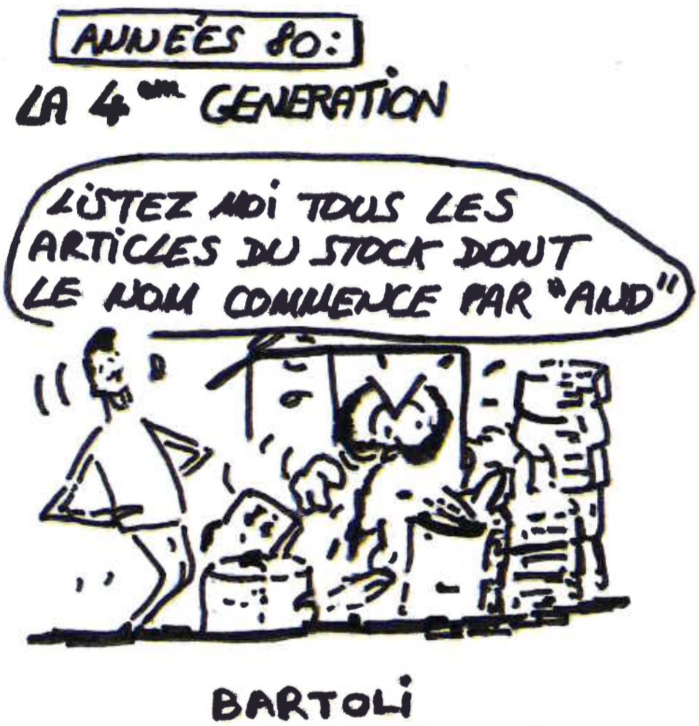
BIAA

Bulletin d'Informatique Approfondie & Applications
Computation - Information
Volumes 2016

103-104-105



Aix✱Marseille université

Publication trimestrielle, gratuite, de l'Université d'Aix Marseille
Dépôt légal : février 2016 - http://www.biaa.eu - ISSN 0291 - 5413

B.I.A.A.

Bulletin d'Informatique Approfondie et Applications
Revue fondée par Edmond Bianco
Publication trimestrielle de l'Université d'Aix-Marseille
ISSN 0291-5413

Le bulletin d'informatique approfondie et applications est une revue pluridisciplinaire destinée à éclairer les connaissances fondamentales informatiques. Les fondements sont un domaine vaste allant de la structure intérieure de l'ordinateur, où se matérialise la machine universelle, à l'algorithme qui devient programme, pour aboutir à la notion de système. Nous contribuons ainsi à ce que les autres disciplines plus anciennes (sciences humaines et de la société, sciences de la matière et de l'énergie, sciences mathématiques, sciences de la nature, sciences de la terre, sciences de l'univers, sciences de la vie, etc.) n'aient pas tendance à considérer l'informatique comme un simple outil définitivement figé. Il importe de continuer à maîtriser les développements fondamentaux de l'informatique pour que nos disciplines puissent en tirer un meilleur parti.

Notre publication est ouverte à l'ensemble de la communauté scientifique. Le périodique est diffusé vers les bibliothèques universitaires de France et vers quelques bibliothèques des cinq continents.

Directeur de la publication
Jean - Michel Knippel

Responsable de l'édition
Eric Olivier

Serveur de publication
Christian Blanvillain

Secrétariat
Kalassoumi Adjilani
Université d'Aix-Marseille
Site St Charles, Case 33
3 place Victor Hugo
F -13331 Marseille Cedex 3
Téléphone : +33 (0) 413 550 252

Dépositaire
Université d'Aix-Marseille
Bibliothèque Universitaire
Site St Charles
3 place Victor Hugo
F -13331 Marseille Cedex 3
Téléphone : +33 (0) 413 550 579

