proposition.** *Soit A une matrice carrée d'ordre n. On suppose A non nul. Alors, si A est inversible, son inverse s'exprime tel que :*

$$A^{-1} = \frac{1}{\det(A)}\operatorname{com}(A)^{\intercal} \tag{141}$$

32.77. **Remarque.** De manière similaire au cas $n = 2$, on déduit que la matrice est inversible si et seulement si son déterminant est différent de zéro.

On ne définit pas pour l'instant ce qu'est ce $\operatorname{com}(A)$, la comatrice, on verra plus tard. Notons toutefois, en fin de compte que la méthode de la proposition 32.76 pour calculer l'inverse d'une matrice se fait supplanter par d'autres méthodes, par exemple celle du pivot de Gauss que l'on aura l'occasion de voir dans quelques mois.

33. Formule BCH vue par un élève de Terminale

Je me demande si la première fois que j'en ai entendu parler n'était pas un document de J. Feydy.
L'histoire semble commencer avec des matrices et finit dans des *algèbres de Lie de groupes de Lie* (qu'est-ce ?).

Commençons par le commencement : on se souvient des problèmes de commutativité que l'on pouvait rencontrer avec les matrices (cf. la section 32.1.2). De manière générale, on est bien loin d'avoir l'égalité suivante :

$$AB = BA \tag{142}$$

où A et B sont deux matrices compatibles pour le produit (on va même dire carrées tant qu'à faire).

Alors alors comment faire ? Comment "*dériver*" la formule de Baker-Campbell-Hausdorff ? Comme un couillon, je pensais qu'une simple étude du Taylor (ordre par ordre) de $\exp(X)\exp(Y)$ suffirait... Un petit programme [10], et c'était réglé :

10. Je ne sais pas si fallait spécifier un truc type *commutative=False*.

```
from sympy import exp, MatrixSymbol, Sum, factorial
from sympy import simplify, expand
from sympy.abc import k

x = MatrixSymbol('X', 2, 2)
y = MatrixSymbol('Y', 2, 2)
z = MatrixSymbol('Z', 2, 2)


def exp_mat(mat, N):
return Sum(mat**k/factorial(k), (k, 0, N)).doit()


for N in range(5):
print(N, " -- ", expand(exp_mat(x, N) * exp_mat(y, N)))
print(" ", " -- ", exp_mat(z, N))
```

C'est peut-être un peu plus compliqué que prévu. Mais à vrai dire, je ne crois même pas réfléchir sur les bons objets (ça va péter dans la théorie des groupes de Lie, géométrie différentielle). L'énoncé précis est le suivant :

33.1. **Proposition.** (Tu, 2004) *Soient G un groupe de Lie et $\mathfrak{g} = T_1G$ son algèbre de Lie, vue comme l'espace tangent à G en l'élément neutre 1. Alors il existe un voisinage V de 0 dans $\mathfrak{g}$ et un voisinage U de 1 dans G tels que la restriction de l'application exponentielle* $\exp : V \to U$ *est un difféomorphisme. Son inverse $U \to V$ s'appelle le logarithme et s'écrit* $\log$*. Si deux éléments A et B de $\mathfrak{g}$ sont suffisamment proches de l'origine, alors la formule de Campbell-Hausdorff donne l'expression de* $\log(\exp(A)\exp(B))$ *en tant que série entière dans l'algèbre de Lie engendrée par A et B :*

$$\exp(A)\exp(B) = \exp\Big(A + B + \frac{1}{2}[A, B] + \frac{1}{12}([A, [A, B]] + [B, [B, A]]) + \dots\Big). \tag{143}$$

De manière tout aussi technique, on trouve l'énoncé suivant :

33.2. **Proposition.** *Let G be a Lie group of Lie algebra $\mathfrak{g}$ and let X, Y be elements of the Lie algebra $\mathfrak{g}$. Then :*

$$\log(\exp(X)\exp(Y)) = X + \int_0^1 \psi(\exp(ad_X)\exp(t ad_Y))Y \mathrm{d}t \tag{144}$$

whenever the $\log(\exp(X)\exp(Y))$ *is defined.*

De manière générale, la formule BCH joue un rôle important dans différents domaines, par exemple :

- Mathematics : theory of linear differential equations, Lie groups, numerical analysis of differential equations ;
- Theoretical Physics : perturbation theory, quantum mechanics, statistical mechanics, quantum computing ; et,
- Control theory : design and analysis of nonlinear control mechanisms, nonlinear filters, stabilization of rigid bodies,...

En attendant, il y a un sujet du second concours (2009, ÉNS de Lyon) qui démontre élémentairement un cas bien particulier.

J'espère avoir l'occasion de revenir grandement sur ce sujet plus tard. Afin de ne pas le laisser trop vacant, je vais juste modifier un petit peu le programme au-dessus car je crois avoir trouvé le moyen de moyenner (et de parvenir au but) !

```
from sympy import Symbol, Sum, factorial
from sympy import expand
from sympy.abc import k
from sympy import latex

x = Symbol('X', commutative=False)
y = Symbol('Y', commutative=False)

def exp_mat(mat, N):
return Sum(mat**k/factorial(k), (k, 0, N)).doit()
```

```
def log_mat(mat, N):
return Sum((−1)**(k + 1) * (mat)**k/k, (k, 1, N)).doit()


for N in range(3):
print(N, " — exp — ", latex(expand(exp_mat(x, N) * exp_mat(y, N), N)))
print(N, " — log(exp) — ", latex(expand(log_mat(− 1 + expand(exp_mat(x, N) * exp_mat(y,
```

En réalité, il fallait totalement changer l'angle d'approche [11] et chercher à calculer $\log(\exp(X)\exp(Y))$ plutôt que simplement $\exp(X)\exp(Y)$. (Ce qui est au fond très logique étant donné que l'on cherche Z plutôt que $\exp(Z)$ dans l'équation $\exp(Z) = \exp(X)\exp(Y)$.)

34. Quelques dates concernant les fonctions elliptiques

La première fois que j'ai entendu parler de ça, j'étais stupéfait. Je me souviens avoir essayé d'expliquer le miracle de la double périodicité à un camarade de première.

L'idée n'est pas de démontrer des choses, juste de découvrir et de s'émerveiller un max !

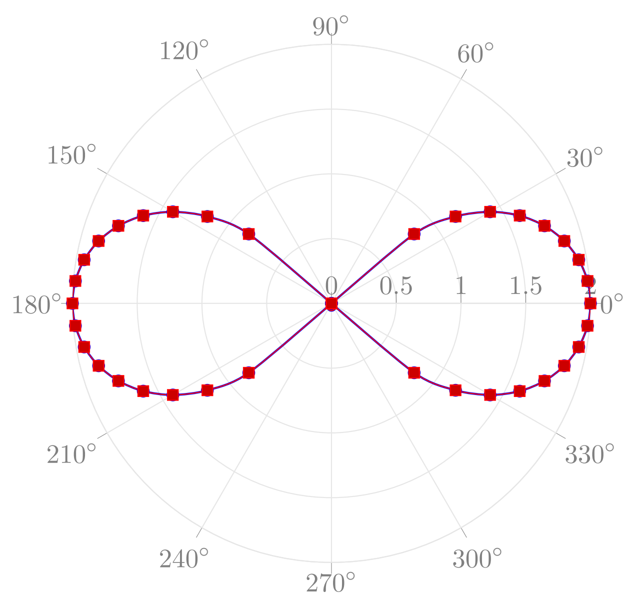
Voici une petite présentation historique compilée par Emmanuel Moreau que je me suis permis d'augmenter :

1665 : **Wallis et Newton** : le premier problème fut d'exprimer la longueur d'une ellipse. Des représentations intégrales fleurissent :

$$L_{\text{ellipse}} = 4\int_0^{\pi/2} \sqrt{a^2\sin^2(t) + b^2\cos^2(t)}\mathrm{d}t = 4a\int_0^{\pi/2} \sqrt{1 - e^2\cos^2(t)}\mathrm{d}t \tag{145}$$

où $e^2 = 1 - \dfrac{b^2}{a^2}$. Le but est d'aller plus loin et de réussir à exprimer L avec des fonctions usuelles... faudra attendre quelques temps pour se rendre compte que ce n'est pas trop possible.

1694 : **Bernoulli** : la fameuse lemniscate de Bernoulli voit le jour et sera apparemment "*l'un des outils principaux dans la découverte des fonctions elliptiques*".



11. Je me demande si je n'ai pas fait une erreur, que ce soit sur le développement en série du logarithme ou ailleurs ? C'est quand même étonnant d'avoir autant de signes moins...

1750 : **Fagnano** : mathématicien amateur apparemment qui parviendra à calculer l'aire définie par la lemniscate de Bernoulli mais dont le calcul de la longueur posait problème. Néanmoins, il parviendra à montrer que :

$$L_{\text{lemniscate}} = \frac{4}{\sqrt{2}} \int_0^{\pi/2} \frac{\mathrm{d}\theta}{\sqrt{1-(1/2)\sin^2(\theta)}}. \tag{146}$$

1751 : **Euler** : lit les travaux de Fagnano. Obtiendra la formule dite *d'addition* et tentera de pousser le travail de Fagnano sur le calcul de $L_{\text{lemniscate}}$ mais ne parviendra pas à aller jusqu'au bout.

1784 : **Lagrange** : a le mérite de définir ce qu'est une intégrale elliptique :

34.1. **Définition.** On appelle **intégrale elliptique** une intégrale de la forme $\int R(x,y)\mathrm{d}x$ où R est une fraction rationnelle en x et y et où $y = \sqrt{P(x)}$ où P est un polynôme de degré 3 ou 4 sans racine multiple.

1791 : **Gauss** : s'intéresse au calcul de la longueur de la lemniscate de Bernoulli. S'intéresse également à la moyenne arithmético-géométrique de a et b notée $M(a,b)$. Prouve la constructibilité du polygone régulier à 17 côtés et la loi de réciprocité quadratique. Réussi à exprimer la longueur de la lemniscate :

$$L_{\text{lemniscate}} = \frac{2\pi}{M(1,\sqrt{2})} \tag{147}$$

[12] toutefois, ce n'est toujours pas une fonction usuelle.

Introduit le *sinus lemniscatique* ($sl(x)$) et le *cosinus lemniscatique* ($cl(x)$) :

$$sl(x) = G^{-1}(x) \text{ où } G(x) = \int_0^x \frac{\mathrm{d}t}{\sqrt{1-t^4}} \tag{148}$$

$$cl(x) = \tilde{G}^{-1}(x) \text{ où } \tilde{G}(x) = \int_x^1 \frac{\mathrm{d}t}{\sqrt{1-t^4}}. \tag{149}$$

Dans le livre d'André, les sinus et cosinus lemniscatiques apparaissent et se pose le problème des relations algébriques que peuvent bien satisfaire les nombres $\Gamma(a)$, $a \in \mathbb{Q}\setminus -\mathbb{N}$, (et de les interpréter dans le contexte des motifs).

Des formules analogues à celles de la trigonométrie classique vont pouvoirs être trouvées (addition, multiplication).

1793 : **Legendre** : "*parvient à démontrer que les intégrales elliptiques peuvent se ramener à l'une des trois formes canoniques suivantes*" :

- intégrale elliptique de première espèce :

$$F(u) = \int_0^u \frac{\mathrm{d}u}{\sqrt{1-k^2\sin^2(u)}} = \int \frac{\mathrm{d}x}{\sqrt{(1-x^2)(1-k^2x^2)}} \tag{150}$$

- intégrale elliptique de seconde espèce :

$$E(u) = \int_0^u \sqrt{1-k^2\sin^2(u)}\mathrm{d}u = \int \sqrt{\frac{1-k^2x^2}{1-x^2}}\mathrm{d}x \tag{151}$$

- intégrale elliptique de troisième espèce :

$$\Pi(u) = \int_0^u \frac{\mathrm{d}u}{(1+n\sin^2(u))\sqrt{1-k^2\sin^2(u)}} = \int \frac{\mathrm{d}x}{(1+nx^2)\sqrt{(1-x^2)(1-k^2x^2)}} \tag{152}$$

1827 : **Abel et Jacobi** : à noter que la période 1827 – 1832 est également celle de la correspondance Jacobi / Legendre. Abel et Jacobi réaliseront indépendamment des travaux similaires. Jacobi introduit une fonction amplitude ($am(x)$) :

$$am(x) = F^{-1}(x) \text{ où } F(x) = \int_0^x \frac{\mathrm{d}\theta}{\sqrt{1-k^2\sin^2(\theta)}} \tag{153}$$

12. Ici l'importance n'est "que" visuelle (gné), mais j'ai un doute sur la formule. Dans le Gourdon, par exemple, on a $L_{\text{lemniscate}} = \dfrac{\pi}{2M(\sqrt{2},1)}$ plutôt hé.

ainsi il définira : $sn(x) = \sin(am(x))$ et $cn(x) = \cos(am(x))$. Les fonctions (thêta) de Jacobi semblent plus commodes que celles de Gauss.

1833 : **Liouville** : "*parvient à montrer que les intégrales elliptiques ne peuvent pas s'exprimer à l'aide des fonctions usuelles*".

1858 : **Hermite** : "*prouve que les solutions d'une équation polynomiale de degré* 5 *peuvent toujours s'exprimer à l'aide des fonctions thêta de Jacobi*".

1864 : **Clebsch** : utilisation de la possibilité de paramétrer toute courbe cubique à l'aide des fonctions elliptiques pour démontrer un théorème de Jacob Steiner.

1867 : **Weierstrass** : fonction elliptique de Weierstrass $\wp$ (début de toute une théorie[13]) qui satisfait à l'équation différentielle suivante : $(\wp')^2 = 4\wp^3 - g_2\wp - g_3$.

1973 : **Lang** : là ça part totalement en couilles, Serge Lang sort son bouquin *Elliptic functions* et depuis 1867 il a du s'en passer (beaucoup) de choses. Pour le beau jeu, je propose de zieuter le livre de Lang.

(Il semblerait que Riemann avec ses surfaces ait également eu une certaine importance (?). De plus, à noter que Steven Wepster semble avoir des documents intéressants un peu cachés sur son page professionnelle. Il y a aussi un texte de Mittag-Leffler qui a l'air bien intéressant ! Boh, on a même pas parlé de double périodicité encore, gné. Je ne vais même pas fourrer le nez dans le texte de Weil mais n'oublions pas Kronecker et Eisenstein ! ! !)

C'est quand même fou que ça ait commencé avec des considérations sur les ellipses et que ça en arrive aujourd'hui à de la théorie algébrique des nombres et cie.

Aujourd'hui, il semblerait que les fonctions elliptiques aient carrément pris un nouveau tournant. Par exemple, dans un bouquin de Stevenhagen, on trouve cette définition :

34.2. **Définition.** An **elliptic function** *with respect to a lattice* $\Lambda = \mathbf{Z}\omega_1 + \mathbf{Z}\omega_2$ in $\mathbf{C}$ is a meromorphic function f on $\mathbf{C}$ that satisfies $f(z+\omega) = f(z)$ for all $\omega \in \Lambda$.

À noter que Lang a une définition qui semble clairement équivalente.

Putain de merde, y'aurait même une fonction zêta de Weierstrass pour le réseau (=*lattice*) $\Lambda = \mathbf{Z}\omega_1 + \mathbf{Z}\omega_2$ dans $\mathbf{C}$:

$$\zeta(z) = \frac{1}{z} + \sum_{\omega \in \Lambda'} \left(\frac{1}{z-\omega} + \frac{1}{\omega} + \frac{z}{\omega^2} \right) \tag{154}$$

Apparemment, cette fonction zêta peut se réécrire comme une somme infinie de séries d'Eisenstein (?), richesse richesse !

J'écris ça à la fin, tiens c'est bien le moment..., ce serait peut-être pas si mal de tout reprendre proprement un jour ? ! En faisant des calculs et plus que juste présenter deux ou trois choses (en se tapant le livre de Lang par exemple ?) ! ! !

Il y a notamment un polycopié d'une réunion APMEP (2006) qui a l'air vraiment bien et simple à suivre.

35. Young et Hölder vus par un élève de Terminale

Exercices 184 et 185 issus de du poly LLG.

Exercice 184 (Inégalité de Young)

Soit p un réel supérieur à 1.

a) Montrer qu'il existe un unique réel q tel que :

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{155}$$

Vérifier que q (que l'on appelle parfois *exposant conjugué de* p) est supérieur à 1. Déterminer q pour $p = 2$, puis pour $p = 4$.

L'existence de q est intuitivement très évidente. En effet, lorsque l'on travaille dans les réels, il est évident que la différence $1 - \dfrac{1}{p}$ existe. (Autrement pour le montrer plus formellement, ne pourrait-on pas utiliser le fait que $\mathbb{R}$ soit archimédien combiné

13. Théorème de Siegel de 1932 ; Schneider 1936 ; Masser 1975 ; Chudnovsky 1976, 1981... cf Waldschmidt.

au fait qu'il soit complet ou quelque chose comme ça ? Sinon on sort directement l'artillerie lourde en utilisant le théorème des valeurs intermédiaires (vu que la fonction représentant une droite est continue).)

Une fois l'existence montrée, il reste à prouver l'unicité. Pour cela, faisons un raisonnement par l'absurde : supposons qu'il existe q et q' deux exposants conjugués de p différents. Alors :

$$\frac{1}{p} + \frac{1}{q} = 1 = \frac{1}{p} + \frac{1}{q'} \tag{156}$$

d'où :

$$\cancel{\frac{1}{p}} + \frac{1}{q} = \cancel{\frac{1}{p}} + \frac{1}{q'} \tag{157}$$

on trouve finalement que $q = q'$. Il est donc absurde de les supposer différents, donc ils sont égaux, ainsi q est unique.

Pour démontrer que q est supérieur à 1, il n'y qu'à manipuler l'expression 155 :

$$\frac{1}{q} = 1 - \frac{1}{p} = \frac{p-1}{p}. \tag{158}$$

En inversant, on obtient finalement que $q > \frac{p}{p-1}$, ce qui est évidemment supérieur à 1.

Il reste tout de même deux petites questions : déterminer q si :

- $p = 2$: $\frac{1}{q} = 1 - \frac{1}{2} = \frac{1}{2}$, d'où $q = 2$.
- $p = 4$: $\frac{1}{q} = 1 - \frac{1}{4} = \frac{3}{4}$, d'où $q = \frac{4}{3}$.

b) On fixe y dans $\mathbb{R}^{+*}$ et on pose :

$$\forall x \in \mathbb{R}^{+*}, f(x) = \frac{x^p}{p} + \frac{y^q}{q} - xy. \tag{159}$$

Donner le tableau de variations de f.

Calculons la dérivée de f (dérivable car somme de fonctions dérivables) :

$$f'(x) = x^{p-1} - y \tag{160}$$

Le tableau de variation est alors :

x	0^+		$y^{\frac{1}{p-1}}$		∞
$f'(x)$		$-$	0	$+$	
$f(x)$		$\searrow$	$f(y^{\frac{1}{p-1}})$	$\nearrow$	

Ainsi, on en déduit que $f(x) \geq f(y^{\frac{1}{p-1}})$ pour tout $x \in \mathbb{R}^{+*}$.

c) Conclure :

$$\forall (x,y) \in (\mathbb{R}^{+*})^2, xy \leq \frac{x^p}{p} + \frac{y^q}{q} \tag{161}$$

Évaluons le minimum de f :

$$f(y^{\frac{1}{p-1}}) = \frac{(y^{\frac{1}{p-1}})^p}{p} + \frac{y^q}{q} - y^{\frac{1}{p-1}}y = \frac{y^{\frac{p}{p-1}}}{p} + \frac{y^q}{q} - y^{\frac{1}{p-1}+1} \tag{162}$$

Or, $\frac{p}{p-1} = q$, d'où :

$$f(y^{\frac{1}{p-1}}) = \frac{y^q}{p} + \frac{y^q}{q} - y^{\frac{1}{p-1}+1} = y^q \underbrace{\left[\frac{1}{p} + \frac{1}{q}\right]}_{=1} - y^{\frac{1}{p-1}+1} \tag{163}$$

Pour conclure, il suffit de remarquer que $\dfrac{1}{p-1}+1=\dfrac{p}{p-1}=q$, d'où : $f(y^{\frac{1}{p-1}})=0$. En somme : $f(x)\geq 0$ pour tout $x\in\mathbb{R}^{+*}$. Ce qui permet d'obtenir le résultat escompté.

Exercice 185 (Inégalité de Hölder pour les intégrales)

Les notations p, q sont celles de l'exercice précédent, dont on utilise également le résultat. Soient a et b deux réels tels que $a < b$, f et g deux fonctions continues de $[a, b]$ dans $\mathbb{R}$. On se propose d'établir l'inégalité de Hölder :

$$\left|\int_a^b f(t)g(t)\mathrm{d}t\right| \leq \left(\int_a^b |f(t)|^p\mathrm{d}t\right)^{1/p}\left(\int_a^b |g(t)|^q\mathrm{d}t\right)^{1/q} \tag{164}$$

On remarquera que, pour $p = 2$, on obtient l'inégalité de Cauchy-Schwarz.

a) En utilisant l'inégalité de Young, montrer, pour $\lambda\in\mathbb{R}^{+*}$:

$$\left|\int_a^b f(t)g(t)\mathrm{d}t\right| \leq \frac{\lambda^p}{p}\int_a^b |f(t)|^p\mathrm{d}t+\frac{1}{q\lambda^q}\int_a^b |g(t)|^q\mathrm{d}t. \tag{165}$$

Il suffit de multiplier astucieusement $|f(t)g(t)|$ pour obtenir le résultat en appliquant l'inégalité de Young :

$$|f(t)g(t)| = |\lambda f(t)\frac{g(t)}{\lambda}| \leq \frac{|\lambda f(t)|^p}{p}+\frac{|\frac{g(t)}{\lambda}|^q}{q} \tag{166}$$

étant donné que λ est strictement positif, on peut le sortir des valeurs absolues. Ensuite, on intègre des deux côtés en utilisant la linéarité de l'intégrale pour le côté de droit et le fait que $\left|\int\cdot\right|\leq\int|\cdot|$ pour conclure.

b) Déterminer le minimum de la fonction :

$$\psi : \lambda\in\mathbb{R}^{+*}\mapsto \frac{\lambda^p}{p}\int_a^b |f(t)|^p\mathrm{d}t+\frac{1}{q\lambda^q}\int_a^b |g(t)|^q\mathrm{d}t. \tag{167}$$

Conclure.

La fonction ψ est dérivable (car somme de fonctions dérivables), sa dérivée vaut :

$$\psi'(\lambda) = \lambda^{p-1}\int_a^b |f(t)|^p\mathrm{d}t-\frac{1}{\lambda^{q+1}}\int_a^b |g(t)|^q\mathrm{d}t \tag{168}$$

On cherche (sauvagement) à résoudre l'équation : $\psi'(\lambda)=0$:

$$\lambda^{p-1}\int_a^b |f(t)|^p\mathrm{d}t = \frac{1}{\lambda^{q+1}}\int_a^b |g(t)|^q\mathrm{d}t \tag{169}$$

Donc :

$$\lambda^{p+q} = \frac{\displaystyle\int_a^b |g(t)|^q\mathrm{d}t}{\displaystyle\int_a^b |f(t)|^p\mathrm{d}t} \tag{170}$$

$$\lambda = \left[\frac{\displaystyle\int_a^b |g(t)|^q\mathrm{d}t}{\displaystyle\int_a^b |f(t)|^p\mathrm{d}t}\right]^{\frac{1}{p+q}}. \tag{171}$$

Et pour finir, calculons $\psi(\lambda)$:

$$\psi(\lambda) = \frac{1}{p}\left[\frac{\displaystyle\int_a^b |g(t)|^q \mathrm{d}t}{\displaystyle\int_a^b |f(t)|^p \mathrm{d}t}\right]^{\frac{p}{p+q}} \int_a^b |f(t)|^p \mathrm{d}t + \frac{1}{q}\frac{1}{\left[\frac{\displaystyle\int_a^b |g(t)|^q \mathrm{d}t}{\displaystyle\int_a^b |f(t)|^p \mathrm{d}t}\right]^{\frac{q}{p+q}}} \int_a^b |g(t)|^q \mathrm{d}t. \tag{172}$$

Il convient de remarquer que $\dfrac{p}{p+q} = \dfrac{p}{pq(\frac{1}{p}+\frac{1}{q})} = \dfrac{1}{q}$ et $\dfrac{q}{p+q} = \dfrac{q}{pq(\frac{1}{p}+\frac{1}{q})} = \dfrac{1}{p}$.

$$\psi(\lambda) = \frac{1}{p}\left(\int_a^b |g(t)|^q \mathrm{d}t\right)^{1/q}\left(\int_a^b |f(t)|^p \mathrm{d}t\right)^{1-1/q} + \frac{1}{q}\left(\int_a^b |g(t)|^q \mathrm{d}t\right)^{1-1/p}\left(\int_a^b |f(t)|^p \mathrm{d}t\right)^{1/p} \tag{173}$$

$$\psi(\lambda) = \frac{1}{p}\left(\int_a^b |g(t)|^q \mathrm{d}t\right)^{1/q}\left(\int_a^b |f(t)|^p \mathrm{d}t\right)^{1/p} + \frac{1}{q}\left(\int_a^b |g(t)|^q \mathrm{d}t\right)^{1/q}\left(\int_a^b |f(t)|^p \mathrm{d}t\right)^{1/p} \tag{174}$$

$$\psi(\lambda) = \left(\int_a^b |f(t)|^p \mathrm{d}t\right)^{1/p}\left(\int_a^b |g(t)|^q \mathrm{d}t\right)^{1/q}\left[\frac{1}{p}+\frac{1}{q}\right] \tag{175}$$

D'où finalement l'inégalité de Hölder :

$$\left|\int_a^b f(t)g(t)\mathrm{d}t\right| \leq \psi(\lambda) = \left(\int_a^b |f(t)|^p \mathrm{d}t\right)^{1/p}\left(\int_a^b |g(t)|^q \mathrm{d}t\right)^{1/q} \tag{176}$$

36. Multiple zêta values (MZV)

37. Équivalence de normes combinées

Une utilisation raisonnée du théorème de Taylor-Lagrange faute dans le Skandalis de *Topologie* nous donne l'équivalence entre les deux normes sur $\mathcal{C}^2([0,1],\mathbf{R})$ suivantes :

$$N_1 : f \mapsto \|f\|_\infty + \|f'\|_\infty, \quad N_2 : f \mapsto \|f\|_\infty + \|f'\|_\infty + \|f''\|_\infty.$$

À leur sujet, notamment leur équivalence, on se référera à 37.

On montre sans soucis que ce sont des normes. Ensuite, une petite astuce (que l'on va revisiter) nous donne, pour tout t dans $[0,1]$, l'existence de x et y dans $[0,1]$ tels que :

$$f(1) - f(0) = f'(t) - \frac{t^2}{2}f''(x) + \frac{(1-t)^2}{2}f''(y). \tag{177}$$

L'équivalence découle de la formule ci-dessus. Exercice : le montrer. (Une majoration est évidente, l'autre utilise effectivement la formule ci-dessus et l'on doit trouver $N_2(f) \leq 3N_1(f)$.) On est alors très tenté de se demander si la méthode se généralise bien ? Énonçons l'idée de généralisation, en sachant pertinemment que le problème est mal posé. Formuler correctement le problème sera le premier des labeurs.

37.1. **Exercice.** Soit I et J deux ensembles d'indices inclus dans $\mathbf{N}$. Nous notons respectivement ι et γ les indices maximaux de I et J. Définissons les deux quantités suivantes pour f dans $\mathcal{C}^\iota([a,b],\mathbf{R})$ ou $\mathcal{C}^\gamma([a,b],\mathbf{R})$:

$$p(f) = \sum_{i\in I}\alpha_i \left\|f^{(i)}\right\|_\infty, \quad q(f) = \sum_{j\in J}\beta_j \left\|f^{(j)}\right\|_\infty. \tag{178}$$

Lorsque p et q définissent des normes, quand sont-elles équivalentes ?

Commençons par donner une condition nécessaire pour que p définisse une norme. Pour l'instant, le discours n'est fait que sur p par commodité. Par symétrie de construction, les énoncés valent également pour q.

37.2. **Proposition.** *Il est nécessaire que l'ensemble I contienne l'indice 0 pour que p puisse définir une norme.*

Démonstration. Supposons que $0 \notin I$. Montrons que dans ce cas p ne vérifie pas la propriété de séparation. Par supposition, on dispose de l'écriture de p sous la forme :

$$p(f) = \alpha_{i_1} \left\| f^{(i_1)} \right\|_\infty + \alpha_{i_2} \left\| f^{(i_2)} \right\|_\infty + \dots \quad \text{avec } i_1 \geq 1 \text{ et } (i_n)_n \text{ une suite strictement croissante.}$$

Cherchons à déterminer les f tels que $p(f) = 0$. Cette équation est équivalente à avoir simultanément :

$$\left| f^{(i_1)}(t) \right| = 0, \quad \left| f^{(i_2)}(t) \right| = 0, \quad \dots \tag{179}$$

Or, $i_1 \geq 1$, donc f est a minima constante (voire linéaire, quadratique...) donc non nécessairement nulle. □

Dès que $0 \in I$, la propriété de séparation sera acquise dès que les coefficients α_i seront strictement positifs, cf. l'équation (1).

37.3. **Proposition.** *L'homogénéité et l'inégalité triangulaire sont vérifiées.*

Démonstration. Soient $\lambda \in \mathbf{R}$, f et g deux fonctions (on donnera plus tard des hypothèses plus précises, se dire que ce sont essentiellement des fonctions suffisamment continûment dérivables). Alors :

$$p(\lambda f) = \sum_{i \in I} \alpha_i \left\| (\lambda f)^{(i)} \right\|_\infty = \sum_{i \in I} \alpha_i |\lambda| \left\| f^{(i)} \right\|_\infty = \lambda p(f), \tag{180}$$

$$p(f+g) = \sum_{i \in I} \alpha_i \left\| (f+g)^{(i)} \right\|_\infty = \sum_{i \in I} \alpha_i \left\| f^{(i)} + g^{(i)} \right\|_\infty \leq \sum_{i \in I} \alpha_i \left(\left\| f^{(i)} \right\|_\infty + \left\| g^{(i)} \right\|_\infty \right) = p(f) + p(g). \tag{181}$$

□

En somme p et q sont des normes si, et seulement si, $0 \in I$ et $0 \in J$ et si les coefficients sont des réels strictement positifs.

Rappelons le théorème de Taylor-Lagrange (on se souvient très bien de l'impression qu'a eu la classe quand J. Royer nous a montré la voie pour démontrer ce théorème : redresser convenablement la bonne fonction).

37.4. **Théorème.** *Soit $n \in \mathbf{N}$. Soient $a, b \in \mathbf{R}$. Soit f une fonction de classe $\mathcal{C}^n$ sur $[a,b]$ et $n+1$ fois dérivable sur $]a,b[$. Alors, il existe $c \in]a,b[$ tel que :*

$$f(b) = \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(c). \tag{182}$$

Pour simplifier le problème, concentrons nous sur l'intervalle $[0,1]$. La formule de Taylor Lagrange devient :

$$f(1) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(0) + \frac{1}{(n+1)!} f^{(n+1)}(c).$$

On peut obtenir un découpage un petit peu plus fin. Soit (h_n) les points d'une subdivision de $[0,1]$, c'est-à-dire que l'on a $[0,1] = [0,h_1] \cup [h_1,h_2] \cup \ldots \cup [h_n,1]$ avec $h_i < h_j$ si $i < j$. Alors, on dispose de l'égalité suivante :

$$f(1) - f(0) = \big(f(1) - f(h_n)\big) + \big(f(h_n) - f(h_{n-1})\big) + \big(f(h_{n-1}) - f(h_{n-2}))\big) + \ldots + \big(f(h_1) - f(0)\big).$$

À vrai dire, on doit pouvoir se contenter de prendre un $t \in [0,1]$ et d'avoir $f(1) - f(0) = \big(f(1) - f(t)\big) + \big(f(t) - f(0)\big)$. Néanmoins, la méthode façon Taylor-Lagrange semble présenter quelques défauts. Attardons nous tout d'abord sur un cas qui *semble a priori* marcher.

37.5. **Proposition.** *Les normes suivantes sont équivalentes sur $\mathcal{C}^n([0,1], \mathbf{R})$:*

$$N_1 : f \mapsto \|f\|_\infty + \|f'\|_\infty, \quad N_2 : f \mapsto \|f\|_\infty + \|f'\|_\infty + \|f''\|_\infty + \|f'''\|_\infty + \ldots + \left\| f^{(n)} \right\|_\infty.$$

Démonstration. □

38. La diagonalisation des fous

Il a déjà était compilé sous bon nombre de formes différentes des abécédaires, formulaires, annuaire... de résultats de réduction. On se propose ici de le faire de nouveau mais sous une forme peut-être quelque peu différente (des formes existantes déjà différentes !). Et, promis Yohan ! on ira se faire la diagonale des fous dans les années qui arrivent.

39. Ouais des groupes y'en a beaucoup

À quoi bon vivre si l'on n'a pas (au moins) classifié les groupes d'ordre inférieur ou égal à 50 ? !

Sur le tableau suivant, on fait figurer (sans redondance, si un nombre apparaît dans plusieurs des catégories ci-dessous on ne le considère que dans une seule) :

— en **gras** les nombres premiers,

— en **bleu** les nombres de la forme $2p$ avec p premier,

— en **rouge** les nombres de la forme p^2 avec p premier,

— en **vert** les nombres de la forme p^3 avec p premier,

— en **violet** les nombres de la forme pq avec p et q premiers, $q \geq 3$.

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
11	12	**13**	**14**	**15**	16	**17**	18	**19**	20
21	**22**	**23**	24	**25**	**26**	**27**	28	**29**	30
31	32	**33**	**34**	**35**	36	**37**	**38**	**39**	40
41	42	**43**	44	45	**46**	**47**	48	**49**	50

Il reste un nombre conséquents de nombres qui n'entrent pas dans les catégories précédentes. Un groupe dont l'ordre n'appartient à aucune des catégories ci-dessus sera appelé (chez nous) groupe récalcitrant (le groupe d'ordre 1 est évidemment exclus de cette catégorie...).

Afin de se mettre les idées en place, l'OEIS (A000001) nous donne le nombre de groupes d'un ordre n donné. On fait apparaître ce numéro entre parenthèse sur le tableau suivant.

1 (1)	**2** (1)	**3** (1)	**4** (2)	**5** (1)	**6** (2)	**7** (1)	**8** (5)	**9** (2)	**10** (2)
11 (1)	12 (5)	**13** (1)	**14** (2)	**15** (1)	16 (14)	**17** (1)	18 (5)	**19** (1)	20 (5)
21 (2)	**22** (2)	**23** (1)	24 (15)	**25** (2)	**26** (2)	**27** (5)	28 (4)	**29** (1)	30 (4)
31 (1)	32 (51)	**33** (1)	**34** (2)	**35** (1)	36 (14)	**37** (1)	**38** (2)	**39** (2)	40 (14)
41 (1)	42 (6)	**43** (1)	44 (4)	45 (2)	**46** (2)	**47** (1)	48 (52)	**49** (2)	50 (5)

40. Symmetric polynomials

One of our first encounter with symmetric polynomials was through the following problem (Thursday, September 28, 2023) :

40.1. **Défi.** Let $p(x) = x^5 + x$ and $q(x) = x^5 + x^2$. Find all $(w, z) \in \mathbf{C}^2$ such that $w \neq z$ and

$$\begin{cases} p(w) = p(z) \\ q(w) = q(z). \end{cases} \tag{183}$$

Partial proof. It is fairly easy to find a relation between w and z : because $w \neq z$ is supposed, we are really tempted to divide $p(w) - p(z)$ and $q(w) - q(z)$ by $w - z$:

$$P(w, z) = \frac{p(w) - p(z)}{w - z} = \frac{w^5 - z^5 + w - z}{w - z} = \frac{w^5 - z^5}{w - z} + 1 \tag{184}$$

$$Q(w, z) = \frac{q(w) - q(z)}{w - z} = \frac{w^5 - z^5 + w^2 - z^2}{w - z} = \frac{w^5 - z^5}{w - z} + (w + z). \tag{185}$$

Thus $w + z = 1$. Brutal calculations may conclude but there is a smarter and classic way. The sum of the two "roots" (w and z) is known and there is a special situation where the knowledge of the sum and the product of roots determines completely a polynomial : the symmetric polynomials (in two variables, in this case). Hence, knowing the value of wz would allow us to find all the possible values (w, z) such that is it solution of 183, as desired. □

The aim of this paper is not to give a detailed proof of problem 40.1 (see IMC 2000). We will only survey some of the main properties of symmetric polynomials (using Gourdon's *Algebra*).

Let $\mathbf{A}$ be an unitary commutative ring.

40.2. **Définition.** A polynomial $P \in \mathbf{A}[X_1, \ldots, X_n]$ is *symmetric* if :

$$P(X_{\sigma(1)}, X_{\sigma(2)}, \ldots, X_{\sigma(n)}) = P(X_1, X_2, \ldots, X_n) \tag{186}$$

for all $\sigma \in \mathfrak{S}_n$ (σ is a permutation of n elements).

Let's see some examples. We will use SageMath (a free open-source mathematics software system) to check properties and play with symmetric polynomials. First, we have to instantiate the ring on which we're working on. Let's say it is $\mathbf{R}[X, Y, Z]$:

```
R.<X, Y, Z> = PolynomialRing(RR)
```

One of the simplest example (putting aside the null function) is :

```
f = X + Y + Z
f.is_symmetric() # True
```

We can also build more complicated example :

```
f = X^4 + Y^4 + Z^4 + X^2*Y*Z + X*Y^2*Z + X*Y*Z^2
f.is_symmetric() # True
```

We can build incredibly complicated example but it'll only be an illusion because each symmetric polynomial can be expressed as a combination of elementary symmetric polynomials. Understanding this class of symmetric polynomials will be sufficient to deal with general symmetric polynomials.

40.3. **Définition.** The k-th elementary symmetric polynomial $e_k \in \mathbf{A}[X_1, \ldots, X_n]$ is :

$$e_k(X_1, X_2, \ldots, X_n) = \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} X_{i_1} X_{i_2} \cdots X_{i_k}. \tag{187}$$

We will produce the list of the k-th symmetric polynomial for small values of n.

```
R = PolynomialRing(RR, n, 'X')
for i in range(n + 1):
f = e([i])
g = f.expand(n, alphabet=['X_' + str(i) for i in range(1, n + 1)])
print(g)
```

We thus obtain a beautiful Christmas tree (or a space invaders ship) !

n	e_k (with $k \in [\![1, n]\!]$)
1	1 X_1
2	1 $X_1 + X_2$ X_1X_2
3	1 $X_1 + X_2 + X_3$ $X_1X_2 + X_1X_3 + X_2X_3$ $X_1X_2X_3$
4	1 $X_1 + X_2 + X_3 + X_4$ $X_1X_2 + X_1X_3 + X_2X_3 + X_1X_4 + X_2X_4 + X_3X_4$ $X_1X_2X_3 + X_1X_2X_4 + X_1X_3X_4 + X_2X_3X_4$ $X_1X_2X_3X_4$
5	$X_1 + X_2 + X_3 + X_4 + X_5$ $X_1X_2 + X_1X_3 + X_2X_3 + X_1X_4 + X_2X_4 + X_3X_4 + X_1X_5 + X_2X_5 + X_3X_5 + X_4X_5$ $X_1X_2X_3 + X_1X_2X_4 + X_1X_3X_4 + X_2X_3X_4 + X_1X_2X_5 + X_1X_3X_5 + X_2X_3X_5 + X_1X_4X_5 + X_2X_4X_5 + X_3X_4X_5$ $X_1X_2X_3X_4 + X_1X_2X_3X_5 + X_1X_2X_4X_5 + X_1X_3X_4X_5 + X_2X_3X_4X_5$ $X_1X_2X_3X_4X_5$

Our problem is now a *symmetrization problem* : given a symmetric polynomial expressed in the usual basis we want to explicit the same polynomial in the basis of elementary symmetric polynomials. That can be done, for instance, using the "SymmetricReduction[...]" function from WolframAlpha. An implementation has been made for SageMath by Federico Lebrón (the source has been slightly modified to be up-to-date) :

```
def symmetrize(p):
if not p.is_symmetric(): raise Error(str(p) + " is not a symmetric polynomial.")
vars = p.variables()
nvars = len(vars)

S = SymmetricFunctions(RR)
e = S.elementary()
sigmas = [e([i]).expand(nvars, alphabet=vars) for i in range(1, nvars+1)]

R = PolynomialRing(p.base_ring(), [’sigma_%s’ % i for i in range(1, nvars+1)])
sigma_vars = R.gens()

def sym(f):
if f == 0: return 0
c = f.lc()
degrees = f.lm().degrees()

exps = [degrees[i]-degrees[i+1] for i in range(len(degrees)-1)]
exps.extend(degrees[-1:])

g = prod([sigma_vars[i]**exps[i] for i in range(len(exps))])
gp = prod([sigmas[i]**exps[i] for i in range(len(exps))])

return c*g + sym(f - c*gp)
return sym(p)
```

We look at the $n = 2$ case, for convenience we let $\sigma_1 = X + Y$ and $\sigma_2 = XY$.
For instance, suppose $f(X, Y) = X^2 + Y^2 + 6XY - 3X^2Y - 3XY^2$. Then, the symmetrized version of f is $\sigma_1^2 - 3\sigma_1\sigma_2 + 4\sigma_2$.
To gain in intuition, at least on $\mathbf{R}[X, Y]$, one may consider constructing many examples.

$P \in \mathbf{R}[X, Y]$	Symmetrized version	$P \in \mathbf{R}[X, Y]$	Symmetrized version
$(X + Y) - X - Y$	0	$X + Y$	σ_1
$(X + Y)^2 - X^2 - Y^2$	$2\sigma_2$	$X^2 + Y^2$	$\sigma_1^2 - 2\sigma_2$
$(X + Y)^3 - X^3 - Y^3$	$3\sigma_1\sigma_2$	$X^3 + Y^3$	$\sigma_1^3 - 3\sigma_1\sigma_2$
$(X + Y)^4 - X^4 - Y^4$	$4\sigma_1^2\sigma_2 - 2\sigma_2^2$	$X^4 + Y^4$	$\sigma_1^4 - 4\sigma_1^2\sigma_2 + 2\sigma_2^2$
$(X + Y)^5 - X^5 - Y^5$	$5\sigma_1^3\sigma_2 - 5\sigma_1\sigma_2^2$	$X^5 + Y^5$	$\sigma_1^5 - 5\sigma_1^3\sigma_2 + 5\sigma_1\sigma_2^2$
$-X - Y + 8X + 8Y$	$7\sigma_1$	$X^2Y + Y^2X$	$\sigma_1\sigma_2$
$-X^2 - Y^2 + 8X^2Y + 8Y^2X$	$-\sigma_1^2 + 8\sigma_1\sigma_2 + 2\sigma_2$	$X^4Y^2 + Y^4X^2$	$\sigma_1^2\sigma_2^2 - 2\sigma_2^3$
$-X^3 - Y^3 + 8X^3Y^2 + 8Y^3X^2$	$-\sigma_1^3 + 8\sigma_1\sigma_2^2 + 3\sigma_1\sigma_2$	$X^6Y^3 + Y^6X^3$	$\sigma_1^3\sigma_2^3 - 3\sigma_1\sigma_2^4$
$-X^4 - Y^4 + 8X^4Y^3 + 8Y^4X^3$	$-\sigma_1^4 + 8\sigma_1\sigma_2^3 + 4\sigma_1^2\sigma_2 - 2\sigma_2^2$	$X^8Y^4 + Y^8X^4$	$\sigma_1^4\sigma_2^4 - 4\sigma_1^2\sigma_2^5 + 2\sigma_2^6$
$-X^5 - Y^5 + 8X^5Y^4 + 8Y^5X^4$	$-\sigma_1^5 + 8\sigma_1\sigma_2^4 + 5\sigma_1^3\sigma_2 - 5\sigma_1\sigma_2^2$	$X^{10}Y^5 + Y^{10}X^5$	$\sigma_1^5\sigma_2^5 - 5\sigma_1^3\sigma_2^6 + 5\sigma_1\sigma_2^7$

More generally, one can proceed the same way for each symmetric polynomial in n variables.

40.4. **Théorème.** *Every symmetric polynomial of* $\mathbf{A}[X_1, \ldots, X_n]$ *is uniquely determined by a polynomial expression of elementary symmetric polynomials.*

Démonstration. See Algebra, class n°33, MIT or Charles Walter's course. □

Let's see an usual application of the theorem 40.4. Set n to 1 and consider the root-expression of a n-th degree polynomial :

$$P_n(X) = \prod_{i=1}^{n} (X - z_i). \tag{188}$$

If we expand it for small values of n, we find :

$$P_1(X) = X - z_1 \tag{189}$$

$$P_2(X) = X^2 - (z_1 + z_2)X + z_1 z_2 \tag{190}$$

$$P_3(X) = X^3 - (z_1 + z_2 + z_3)X^2 + (z_1 z_2 + z_2 z_3 + z_3 z_1)X - z_1 z_2 z_3. \tag{191}$$

We obviously recognize some of the elementary symmetric polynomials e_k ! More generally, one can prove the following result :

$$P(\sigma_1, \sigma_2, \ldots, \sigma_n) := P_n(X) = X^n - \sigma_1 X^{n-1} + \sigma_2 X^{n-2} - \ldots + (-1)^n \sigma_n. \tag{192}$$

We thus have a *duality result* :

$$\begin{array}{ccc} \text{Symmetric polynomial} & & \text{One-variable polynomial} \\ \text{in terms of roots of } P & \longleftrightarrow & \text{in terms of coeffs of } P \\ P(\sigma_1, \sigma_2, \ldots, \sigma_n) & & P_n(X) \end{array}$$

As an other example of this duality, one can the consider generating function of the elementary symmetric polynomials :

$$g(t) := \sum_{k \geq 0} e_k(X_1, X_2, \ldots, X_n) t^k = \prod_{i=1}^{n} (1 + t z_i). \tag{193}$$

We then have :

$$P_n(t) = t^n g(-1/t). \tag{194}$$

We can also attach some special symmetric polynomial functions to a given, for instance, one-variable polynomial. A canonical example is the discriminant. Just think of the well-known quantity $\Delta = b^2 - 4ac$. But $\Delta(a, b, c)$ is not a symmetric polynomial of $\mathbf{A}[a, b, c]$. Instead of the (a, b, c)-triplet we shall consider the roots of z_1 and z_2 of a two-degree polynomial. Computing $(z_1 - z_2)^2$ gives the desired result $b^2 - 4ac$ (up to a division by a, a minor detail without more significance than a scaling constant). One sees that $(z_1 - z_2)^2$ is a symmetric polynomial !

More generally, one defines the discriminant $\Delta(P)$ of $P(X) = \prod_{i=1}^{n} (X - z_i)$ such that :

$$\Delta(P) = \prod_{i<j} (z_i - z_j)^2. \tag{195}$$

41. Dans la Vallée du Poussin il y a de l'uniforme intégrabilité

42.

43.

44.

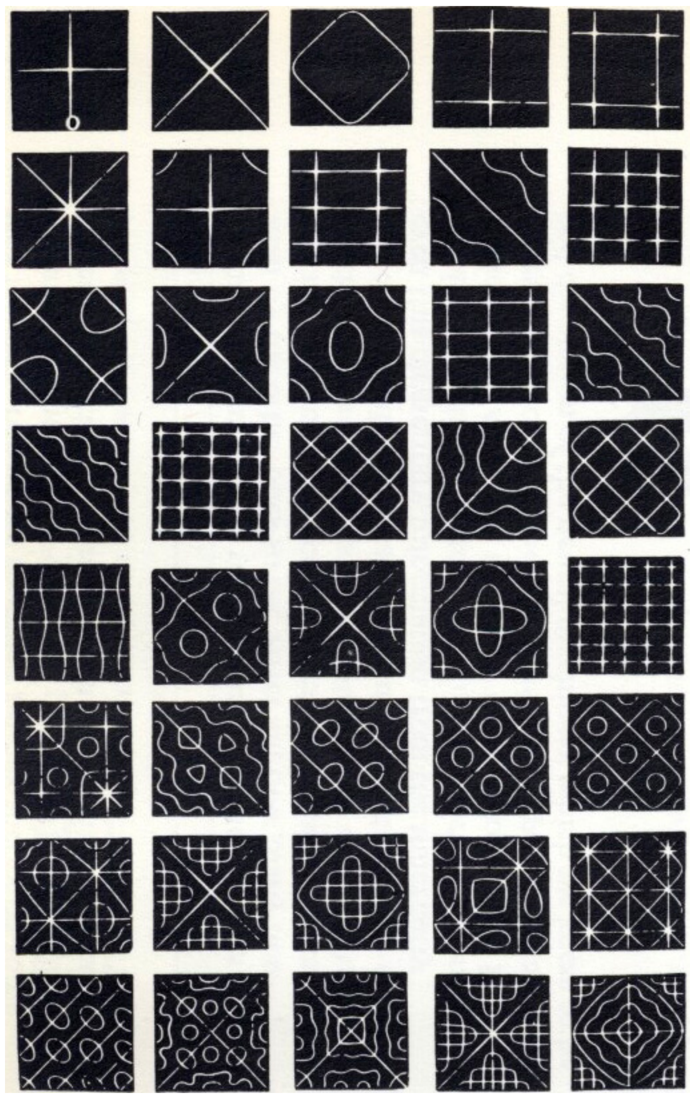
45. Counting zeros of graphs eigenfunctions, a mini-course by G. Berkolaiko

45.1. Lecture 1. Eigenvalue interlacing from Cauchy to Weyl to Dirichlet-Neumann bracketing.

Overview. In this lecture we will introduce the concept of eigenvalue interlacing, starting with symmetric matrices (Cauchy and Weyl interlacing) and proceeding to the Laplacian of a metric graph. The proofs proceed using the well-known tools : quadratic forms and the minimax characterisation of the eigenvalues.

The focus will be put on the discrete Laplacian but the methods involved also extend to differential operator on metric graph (but won't have enough time to treat them).

Before the modern foundation of the theory by Sturm in the 1840s (*oscillation theory of Sturm-Liouville operators*), the subject goes back to observations of high vibrational modes which oscillate on a finer scale thus they have many zeros (need to properly explain what is intended). The first observations were made by da Vinci, Galileo and Hooke and it was Chladni who really experienced the visualisation of zeros of eigenfunctions. See for instance the following pictures from physics exchange :

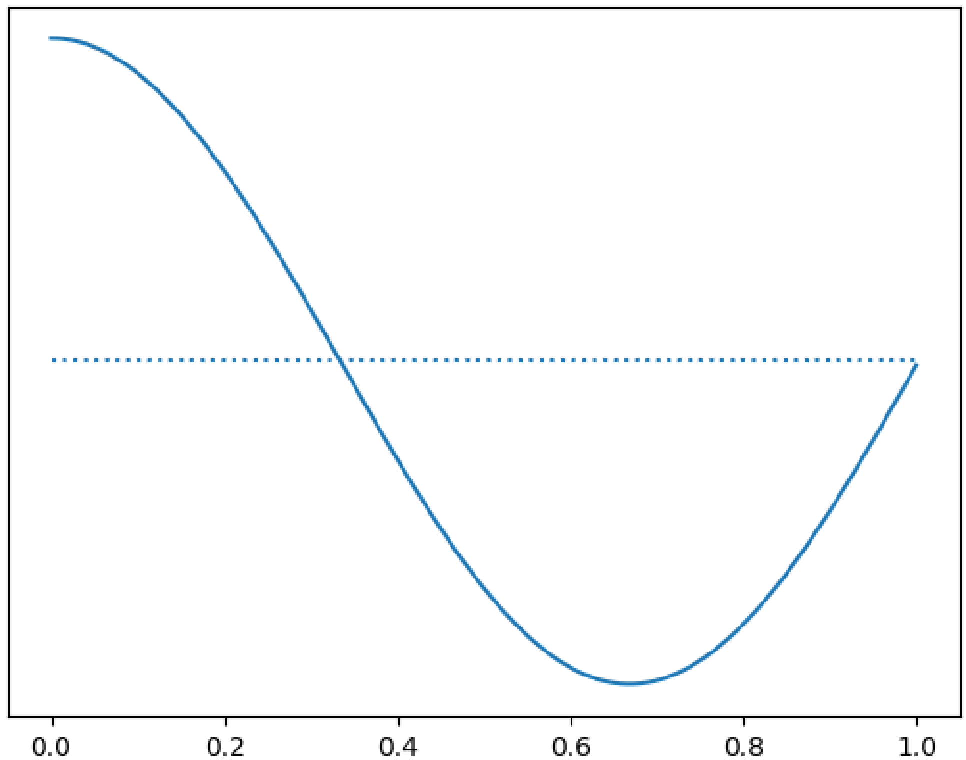


Motivations. One of the main idea inherited from the modern theory is, broadly speaking, that on an interval, the k-th eigenfunction has $k-1$ zeros.

For instance, in the one dimensional case, consider the interval $[0,1]$ with boundary conditions $f'(0)=0$ and $f(1)=0$. On the interval, acts the following differential operator : $-\dfrac{\mathrm{d}^2}{\mathrm{d}x^2}$, i.e. the one dimensional Laplace operator. One would like to solve to generalised eigenvalue equation :

$$-\frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2} = \lambda f(x) \tag{196}$$

with the imposed boundary conditions. Here is the plot of the shape of the second eigenvector $(\lambda = 2)$:



One effectively sees only one zero (the boundary condition do not count).

In the two dimensional case, a similar problem is posed still for the Dirichlet condition. The main result can be traced back to Courant in 1923 : the number of nodal domains (seen as connected components) of the k-th eigenfunction is bounded above by k. For Neumann boundary condition, a similar result exists and it is quite recent. Many questions arise from this setting.

Some questions. Denote by $\nu(f_k)$ the number of nodal domains of the k-th eigenfunction. Can one estimate the following quantity :

$$\limsup_{k\to+\infty} \frac{\nu(f_k)}{k} = \ ? \tag{197}$$

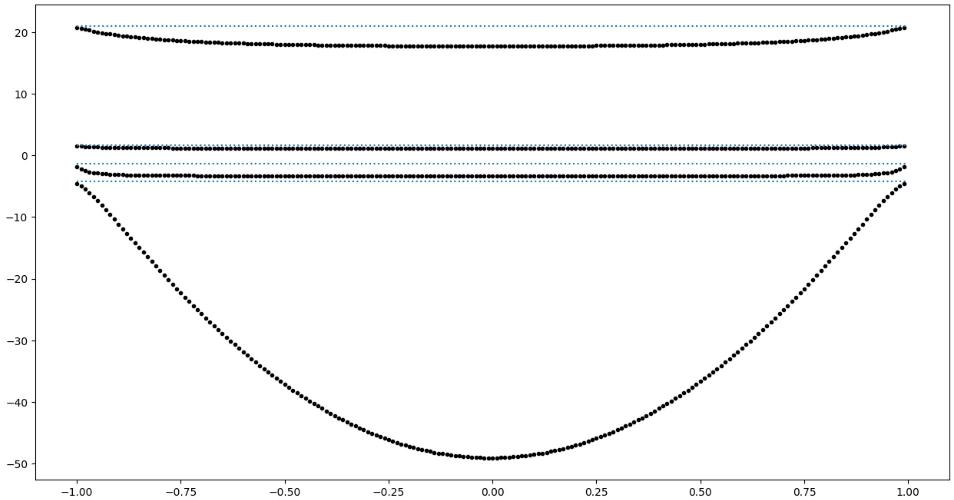
A bound was given by Pleijel in the fifties in terms of zero of the Bessel function. We thus have $\limsup_{k\to+\infty} \frac{\nu(f_k)}{k} \lessapprox 0.7$. (This seems to be related to Faber-Krahn inequality and Weyl asymptotic.) More directly, one can ask for the superior limit of $\nu(f_k)$. It was investigated by Hoffmann-Ostenhof and solved by Jung-Zelditch. An other problem relies on giving a distribution of $\frac{\nu(f_k)}{k}$. Numerics shows that the distribution looks different in function of the shape of the studied domain. There are works of Smilansky and a famous conjecture by Bogonolny with progress by Nazarov-Sodin. At last, one can expect *universality* of nodal surplus distribution on graphs (see Alon, Band, Berkolaiko and Goresky).

Some problems treat the measure of the nodal set. See Yau, Donnely-Feffermann and recent progresses by Logunov-Malinnikova. Things can get really wild (for instance in three dimensions).

Jerison and his collaborators studied the morphology of the second nodal curve.

A domain where Helffer contributed a lot with his collaborators relates to spectral minimal partitions.

Eigenvalue interlacing. Given a matrix M, form the symmetric matrix $H = \frac{M + {}^tM}{2}$. Consider a perturbation of H (for instance, we replace $H_{2,2}$ by $H_{2,2} + s$, where s is the perturbation parameter) denoted by $H(s)$. We want to see how eigenvalues evolve in terms of s and compare it with the non perturbed matrix H. Here is an example of a plot for a random matrix M.



One can guess that $\lambda_k(s)$, the k-th eigenvalue of $H(s)$, is increasing if $s \geq 0$ and $\lambda_k(0) \leq \lambda_k(s) \leq \lambda_{k+1}(0)$ for $s \geq 0$. But there is nothing special with zero, we must have more generally $\lambda_k(s') \leq \lambda_k(s) \leq \lambda_{k+1}(s')$ for $s \geq s'$.

Since the sequence $(\lambda_k(s))_k$ is bounded and increasing, it must have a limit. Is it $\lambda_{k+1}(0)$? No. May it be $\lim\limits_{s' \to -\infty} \lambda_{k+1}(s)$? To investigate the behaviour at $\pm\infty$, we can make the change of variable $s = \tan(t)$.

From numeric we can guess the following Weyl interlacing theorem.

45.1. **Théorème.** *Let H be a real symmetric matrix, P a rank one perturbation (equal to $(\vec{z}, \cdot)\vec{z}$, where the dot product is linear in the second variable). We thus have :*

$$\lambda_k H \leq \lambda_k(H + P) \leq \lambda_{k+1}(H). \tag{198}$$

Here is the proof idea : use minimax characterisation of the eigenvalues :

$$\lambda_k(A) = \min_{\substack{V \subset \mathcal{C}^n \\ \dim(V)=k}} \max_{\substack{||u||=1 \\ u \in V}} Q_A(\vec{u}) = \min_{\substack{V \subset \mathcal{C}^n \\ \dim(V)=k}} \max_{\substack{||u||=1 \\ u \in V}} (\vec{u}, A\vec{u}). \tag{199}$$

Apply to $H + P$, then find a one codimensional subspace ignoring P.

Here is an other interlacing inequality : Cauchy interlacing.

45.2. **Théorème.** *Suppose H' is a principal minor of H. Then :*

$$\lambda_k(H) \leq \lambda_k(H') \leq \lambda_{k+1}(H). \tag{200}$$

The proof idea is the same : use minimax and observe that $Q_{H'} = Q_H$ restricted to subspace of codimension 1.

We introduce some notations. We say $V \subset_1 W$ if W/V is one dimensional. Let $\mathcal{H}$ be a not necessarily finite dimensional Hilbert space, $Q : \mathrm{dom}(Q) \times \mathrm{dom}(Q) \to \mathbf{C}$ a sesquilinear form bounded from below, closed and hermitian (i.e. $Q(u, v) = \overline{Q(v, u)}$) with discrete spectrum. We define f as an eigenvector with eigenvalue λ if $Q(u, f) = \lambda(u, f)_{\mathcal{H}}$, for all $u \in \mathrm{dom}(Q)$.

Note that there are other characterisations of the minimax principle (maximin !) :

$$\begin{aligned}\lambda_k(A) &= \min_{\substack{V \subset \mathrm{dom}(Q) \\ \dim(V)=k}} \max_{\substack{||u||=1 \\ u \in V}} Q_A(\vec{u}, \vec{u}) \\ &= \max_{\substack{W \subset \mathrm{dom}(Q) \\ \dim(W)=k-1}} \min_{\substack{u \in W^\perp \\ ||u||=1}} Q_A(\vec{u}, \vec{u}).\end{aligned}$$

45.3. **Définition.** (see Birman-Solomyak book) The operator $\tilde{Q}$ is a rank one non negative perturbation of Q if :

— $\exists Z \subset \mathrm{dom}(Q)$ such that $Q\big|_Z = \tilde{Q}\big|_Z$ and

— either $\mathrm{dom}(\tilde{Q}) \subset_1 \mathrm{dom}(Q)$ (and $Z = \mathrm{dom}(\tilde{Q})$) or $\mathrm{dom}(\tilde{Q}) = \mathrm{dom}(Q)$ and $\tilde{Q} \geq Q$ (for every vector).

Here is a general interlacing inequality.

45.4. **Théorème.** *Let $\tilde{Q}$ be a non negative rank one perturbation of Q. Then :*

$$\lambda_k(Q) \leq \lambda_k(\tilde{Q})\lambda_{k+1}(Q). \tag{201}$$

Furthermore, if Λ has multiplicity m and $\tilde{m}$ is in the spectrum (resp.) of Q, $\tilde{Q}$, then :

$$|m - \tilde{m}| \leq 1 \text{ and } \dim(E(Q,\lambda) \cap E(\tilde{Q},\lambda)) = \min(m, \tilde{m}) \tag{202}$$

where $E(Q,\lambda)$ is the eigenspace of Q for the eigenvalue λ.

For more informations, see *Surgery principles for the spectral analysis of quantum graphs* by Berkolaiko-Kennedy-Kurasov-Mugnolo, TAMS 2019. (Due to the title of the article, Berkolaiko got surprisingly invited to many medical conferences !)

45.2. **Lecture 2. Interlacing inequalities and nodal estimates.**

> **Overview.** In this lecture we will use interlacing inequalities to derive the estimates on the number of zeros of the n-th eigenfunction on a graph (both discrete and metric). That the number of zeros of an eigenfunction is related to the label of the corresponding eigenvalue is an old idea of Sturm, developed in higher dimensions by Courant, Pleijel and others. We will see that metric graphs behave as if they have a little more than one dimension to them, quantified by the first Betti number of the graph.

Our objective is to get non trivial estimates of zeros of eigenfunctions on grpahs $\Gamma = (V, E)$ where $V = \{1, \ldots, n\}$ is the set of vertices and E the set of edges. (The graph is considered by no loops and even connected).

The graph Laplacian is the following operator (there exists many alternatives definitions in terms of what we expect about it) : $L : \mathbf{C}^V \to \mathbf{C}^V$ such that

$$(L\psi)_u = \sum_{v \sim u} (\psi_u - \psi_v). \tag{203}$$

A (generalised discrete) Schrödinger operator $H : \mathbf{C}^V \to \mathbf{C}^V$ is real symmetric and respect for all $u \neq v$, $H_{u,v} \neq 0 \iff u \sim v$. Pay attention, $H_{u,u}$ can be zero.

Let $\psi \in \mathbf{C}^V$, H be as above. We define the edge count function :

$$\phi_H(\psi) := \#\big\{u \sim v | \psi_u\big(-H_{u,v}\big)\psi_v < 0\big\}. \tag{204}$$

It counts the number of edges on which the function changes sign. It is quite analogous to the number of zeros.

45.5. **Théorème.** *Let H be a generalised Schrödinger operator, Γ a connected graphs. Let $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \ldots$ with corresponding eigenfunctions. If λ_k is simple and $(\psi_k)_v \neq 0$ for all $v \in V$. Then :*

$$k - 1 \leq \phi_H(\psi_k) \leq k - 1 + \beta_1 \tag{205}$$

with $\beta_1 = |E| - |V| + 1 = \#$cycles the first Betti number (generalisations for Betti numbers seem not to exist).

For $\beta = 0$ it is known since Fiedler, 1975. There was in independent proof by Biyikoglu in 2003. For $\beta > 0$, see Berkolaiko, 2008. One can drop some assumptions, see H. Xu, S.-T. Yau, 2012 and two articles in 2023 by C. Ge, S. Lin. See also Deidda-Putti-Tudisco-Hein for p-Laplacians.

Define the nodal surplus $\sigma(H,k) := \phi_H(\psi_k) - (k-1)$. Here is a really beautiful inverse result.

45.6. **Théorème.** (Band, 2014)

$$\sigma(H,k) = 0, \text{ for all } k, \text{ then } H \sim \text{a tree graph.} \tag{206}$$

Here is a converse result based on the two following lemmas.

45.7. **Lemme.** *Assume Γ is a tree graph, ψ an eigenvector of eigenvalue λ of H with $\psi_u \neq 0$ for all $u \in V$. Then λ is simple, i.e. of multiplicity one.*

Démonstration. By induction on $n = |V|$. The $n = 1$ case is obvious because there is only one eigenvalue. For the induction : pick u such that $\deg(u) = 1$ (it is always possible because we consider finite graphs ; choosing a leaf is maybe not absolutely mandatory for the sake of the proof). Consider the perturbation : $\tilde{H}_{u,v} = 0$, $\tilde{H}_{v,u} = 0$, $\tilde{H}_{u,u} = H_{u,u} + \frac{\psi_v}{\psi_u} H_{u,v}$, $\tilde{H}_{v,v} = H_{v,v} + \frac{\psi_u}{\psi_v} H_{v,u}$. It has been built such that $\tilde{H}\psi$ is still equal to $\lambda\psi$ and $\tilde{H}$ is a rank one perturbation of H (could be positive or negative but at this point we don't care).

Since $\lambda \in \mathrm{spec}(\tilde{H})$ is of multiplicity two (no more than two by induction hypothesis ; not understood this part). Assume λ is not simple for H. Then, by interlacing theorem (the dimensional equality) :

$$E(\tilde{H}, \lambda) \subset E(H, \lambda). \tag{207}$$

Note that $E(\tilde{H}, \lambda)$ contains a "strange" eigenvector :

$$\tilde{\psi} = \begin{cases} 1 & \text{on } u \\ 0 & \text{otherwise} \end{cases} \tag{208}$$

But, we have :

$$0 = (\lambda\tilde{\psi})_v = (H\tilde{\psi})_v = \sum_{w \sim v} H_{v,w}\psi_w = H_{v,u}. \tag{209}$$

This is a contradiction since $H_{v,u}$ cannot be 0. So λ is simple. □

45.8. **Lemme.** *Assume Γ is a tree ($\beta = 0$), ψ an eigenvector with $\psi_u \neq 0$ for all $u \in V$. Then :*

$$\phi_H(\psi_k) = k - 1 = N(H, \lambda) \tag{210}$$

where

$$N(H, \lambda) := \#\{\lambda' \in \mathrm{spec}(H) \mid \lambda' < \lambda\}. \tag{211}$$

Démonstration. We also proceed by induction on $n = |V|$. The proof is omitted since the author did not understood it. □

45.3. Lecture 3. Transversal variation and magnetic characterisation of the nodal count.

Overview. While eigenvalue interlacing deals with a single rank-r perturbation to the graph operator, more precise information is available by going sideways : allowing the rank-r perturbation to vary in a small neighborhood and viewing the eigenvalue as a real-valued function on the space of rank-r perturbations. This idea allows us to compute the number of zeros of the graph eigenfunction as a Morse index of the corresponding eigenvalue with respect to magnetic perturbations of the operator, opening up the problem to tools from Morse theory and, more widely, algebraic topology.

Some new notations :

- H is the Laplacian associated with the graph $\Gamma = (V, E)$,
- $S(\Gamma; \mathbf{R})$ is the set of real symmetric H supported on Γ,
- $S(H)$ is the set of all "signings" of H :

$$S(H) = \{H' \in S(\Gamma; \mathbf{R}) \mid |H'_{u,v}| = |H_{u,v}|, H'_{u,u} = H_{u,u}\}. \tag{212}$$

45.9. **Conjecture.** (Berkolaiko and collaborators, see Alon-Goresky, 2022) *Let $H \in S(\Gamma; \mathbf{R})$ be "generic" (technical meaning, not important here). Consider $\sigma(H', k)$ as a random variable with k uniform in $\{1, \ldots, n\}$, $H' \in S(H)$ uniform.*

Then $\sigma(H', k)$, standardised (subtract mean and divide by variance), converges to $\mathcal{N}(0, 1)$ as $\beta \to +\infty$ uniformly on Γ (pick a sequence of graph with $\beta \to +\infty$, uniformly means the choice of graphs is not taken into consideration).

For Berkolaiko it sounds a little bit too good to be true, but :

45.10. **Théorème.** (Alon-Goresky, 2024) *Let Γ have disjoint cycles, $H \in S(\Gamma; \mathbf{R})$ "generic". For any fixed k (the result is the same of uniformly distributed k), $H' \in S(H)$, then the distribution is* $\mathrm{Bin}(1/2, \{0, \ldots, \beta\})$.

Note that if $\beta \to +\infty$, it converges to the gaussian !

The conjecture was inspired by work of Alon-Band-Berkolaiko on quantum graphs. The tools used to prove the theorem are eigenvalue perturbations and Morse theory (counts number of critical points of a given index on a compact manifold).

What is a transversal variation ? We want to compare $\mathrm{spec}(A + L(0))$ with $\mathrm{spec}(A)$, where A is a selfadjoint matrix and $L(0)$ is a perturbation maybe depending on parameters, for instance : $\mathrm{spec}(A + tL)$. We want to vary t and track the eigenvalues !

45.11. **Définition.** Let $\mathcal{U} \subset \mathbf{R}^d$ a neighbourhood of 0, let $L(x)$ with $x \in \mathcal{U}$ a family of operators on $\mathcal{H}$, let $g_0 \in \ker(L_0)$. Then $L(x)$ is a transversal variation of $L(0)$ with respect to g_0 if :

— the range of $L(x)$ is smooth in x,

— $\chi : \mathcal{U} \to \mathcal{H}$, $x \mapsto L(x)g_0$ is transversal to $\ker(L(0))^* = (\mathrm{range} L(0))^{\perp}$.

45.12. **Théorème.** (Berkolaiko, Kuchment ; work in progress Berkolaiko, Latushkin, Sukhtaiev) *Let A be a selfadjoint operator on $\mathcal{H}$, $Ag_0 = \lambda_0 g_0$ and assume λ_0 is simple. The family of operators $L(x)$ is a transversal variation of $L(0)$ with respect to $g_0 \in \ker L_0$. Assume $\lambda_0 \in \mathrm{spec}(A + L(0))$ is simple, let $\Lambda(x) \in \mathrm{spec}(A + L(x))$ (thus $\Lambda(0) = \lambda_0$). Then : $d\Lambda(0) = 0$ and*

$$n_-(d^2\Lambda(0)) = \underbrace{n_-(A - \lambda_0) - n_-(A + L(0) - \lambda_0)}_{\text{the spectral shift}} + n_-(L(0)) \tag{213}$$

where $n_-(\cdot)$ counts the number of negatives eigenvalues, d^2 is the hessian. We are thus computing the Morse index of Λ at the critical point $x = 0$.

The hypothesis $Ag_0 = \lambda_0 g_0$ can be dropped and it would be natural to drop the hypothesis of simplicity of λ. Note that the key ingredient is operators sharing the same eigenvector.

Fun fact, and that's why Berkolaiko finds this formula beautiful, if we open the brackets, then everything disappears and simplifies.

Here is a former result who motivated a little this work. Berkolaiko sent a mail to Y. Colin de Verdière and no more than two days later, he replied with a more beautiful proof (but Berkolaiko preferred his proof) using de Rham cohomology.

45.13. **Théorème.** (Berkolaiko, Colin de Verdière, 2013) *Let λ_k be a simple eigenvalue, $\psi_k \neq 0$ on vertices. Consider $\Lambda(x) := \lambda_k(H(\vec{x}))$, where $H(\vec{x})$ is the discrete magnetic Schrödinger operator. Then $\Lambda(x)$ has a critical point at $\vec{x} = \vec{0}$ with Morse index $\sigma(H, k)$.*

For Berkolaiko, there is magic in constant rank operators !

46. Solution of the wave equation

This subsection will be very classical but is quite new for the author and he needs to treat it carefully at least one time. Here is our problem : let $G = (V, E)$ a locally finite weighted graph and Ω a bounded domain. We want to solve the inhomogeneous wave equation on the graph, i.e.

$$\begin{cases} \partial_t^2 u(t,x) - \Delta u(t,x) = f(t,x), & (t,x) \in [0, +\infty[\times \Omega^{\mathrm{o}} \\ u(0,x) = g(x), \partial_t u(0,x) = h(x), & x \in \Omega^{\mathrm{o}} \\ u(t,x) = 0, & (t,x) \in [0, +\infty[\times \partial\Omega \end{cases} \tag{214}$$

where the domain of the Laplacian is restricted to Ω^{o}, the function $f : [0, +\infty[\times \Omega^{\mathrm{o}} \to \mathbf{R}$ is given and continuous in t, and g and h are unknown function from Ω^{o} to $\mathbf{R}$.

We treat the homogeneous setting (i.e. $f = 0$). We make the assumption that u can be decomposed in a basis of eigenvectors ϕ so that

$$u(t,x) = \sum_n c_n(t)\phi_n(x). \tag{215}$$

The homogeneous equation can be rewritten :

$$0 = \partial_t^2\Big(\sum_n c_n(t)\phi_n(x)\Big) - \sum_n c_n(t)\Delta\phi_n(x) = \sum_n c_n''(t)\phi_n(x) - \sum_n c_n(t)\lambda_n\phi_n(x). \tag{216}$$

Since $\{\phi_n\}_n$ is a basis of $\ell^2(V, \mu)$ (we are not perfectly rigorous on this point), we must have the following implication :

$$\sum_n \big(c_n''(t) - c_n(t)\lambda_n\big)\phi_n(x) = 0 \Longrightarrow c_n''(t) - c_n(t)\lambda_n = 0. \tag{217}$$

Hence :

$$c_n(t) = C_{1,n}\mathrm{e}^{\sqrt{\lambda_n}t} + C_{2,n}\mathrm{e}^{-\sqrt{\lambda_n}t}. \tag{218}$$

To determine the constant (for t) $C_{1,n}$ and $C_{2,n}$ we use the initial value conditions. Expand the initial value in the orthogonal basis $\{\phi_n\}_n$:

$$\begin{aligned} u(0,x) = g(x) &= \sum_n g_n \phi_n(x), \\ \partial_t u(0,x) = h(x) &= \sum_n h_n \phi_n(x). \end{aligned}$$

By uniqueness of the decomposition, we have : $C_{1,n} + C_{2,n} = g_n$. The same way, we find that $h_n = \sqrt{\lambda_n}\big(C_{1,n} - C_{2,n}\big)$ because :

$$\partial_t u(t,x) = \sum_n \left(C_{1,n}\sqrt{\lambda_n}\mathrm{e}^{\sqrt{\lambda_n}t} - C_{2,n}\sqrt{\lambda_n}\mathrm{e}^{-\sqrt{\lambda_n}t} \right)\phi_n(x) \tag{219}$$

and evaluating at $t = 0$ yields the result.

Solving the system hereinbelow, we find an expression for $C_{1,n}$ and $C_{2,n}$:

$$\begin{cases} C_{1,n} + C_{2,n} = g_n \\ C_{1,n} - C_{2,n} = \dfrac{h_n}{\sqrt{\lambda_n}} \end{cases} \tag{220}$$

for $\lambda_n \neq 0$.

Injecting this altogether, we find the following homogeneous solution :

$$\begin{aligned} u(t,x) &= \sum_n \left(C_{1,n}\mathrm{e}^{\sqrt{\lambda_n}t} - C_{2,n}\mathrm{e}^{-\sqrt{\lambda_n}t} \right)\phi_n(x) \\ &= \sum_n \left(\Big(\frac{g_n}{2} + \frac{h_n}{2\sqrt{\lambda_n}}\Big)\mathrm{e}^{\sqrt{\lambda_n}t} + \Big(\frac{g_n}{2} - \frac{h_n}{2\sqrt{\lambda_n}}\Big)\mathrm{e}^{-\sqrt{\lambda_n}t} \right)\phi_n(x) \\ &= \sum_n \frac{g_n}{2}\mathrm{e}^{\sqrt{\lambda_n}t}\phi_n(x) + \sum_n \frac{h_n}{2\sqrt{\lambda_n}}\mathrm{e}^{\sqrt{\lambda_n}t}\phi_n(x) + \sum_n \frac{g_n}{2}\mathrm{e}^{-\sqrt{\lambda_n}t}\phi_n(x) - \sum_n \frac{h_n}{2\sqrt{\lambda_n}}\mathrm{e}^{-\sqrt{\lambda_n}t}\phi_n(x) \\ &= \sum_n g_n \frac{\mathrm{e}^{\sqrt{\lambda_n}t} + \mathrm{e}^{-\sqrt{\lambda_n}t}}{2}\phi_n(x) + \sum_n \frac{h_n}{\sqrt{\lambda_n}}\frac{\mathrm{e}^{\sqrt{\lambda_n}t} - \mathrm{e}^{-\sqrt{\lambda_n}t}}{2}\phi_n(x) \\ &= \sum_n g_n \cosh\big(\sqrt{\lambda_n}t\big)\phi_n(x) + \sum_n \frac{h_n}{\sqrt{\lambda_n}}\sinh\big(\sqrt{\lambda_n}t\big)\phi_n(x) \\ &= \sum_n \left(g_n \cosh\big(\sqrt{\lambda_n}t\big) + \frac{h_n}{\sqrt{\lambda_n}}\sinh\big(\sqrt{\lambda_n}t\big) \right)\phi_n(x) \end{aligned}$$

We could be surprised to not find the same result as theorem 1.4 of *The existence of the solution of the wave equation on graphs* by Lin and Xie. It is just a matter of convention (physics or not, that is the question) : in equation 216, we assumed that $\Delta\phi_n(x) = \lambda_n\phi_n(x)$. But we could have considered the eigenvalues and eigenvectors of $-\Delta$. That way, we would have found :

$$u(t,x) = \sum_n \left(g_n \cos\big(\sqrt{\lambda_n}t\big) + \frac{h_n}{\sqrt{\lambda_n}}\sin\big(\sqrt{\lambda_n}t\big) \right)\phi_n(x) \tag{221}$$

Some facts were not completely clear, not sufficiently precise... For instance, what is the range of the index n in every sum ? What precisely are the spaces considered ? This are important questions that must be treated.

One can remark that nothing was really particular to the graph setting (despite the eluded question above). Most of the question lies in determining the eigenvectors ϕ_n ! That's precisely the problem in general !

47.

48.

49.

50.

51. Mesures accélérées pour physiciens vigoureux

Ma (la ?) Bible, c'est le Briane-Pagès. On suit les chapitres 4 à 13 et en fait une exposition sommaire et accélérée pour le camarade physicien Raphaël Decan de Chatouville. On ne trouvera donc pas plus que des idées de preuve qui, au demeurant, sont très naturelles mais peuvent parfois s'avérer astucieuses. N'est présent que le minimum vital. Moralement, on va passer de ces horribles sommes de Riemann relatives à une subdivision $S(f,\sigma,\chi_\sigma) := \sum_{i=1}^n (a_i - a_{i-1}) f(\chi_\sigma)$ à une bien belle théorie qui s'assume parfaitement.

On rappelle simplement qu'un ensemble X est dit dénombrable s'il existe une injection de X dans $\mathbf{N}$, i.e. $\operatorname{card}(X) \leq \operatorname{card}(\mathbf{N})$. Exercice : montrer que $\mathbf{N}^2$ est dénombrable. Indication : penser à un escargot se déplaçant sur $\mathbf{N}\times\mathbf{N}$.

Il peut également être bon d'avoir en tête qu'un espace métrique est séparable si et seulement si il est à base dénombrable d'ouverts. C'est le cas de notre cher L^2. On y reviendra pas beaucoup mais bien avoir en tête que l'espace L^2 est vraiment très agréable (notamment pour faire de la mécanique quantique), c'est un joli Hilbert.

51.1. **Définition.** Soit X un ensemble. On appelle tribu (ou σ-algèbre) sur X toute famille $\mathscr{A}$ de parties de X vérifiant :

(i) $\emptyset \in \mathscr{A}$,

(ii) si $A \in \mathscr{A}$, alors $A^c \in \mathscr{A}$ [stabilité par complémentaire],

(iii) si $(A_n)_n$ est une suite d'éléments de $\mathscr{A}$, alors $\bigcup_{n\geq 1} A_n \in \mathscr{A}$ [stabilité par union *dénombrable*].

Le couple $(X,\mathscr{A})$ est appelé **espace mesurable**. On construira plus tard les espaces mesurés.

E.g. tribu grossière : $\mathscr{A} = \{\emptyset, X\}$; tribu triviale : $\mathscr{A} = \mathscr{P}(X)$; pour $A\subset X$ fixé, $\mathscr{A} = \{\emptyset, A, A^c, X\}$. Une intersection de tribus est une tribu. La plus petite tribu au sens de l'inclusion contenant $\mathscr{E}$ est appelée **tribu engendrée** par $\mathscr{E}$, on la note $\sigma(\mathscr{E})$. Par exemple, la tribu engendrée par les singletons de X est $\mathscr{A} = \{A \in \mathscr{P}(X) : A \text{ est dénombrable ou } A^c \text{ est dénombrable}\}$. Une tribu particulièrement importante (en particulier en probabilités) est la **tribu image**. Soit $f : X \to Y$ et $\mathscr{A}$ une tribu sur X. On appelle tribu image de $\mathscr{A}$ par f la tribu sur Y définie par $\mathscr{B} := \{B\in\mathscr{P}(Y) : f^{-1}(B)\in\mathscr{A}\}$. Attention, la terminologie "tribu *image*" est trompeuse. On remarquera que $\{f(A), A\in\mathscr{A}\}$ n'est pas une tribu sur Y en général.

51.2. **Définition.** La tribu engendrée la plus importante est sans doute la tribu borélienne. Soit $(X,\mathscr{O}(X))$ un espace topologique [14]. La tribu engendrée par les ouverts de X est appelée **tribu borélienne** de X. On la note $\mathscr{B}(X) := \sigma(\mathscr{O}(X))$.

Lorsque l'on parlera de tribu borélienne, on commettra l'abus fréquent qu'est celui de considérer $X = \mathbf{R}$. (Remarque : il n'est pas évidentissime d'exhiber un ensemble non borélien, i.e. qui n'est pas un élément de la tribu borélienne.) Pourquoi est-ce utile de se coltiner des tribus engendrées ? Si le monde est bien fait, on s'attend à ce que bien comprendre un "comportement" sur les éléments qui engendrent la tribu permette de comprendre totalement ce même "comportement" sur toute la tribu. Le monde est heureusement bien fait (détail mathématique, cf le **lemme de transport** : si $f : X\to Y$ et $\mathscr{E}\subset\mathscr{P}(Y)$, alors $\sigma(f^{-1}(\mathscr{E})) = f^{-1}(\sigma(\mathscr{E}))$). Dans le cas précis de $X=\mathbf{R}$, on comprend aisément qu'il suffit de saisir ce qu'il se passe sur les intervalles (car $\mathbf{R}$ est à base dénombrable d'ouverts). Remarquer que l'on a donc (en utilisant la densité de $\mathbf{Q}$ dans $\mathbf{R}$) :

$$\mathscr{B}(\mathbf{R}) = \sigma(\{[a,+\infty[, a\in\mathbf{Q}\}) = \sigma(\{]a,+\infty[, a\in\mathbf{Q}\}) = \sigma(\{]-\infty,a[\in\mathbf{Q}\}) = \sigma(\{]-\infty,a]\in\mathbf{Q}\}). \tag{222}$$

Au fond, ce que l'on veut c'est intégrer des choses. Autrement dit, les mesurer. L'intégrale de Riemann est légèrement embêtante car elle ne permet pas de mesurer des objets auxquels naturellement on saurait donner une valeur (e.g. l'indicatrice de $\mathbf{Q}$ est nulle *presque partout* donc...). On veut même un peu plus ! On veut élargir la classe des fonctions intégrables [15] tout en espérant obtenir d'agréables théorèmes de convergence !

51.3. **Définition.** Soient $(X,\mathscr{A})$ et $(Y,\mathscr{B})$ deux espaces mesurables. Une fonction $f : (X,\mathscr{A}) \to (Y,\mathscr{B})$ est **mesurable** si pour tout $B\in\mathscr{B}$, on a $f^{-1}(B)\in\mathscr{A}$. Si on considère des tribus boréliennes sur X et Y, on parlera plutôt de **fonctions boréliennes**.

---

14. L'ensemble $\mathscr{O}(X)$ correspond évidemment aux ouverts de X.

15. Faut pas croire, mais on perd du monde en chemin. Par exemple $x\mapsto \sin(x)/x$ est Riemann-intégrable mais pas Lebesgue-intégrable sur $[0,1]$.

Une fonction constante est évidemment mesurable, de même l'indicatrice d'un ensemble $A \in \mathscr{A}$ est mesurable. Un petit jeu sur les ouverts de $\mathbf{R}$ et leur tribu engendrée nous donne que toute fonction continue est borélienne. On a les propriétés classiques : la composition, somme, produit de fonctions mesurables est mesurable [16], le min, le max, la valeur absolue de fonctions mesurables est mesurable. Que demande donc le peuple ? Et bien il aimerait bien que tout se passe bien en dimension supérieure (par exemple quand il joue avec des lois de couples en probabilités). C'est effectivement le cas : soit $f := (f_1, f_2) : (X, \mathscr{A}) \to (\mathbf{R}^2, \mathscr{B}(\mathbf{R}^2))$. Alors f est mesurable si et seulement si f_1 et f_2 le sont.

Spoilons un petit peu. On dispose de fonctions mesurables mais pour l'instant on sait pas vraiment comment les manier (en revanche, on a vu qu'elles avaient tout un tas de propriétés habituelles). L'idée est de se ramener à des suites d'éléments d'une sous-classe de fonctions mesurables qui, modulo quelques hypothèses, tend (en croissant, voire même uniformément) vers une fonction mesurable. C'est surtout le mode de convergence qui nous intéresse (et qui fait la force de la théorie de l'intégration de Lebesgue). La construction est alors très naturelle. On part de **fonctions mesurables étagées** (i.e. qui ne prennent qu'un nombre fini de valeurs) et... c'est tout.

51.4. **Lemme.** (fondamental d'approximation) *Soit $f : (X, \mathscr{A}) \to \mathbf{R}$ mesurable. Il existe alors une suite $(f_n)_{n\in\mathbf{N}}$ de fonctions étagées telle que $\lim_n f_n(x) = f(x)$ pour tout $x \in X$. Si de plus $f \geq 0$, on peut choisir la suite croissante et positive (i.e. $\forall n \geq 1, 0 \leq f_n \leq f_{n+1}$). Si f est bornée, on peut choisir la suite de façon à avoir une convergence uniforme (i.e. $\lim_n \sup_{x\in X} |f_n(x) - f(x)| = 0$).*

Dans les faits, les fonctions étagées sont de la forme suivante pour I **fini** : $f = \sum_{i\in I} \alpha_i 1_{A_i}$. Remarquons que la traditionnelle image du découpage régulier selon l'axe des abscisses chez Riemann est ici remplacé par un découpage régulier selon l'axe des ordonnées chez Lebesgue. La puissance de ce lemme est qu'il permet de ramener un problème à des considérations sur des indicatrices. En effet, supposons que l'on veuille prouver une propriété sur une classe de fonctions mesurables alors : on le prouve d'abord pour une indicatrice, puis une somme finie d'indicatrice (i.e. une fonction étagée) et ensuite on conclut par le lemme fondamental d'approximation. C'est par exemple comme cela que l'on montre la propriété d'additivité de l'intégrale. Bref, continuons notre chemin.

51.5. **Définition.** Soit $(X, \mathscr{A})$ un espace mesurable. On appelle **mesure** sur $(X, \mathscr{A})$ toute application $\mu : \mathscr{A} \to \mathbf{R}_+ \cup \{+\infty\}$ vérifiant :

(i) $\mu(\emptyset) = 0$,

(ii) si $(A_n)_n$ est une suite d'éléments de $\mathscr{A}$ deux à deux disjoints, alors : $\mu\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n\geq 1} \mu(A_n)$.Cette dernière propriété est appelée σ-additivité (le σ référant évidemment à l'aspect dénombrable).

Lorsque $\mu(X) < +\infty$ on parle de mesure finie. En particulier, si $\mu(X) = 1$, on parle de mesure de probabilité (d'ailleurs, comment ne pas reconnaître ci-dessous les axiomes définissant une probabilité ? !). La mesure nulle valant 0 pour tout élément A de $\mathscr{P}(X)$ est effectivement une mesure. De même pour la mesure de Dirac ou encore la mesure de comptage valant $\operatorname{card}(A)$ si A est fini et $+\infty$ autrement. Néanmoins une mesure surpasse toutes les autres, la voici.

51.6. **Théorème.** *Il existe une unique mesure sur $(\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d))$ dite **mesure de Lebesgue** sur $\mathbf{R}^d$, notée λ_d, vérifiant :*

(i) $\lambda_d([0,1]^d) = 1$,

(ii) $\forall a \in \mathbf{R}^d, \forall A \in \mathscr{B}(\mathbf{R}^d), \lambda_d(a + A) = \lambda_d(A)$.

Lorsque d vaut 1 on se permet de noter la mesure de Lebesgue λ. L'invariance par translation est beaucoup plus importante que ce qu'elle paraît. La construction et surtout l'unicité de cette mesure sont un petit peu pénibles et techniques à démontrer. (Il existe par ailleurs tout un procédé pour construire des mesures à coup de π-systèmes, classes monotones et théorème de Carathéodory, faudrait que je creuse cet aspect là un jour, Briane et Pagès prennent à peu près 20 pages bien denses pour le faire !) On esquive les propriétés de régularité intérieure et extérieure de la mesure de Lebesgue mais faut essentiellement avoir en tête les principales propriétés d'une mesure (quelconque) : **croissante pour l'inclusion**, **forte additivité** (i.e. $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ pour tous A et B dans $\mathscr{A}$), **sous-additivité** (i.e. pour $(A_n)_{n\in\mathbf{N}}$ une suite quelconque d'éléments de $\mathscr{A}$, on a $\mu(\bigcup_n A_n) \leq \sum_n \mu_{A_n}$), **continuité à gauche** (i.e. pour une suite *croissante* $(A_n)_{n\in\mathbf{N}}$ d'éléments de $\mathscr{A}$, on a $\mu(\bigcup_n A_n) = \lim_n \mu(A_n)$), **continuité (partielle) à droite** (i.e. pour une suite *décroissante* $(A_n)_{n\in\mathbf{N}}$ d'éléments de $\mathscr{A}$

16. L'ensemble des fonctions mesurables de $(X, \mathscr{A})$ dans $(\mathbf{K}, \mathscr{B}(\mathbf{K}))$ muni des opérations usuelles $+, \cdot, \times$ sur les fonctions forme une $\mathbf{K}$-algèbre.

telle qu'il existe $n_0 \in \mathbf{N}$ avec $\mu(A_{n_0}) < +\infty$, on a $\mu(\bigcap_n A_n) = \lim_n \mu(A_n)$).

La construction de la mesure de Lebesgue prolonge les résultats intuitifs déjà connus. Par exemple, pour $I \subset \mathbf{R}$ un intervalle, on a $\lambda(I)$ qui est égal à la longueur de I. De même $\lambda_d(P)$ est égal au volume de P, où P est un pavé, c'est-à-dire que P est un produit cartésien d'intervalle. On trouve immédiatement que $\lambda_d(P)$ est alors égal au produit des longueurs de chaque intervalle constituant le pavé.

Je ne sais pas si cela te sera utile mais la mesure de Lebesgue est en particulier une **mesure de Borel**, c'est-à-dire que pour tout compact K de $\mathbf{R}^d$, $\mu(K)$ est fini. En effet, tout compact de $\mathbf{R}^d$ étant borné, il existe $M \in \mathbf{R}$ tel que $K \subset [-M, M]^d$ et donc $\lambda(K) \leq (2M)^d < \infty$ (remarquons que l'on a utilisé la croissance pour l'inclusion de la mesure).

Suivant les idées précédemment introduites, on comprend aisément que pour définir **l'intégrale de Lebesgue**, il suffit de savoir le faire sur les fonctions étagées (en réalité, il suffit de savoir le faire sur les fonctions étagées *positives*). L'intégrale d'une fonction étagée (positive) f par rapport à la mesure μ est définie par :

$$\int_X f \mathrm{d}\mu = \sum_{\alpha \in f(X)} \alpha \mu(\{f = \alpha\}) \in \mathbf{R}_+ \cup \{+\infty\}. \tag{223}$$

On trouvera également les notations suivantes $\int_X f(x)\mathrm{d}\mu(x)$ ou $\int_X f(x)\mu(\mathrm{d}x)$. Par exemple, si μ est la mesure de Dirac en a, on trouve que l'intégrale d'une fonction f par rapport à μ vaut $f(a)$. C'est physiquement logique. Pour la mesure de comptage m, lorsque $X = \mathbf{N}$, on constate que l'intégration de Lebesgue généralise toute la théorie des séries bien connue. En effet : $\int_{\mathbf{N}} f(n)\mathrm{d}m(n) = \sum_{n \in \mathbf{N}} f(n)$. L'intégrale a le bon goût d'être **additive** (i.e. $\int_X (f+g)\mathrm{d}\mu = \int_X f\mathrm{d}\mu + \int_X g\mathrm{d}\mu$), **croissante** (i.e. $f \leq g \longrightarrow \int_X f\mathrm{d}\mu \leq \int_X g\mathrm{d}\mu$) et **homogène** (i.e. $\forall \in \mathbf{R}$, $\int_X \lambda f\mathrm{d}\mu = \lambda \int_X f\mathrm{d}\mu$). On peut également intégrer sur une partie de X, il suffit alors de multiplier f par une indicatrice de cette partie. Il faut donc désormais définir l'intégrale d'une fonction mesurable non nécessairement étagée.

51.7. **Définition.** Si f est mesurable, on définit $\displaystyle\int_X f\mathrm{d}\mu = \sup\left\{\int_X \varphi\mathrm{d}\mu, \varphi \leq f, \varphi \text{ mesurable étagée (positive)}\right\}$.

Un théorème centralissime (qui sert d'ailleurs à démontrer beaucoup beaucoup des théorèmes suivants) est celui de Beppo Levi, également appelé théorème de convergence monotone.

51.8. **Théorème.** *Soit $(f_n)_{n \in \mathbf{N}}$ une suite croissante de fonctions mesurables*[17] *(i.e. $0 \leq f_n \leq f_{n+1}$), alors $f := \lim\limits_n f_n$ est mesurable et $\int_X f\mathrm{d}\mu := \int_X \lim_n f_n\mathrm{d}\mu = \lim_n \int_X f_n\mathrm{d}\mu$.*

Dit autrement, on va pouvoir permuter limite et intégrale ! Une première application de ce grandissime théorème est la suivante : pour f mesurable, on a l'équivalence suivante : $\int_X f\mathrm{d}\mu = 0 \iff \mu(f \neq 0) = 0$. On obtient donc une caractérisation des tels f plus fine que ce que l'on aurait obtenu pour f continue (par morceaux) avec la théorie de Riemann ! Avec Lebesgue, on sent bien que f va être nul mais pas partout... f va être nul (μ-)presque partout ! Plus généralement, une fonction va vérifier une propriété $\mathscr{P}$ μ-presque partout si la mesure des éléments ne vérifiant pas cette propriété est nulle ! Le μ-p.p. (abrégé de "presque partout") c'est vraiment cool, ça permet d'oublier des points *qui ne comptent pas vraiment.* Par exemple, pour f et g deux fonctions mesurables, si $f = g$ μ-p.p., alors $\int_X f\mathrm{d}\mu = \int_X g\mathrm{d}\mu$. Logique.

On dit qu'une fonction est **μ-intégrable** si $|f|$ est μ-intégrable, i.e. $\int_X |f|\mathrm{d}\mu < +\infty$. On note comme suit l'ensemble des fonctions μ-intégrables définies sur X à valeurs dans $\mathbf{K}$:

$$\mathscr{L}^1_{\mathbf{K}}(X, \mathscr{A}, \mu) := \{f : (X, \mathscr{A}) \to (\mathbf{K}, \mathscr{B}(\mathbf{K})) \text{ mesurable et } \mu\text{-intégrable}\}. \tag{224}$$

Par commodité et comme on ne travaillera essentiellement qu'avec des boréliens, on prendra l'habitude de ne noter que $\mathscr{L}^1(\mu)$. Octroyons nous le droit une fois de déroger à cette commodité ! Si on munit l'espace mesurable $(\mathbf{N}, \mathscr{P}(\mathbf{N}))$ de la mesure de comptage m, alors on retombe évidemment sur nos pattes : $\mathscr{L}^1(m) = \ell^1(m)$ l'espace des suites absolument convergentes.

Remarquons qu'on sent qu'il commence à se passer de drôles de choses par rapport à ce que l'on connaît. En effet, sur $\mathscr{L}^1(\mu)$, $||f||_1 := \int_X |f|\mathrm{d}\mu$ n'est pas une norme mais une semi norme (on n'a pas $||f||_1 = 0 \implies f = 0$ mais seulement f nul p.p.). Mais néanmoins, pour des fonctions bien régulières (continues par exemple), rien de nouveau sous le soleil ! Dans le cas des fonctions Riemann intégrables définies sur un compact, théorie de Lebesgue et Riemann coïncident.

---

17. Remarquer que je ne précise pas sur quoi qui est mesurable tant cela apparaît évident, puis essentiellement, on se limite à la tribu borélienne.

Passons désormais sur un arsenal de théorèmes de convergence (et c'est typiquement ce qui fait la force de la théorie de Lebesgue, rappelons le !). Il va bien falloir garder à l'esprit que le **théorème de Beppo Levi** également appelé **théorème de convergence monotone** permet de prouver les résultats qui vont suivre.

51.9. **Lemme.** (Fatou) *Soit* $(f_n)_{n\in\mathbf{N}}$ *une suite de fonctions mesurables* ***positives****, alors :* $\int_X \liminf f_n \mathrm{d}\mu \leq \liminf \int_X f_n \mathrm{d}\mu$.

51.10. **Théorème.** (**convergence dominée**) *Soit* $(f_n)_{n\in\mathbf{N}}$ *une suite d'éléments de* $\mathscr{L}^1(\mu)$ *vérifiant :*

(i) lorsque n *tend vers l'infini,* $f_n(x)$ *converge* μ*-p.p.,*

(ii) il existe $g \in \mathscr{L}^1(\mu)$ *telle que, pour tout* $n \geq 1$*, on a* $|f_n(x)| \leq g(x)$ μ*-p.p.*

Alors, il existe $f \in mathscrL^1(\mu)$ *tel que* $f_n(x)$ *converge vers* $f(x)$ μ*-p.p. et on a :*

$$\lim_n \int_X f_n \mathrm{d}\mu = \int_X f \mathrm{d}\mu \quad \text{et même} \quad \lim_n \int_X |f - f_n| \, \mathrm{d}\mu = 0. \tag{225}$$

Encore une fois, on peut agréablement permuter limite et intégrale ! En bidouillant un petit peu ce théorème et celui de convergence monotone, on obtient que pour des fonctions φ_n mesurables, si $\sum_{n\geq 1} \int_X |\varphi_n| \mathrm{d}\mu < +\infty$, alors les fonctions φ_n, $\sum_{n\geq 1} |\varphi_n|$ et la fonction définie μ-p.p. $\sum_{n\geq 1} \varphi_n$ sont μ-intégrables. En bonus, on a :

$$\int_X \left(\sum_{n\geq 1} \varphi_n \right) \mathrm{d}\mu = \sum_{n\geq 1} \int_X \varphi_n \mathrm{d}\mu. \tag{226}$$

Grâce à cela, on peut montrer le lemme de Borel-Cantelli.

51.11. **Lemme.** *Soit* $(A_n)_{n\in\mathrm{N}}$ *une famille de parties de* $\mathscr{A}$*, alors :*

$$\sum_{n\geq 1} \mu(A_n) < +\infty \implies \mu(\limsup A_n) = 0. \tag{227}$$

Terminons cette série de lemme-théorème par deux résultats qui ne sont que des applications du théorème de convergence dominée.

51.12. **Théorème.** (continuité sous le signe intégrale) *Soit* $f : \mathbf{R} \times X \to \mathbf{R}$ *et* $\tilde{u} \in \mathbf{R}$*. Si :*

(i) pour tout $u \in \mathbf{R}$*, l'application* $x \mapsto f(u, x)$ *est mesurable,*

(ii) l'application $u \mapsto f(u, x)$ *est continue en* $\tilde{u}$ μ*-p.p.,*

(iii) il existe $g \in \mathscr{L}^1(\mu)$ *telle que* $\forall u \in \mathbf{R}, |f(u, x)| \leq g(x)$ μ*-p.p.*

Alors la fonction $F(u) := \int_X f(u, x)\mu(\mathrm{d}x)$ *est définie en tout point* $u \in \mathbf{R}$ *et est continue en* $\tilde{u}$*.*

51.13. **Théorème.** (dérivation sous le signe intégrale) *Soit* $I \subset \mathbf{R}$ *un intervalle ouvert non vide,* $f : I \times X \to \mathbf{R}$ *et* $\tilde{u} \in I$*. Si :*

(i) pour tout $u \in I$*,* $f(u, \cdot) \in \mathscr{L}^1(\mu)$*,*

(ii) $\frac{\partial f}{\partial u}(\tilde{u}, x)$ *existe pour* μ *presque tout* x*,*

(iii) il existe $g \in \mathscr{L}^1(\mu)$ *telle que* $\forall u \in I, |f(u, x) - f(\tilde{u}, x)| \geq g(x)|u - \tilde{u}|$ *pour* μ *presque tout* x*.*

Alors la fonction $F(u) := \int_x f(u, x)\mu(\mathrm{d}x)$ *est définie en tout point* u *de* I*, et est dérivable en* $\tilde{u}$ *de dérivée* $F'(\tilde{u}) = \int_X \partial_u f(\tilde{u}, x)\mu(\mathrm{d}x)$*.*

On peut désormais considérer le "récital" de théorie de la mesure terminé. [18] On va considérer les espaces canoniques dans lesquels se déploie la théorie de l'intégration... sachant que l'on en a déjà vu un, le beau mais ténébreux $\mathscr{L}^1(\mu)$. En fait, j'ai menti, ce ne sont pas eux les L *ronds* les bons espaces mais les L *droits*. Venons y petit à petit.

51.14. **Définition.** Pour tout réel $p > 0$, on définit l'espace suivant :

$$\mathscr{L}^p_{\mathbf{K}}(X, \mathscr{A}, \mu) := \left\{ f : (X, \mathscr{A}) \to (\mathbf{K}, \mathscr{B}(\mathbf{K})) \text{ mesurable et } \int_X |f|^p \mathrm{d}\mu < +\infty \right\}. \tag{228}$$

---

18. Pour un point de vue de dynamicien sur le sujet, se référer au chapitre 0 du Mañé, *Ergodic theory and Differentiable dynamics*. Merci Caleb pour la référence !

Une nouvelle fois, le contexte rendant claire la situation, on prendra l'habitude de ne noter que $\mathscr{L}^p(\mu)$. Cet espace est évidemment un **espace vectoriel** (on verra ensuite qu'il a quelques propriétés supplémentaires qui le rendent enviable à merci à beaucoup de ses confrères ; par ailleurs, bienvenue dans l'analyse dans des espaces de dimension infinie !). On fera bien attention à ne pas trop croire en ses rêves : en général, on n'a pas $0 < p \leq q \implies \mathscr{L}^q(\mu) \subset \mathscr{L}^p(\mu)$ (pour un exemple [19]). Il est nécessaire (mais non suffisant ! penser à la mesure de comptage) que μ soit de masse finie (c'est par exemple le cas pour une mesure de probabilité).

Dans l'optique de faire de l'analyse, on se souvient du crédo de J. Dieudonné résumant en ces mots la démarche d'un analyste : *majorer, minorer, approximer.* Présentons donc les deux principales inégalités pour la (semi-)**norme** $\mathscr{L}^p$ usuelle $||f||_p := \left(\int_X |f|^p \mathrm{d}\mu\right)^{1/p}$. On verra comment passer d'une semi norme à une norme, ce n'est pas très important pour l'instant.

51.15. **Proposition.** (**Inégalité de Hölder**) *Soient f et g mesurables, p et q des exposants conjugués (i.e. $1/p + 1/q = 1$). Si $f \in \mathscr{L}^p(\mu)$ et $g \in \mathscr{L}^q(\mu)$, alors $fg \in \mathscr{L}^1(\mu)$ et $||fg||_1 \leq ||f||_p ||g||_p$. On a égalité si et seulement si il existe $(\alpha, \beta) \in \mathbf{R}_+^2 \setminus \{(0,0)\}$ tels que $\alpha|f|^p = \beta|g|^q$ μ-p.p.*

Dans le cas où $p = q = 2$, on retrouve la célèbre inégalité de Cauchy-Schwarz. Bien noter que l'on ne demande pas aux exposants conjugués d'être forcément des entiers ! (Dans les faits, on considèrera les p et q **réels** tels que $1 \leq p \leq q \leq +\infty$, mais bien faire attention aux cas $p = 1$ ou $p = +\infty$.)

51.16. **Proposition.** (**Inégalité de Minkowski**) *Si $p \in [1, +\infty[$, alors $(\mathscr{L}^p(\mu), ||\cdot||_p)$ est un espace vectoriel (semi-)normé. En particulier $\forall f, g \in \mathscr{L}^p(\mu), ||f + g||_p \leq ||f||_p + ||g||_p$.*

Ces deux inégalités sont absolument fondamentales ! (Une troisième est fondamentale lorsque l'on dispose d'une mesure de masse finie : l'inégalité de Jensen. Je ne la présente pas car je doute que tu aies souvent à traiter des inégalités de convexité, au besoin, tu sais que ça existe.)

Bon, pourquoi je m'embête depuis tout à l'heure à dire que l'on n'a pas une norme mais une semi-norme ? Simplement car $||f||_p = 0$ n'est pas équivalent à $f = 0$. Tout ce que l'on peut dire c'est que f va être nulle **presque partout**. Donc, sans surprise, pour obtenir une norme on va devoir raisonner sur un représentant d'une classe d'équivalence. La relation d'équivalence est toute trouvée : être égal presque partout. On quotiente $\mathscr{L}^p(\mu)$ par cette relation d'équivalence et l'on définit l'espace vectoriel quotient obtenu comme étant $\mathrm{L}^p(\mu)$. Prudence est mère de sûreté... Les éléments de cet espace sont des classes d'équivalence, donc dire que deux éléments sont égaux dans $\mathrm{L}^p(\mu)$ est un peu casse gueule. On commettra un nouvel abus qu'est celui d'oublier que l'on travaille avec des représentants de classes d'équivalence. On se dira naïvement que ce sont des éléments tout ce qu'il y a de plus normal (et dans l'absolue majorité des cas, ça ne pose pas de soucis).

L'espace vectoriel normé ainsi construit a le bon goût [20] d'être **complet** (théorème de Riesz-Fisher, 1907), c'est un Banach. Cela signifie que toute suite de Cauchy converge au sens de la norme $||\cdot||_p$. On vient de débloquer un nouveau mode de convergence ! Un corollaire immédiat, particulièrement utile en physique est le suivant.

51.17. **Corollaire.** *L'espace $L^2(\mu)$ muni du produit scalaire canonique est un espace de Hilbert.*

C'est en théorie à ce moment là que les formes linéaires continues font leur apparition. Tu as de très jolis résultats de représentations (Riesz-Fisher, Riesz) et de décomposition (Radon, Radon-Nikodym). Je passe (au moins momentanément), c'est trop technique pour moi pour l'instant. Il va falloir que je m'y penche sérieusement dessus.

Il ne manque qu'un seul petit ingrédient pour pouvoir intégrer sur des **espaces produits** : trouver une tribu naturelle à adjoindre à l'espace produit.

51.18. **Définition.** Soient $(X, \mathscr{A})$ et $(Y, \mathscr{B})$ deux espaces mesurables. On appelle **tribu produit** de $\mathscr{A}$ et $\mathscr{B}$ la tribu $\mathscr{A} \otimes \mathscr{B}$ égale à $\sigma(\{A \times B, A \in \mathscr{A}, B \in \mathscr{B}\})$.

D'un point de vue pédestre, la tribu produit n'est rien d'autre que la tribu engendrée par les pavés à côtés mesurables. Et là c'est banco ! La tribu est suffisamment bien faite pour qu'elle soit *géométriquement naturelle* [21], qu'elle jouisse de propriétés

19. Sur $(\mathbf{R}, \mathscr{B}(\mathbf{R}))$ muni de la mesure de Lebesgue λ, on n'a aucune inclusion entre $\mathscr{L}^1$ et $\mathscr{L}^2$:

$$x \mapsto \frac{1_{]0,1]}}{\sqrt{x}} \in \mathscr{L}^1(\lambda) \setminus \mathscr{L}^2(\lambda) \quad \text{et} \quad x \mapsto \frac{1}{\sqrt{x^2+1}} \in \mathscr{L}^2(\lambda) \setminus \mathscr{L}^1(\lambda). \tag{229}$$

20. Il a également le bon goût de disposer de propriétés de densité bien joyeuses. Par exemple, dans $(\mathbf{R}, \mathscr{B}(\mathbf{R}))$ muni de la mesure de Lebesgue λ, l'ensemble des fonctions en escalier à support compact est dense dans tous les espaces $\mathrm{L}^p(\lambda)$, $1 \leq p < +\infty$. C'est également le cas des fonctions continues à support compact. Remarque que les cas $+\infty$ ont été mis de côté. Il ne semble pas possible (?) de dire plus que les fonctions étagées sont denses dans $\mathrm{L}^\infty(\mu)$. De même, tu as des résultats de dualité qui foirent pour $p = +\infty$ (cf. les espaces réflexifs), de même pour des résultats de séparabilité.

21. *Géométriquement* en un sens volontairement vague, i.e. on cherche la plus petite tribu rendant les projections canoniques mesurables.

que l'on est bien en droit de lui demander. Par exemple une fonction est mesurable sur l'espace produit si et seulement si chacune de ses composante est mesurable. Néanmoins, on se gardera de ne pas manquer de quelques attentions. Pour X et Y deux espaces topologiques, on a toujours $\mathscr{B}(X) \otimes \mathscr{B}(Y) \subset \mathscr{B}(X \times Y)$. L'inclusion réciproque n'est pas toujours vraie, si X et Y sont à base dénombrable d'ouverts elle l'est par exemple. Par exemple, $\mathscr{B}(\mathbf{R}^2) = \mathscr{B}(\mathbf{R}) \otimes \mathscr{B}(\mathbf{R})$. Plus généralement, $\mathscr{B}(\mathbf{R}^d) = \mathscr{B}(\mathbf{R}) \otimes \cdots \otimes \mathscr{B}(\mathbf{R})$ d fois, noté $\mathscr{B}(\mathbf{R})^{\otimes d}$.

Le dernier problème à régler est celui de la mesure produit. Peut-on trouver une mesure sur l'espace produit $(X \times Y, \mathscr{A} \otimes \mathscr{B})$ héritant des propriétés inhérentes à $(X, \mathscr{A})$ et $(Y, \mathscr{B})$? Parfaitement, oui.

51.19. **Proposition.** *Soient $(X, \mathscr{A}, \mu)$ et $(X, \mathscr{B}, \nu)$ deux espaces mesurés σ-finis*[22]*. Il existe alors une* ***unique*** *mesure sur $(X \times Y, \mathscr{A} \otimes \mathscr{B})$ notée $\mu \otimes \nu$ vérifiant : $\forall A \in \mathscr{A}, B \in \mathscr{B}, (\mu \otimes \nu)(A \times B) = \mu(A)\nu(B)$.*

Application directe à la mesure de Lebesgue ! On munit $(\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d))$ de la mesure produit $\lambda_d := \lambda^{\otimes d}$ égale à d produit de la mesure de Lebesgue sur $\mathbf{R}$. Essentiellement, tu n'as *que* deux théorèmes à bien savoir manipuler. On fera bien attention à remarquer les hypothèses de **positivité** et d'**intégrabilité.**

51.20. **Théorème.** (**Fubini-Tonelli**) *Soient $f : (X \times Y, \mathscr{A} \otimes \mathscr{B}) \to \mathbf{R}_+$ une fonction mesurable, μ et ν deux mesures σ-finis sur $(X, \mathscr{A})$ et $(Y, \mathscr{B})$. Alors les fonctions suivantes sont mesurables :*

$$x \mapsto \int_Y f(x,y)\nu(\mathrm{d}y), \quad y \mapsto \int_X f(x,y)\mu(\mathrm{d}x). \tag{230}$$

De plus, on a le classique :

$$\int_{X \times Y} f(x,y)\mathrm{d}(\mu \otimes \nu) = \int_X \left(\int_Y f(x,y)\nu(\mathrm{d}y) \right) \mu(\mathrm{d}x) = \int_Y \left(\int_X f(x,y)\mu(\mathrm{d}x) \right) \nu(\mathrm{d}y) \tag{231}$$

51.21. **Théorème.** (**Fubini-Lebesgue**) *Considérons l'espace mesuré $(X \times Y, \mathscr{A} \otimes \mathscr{B}, \mu \otimes \nu)$. Soit $f \in \mathrm{L}^1(\mu \otimes \nu)$. Alors :*

$$x \mapsto f(x,y) \in \mathrm{L}^1(\mu)\ \nu(\mathrm{d}y)\text{-p.p.}, \quad y \mapsto f(x,y) \in \mathrm{L}^1(\nu)\ \mu(\mathrm{d}x)\text{-p.p.} \tag{232}$$

$$x \mapsto \int_Y f(x,y)\nu(\mathrm{d}y) \in \mathrm{L}^1(\mu)\ \mu\text{-p.p.}, \quad y \mapsto \int_X f(x,y)\mu(\mathrm{d}x) \in \mathrm{L}^1(\nu)\ \nu\text{-p.p.} \tag{233}$$

On a alors :

$$\int_{X \times Y} f(x,y)\mathrm{d}(\mu \otimes \nu) = \int_X \left(\int_Y f(x,y)\nu(\mathrm{d}y) \right) \mu(\mathrm{d}x) = \int_Y \left(\int_X f(x,y)\mu(\mathrm{d}x) \right) \nu(\mathrm{d}y) \tag{234}$$

52. Probabilités avec Laure Coutin

53. Inégalité de Hoeffding

54.

22. Hypothèse technique mais absolument nécessaire. Lorsque la mesure n'est pas de masse finie, i.e. $\mu(X)$ non fini, on peut tout de même espérer que la mesure soit une agglutination de composantes finies. On dira qu'une mesure sur un espace mesurable $(X, \mathscr{A})$ est σ-finie, s'il existe une suite croissante $(E_n)_n$ d'éléments de $\mathscr{A}$ tels que X soit l'union croissante des E_n et que leur mesure soit finie pour tout n. Par exemple, la mesure de Lebesgue est σ-finie. En effet, considérer $E_n =]-n, n[$. C'est bien de mesure finie $\lambda(E_n) = 2n < +\infty$ pour tout $n \in \mathbf{N}$.

Deuxième partie 2. **Stage de L3**

55.

56.

57.

58.

59.

60.

61.

62.

63.

64.

65.

66.

67.

68.

69.

70.

71.

72.

73.

74.