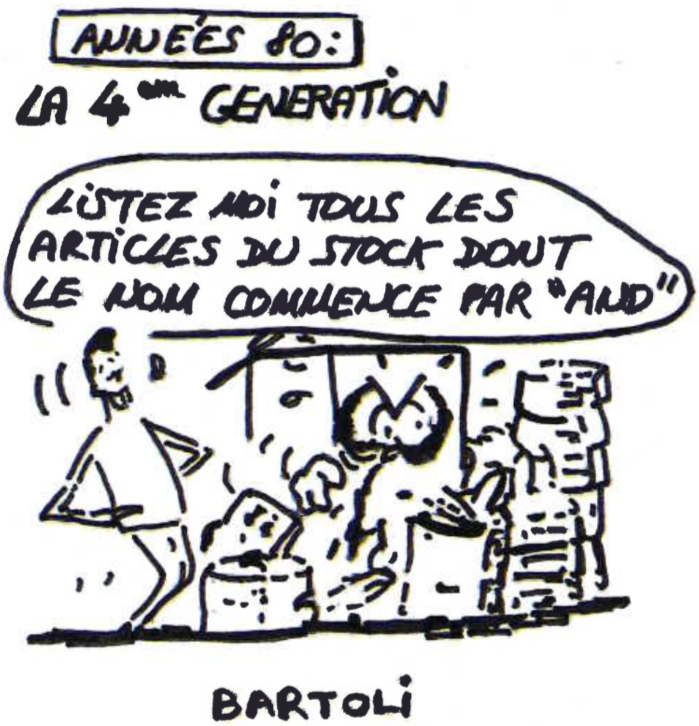
Imprimeur
Université d'Aix-Marseille
Service Reprographie
Site St Charles
3 place Victor Hugo
F -13331 Marseille Cedex 3
Téléphone : +33 (0) 413 551 569

Table Des Matières

BIAA

Bulletin d'Informatique Approfondie & Applications
Computation - Information
Volumes 2016

103-104-105



Publication trimestrielle, gratuite, de l'Université d'Aix Marseille
Dépôt légal : mars 2016 - http://www.biaa.eu - ISSN 0291 - 5413

Couverture : dessin de Michel Avezard (Zevar) interprété par la MMIAGe (1985-1987)

(2016) B.I.A.A. no **103** BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
ISSN 0291-5413 (Université d'Aix-Marseille)

ÉDITORIAL : Information, Communication, James Clerk Maxwell et Albert Einstein

Jean - Michel KNIPPEL

En 1987, Jacques Arsac arrive de manière inéluctable à la définition suivante de l'information : *"L'information est un texte susceptible d'apporter une connaissance."* [Ars87]. Comme le dit Wolfgang Wildgen, "le discours peut être oral ou écrit dans le cas du récit et du texte, il peut être visuel comme le film muet ou essentiellement visuel, mais soutenu et élaboré par la parole, les bruits et la musique dans le film moderne. Enfin il peut être mimique, gestuel ou dansé (dans la pantomime et le ballet). Il est difficile de s'imaginer un récit olfactif ou gustatif ; pourtant une suite de repas ou de vins dégustés pourrait à la limite constituer un récit" [Wil05].

Retournons en 1989, John Archibald Wheeler, pionnier de la fission nucléaire, sortit sa dernière formule : *"It From Bit"*. Chaque *it* - chaque particule, champ de forces, même le continuum de l'espace-temps - tire sa fonction, son sens, son existence même [...] de *bits*. L'information est quantifiée. Le bit serait l'ultime particule insécable [Gle11].

Arrivons à la communication, le sens retenu ici est celui des télécommunications, par le moyen desquelles il y a des échanges de messages avec un contenu intellectuel [Lich83]. L'électricité ne porte pas notre texte, par exemple, elle permet à l'opérateur à l'autre bout de la ligne d'en écrire un semblable.

James Clerk Maxwell au milieu du XIXème siècle décrit le comportement classique des champs électrique et magnétique, mais également celui de la lumière : théorie électromagnétique. Là, nous arrivons à Albert Einstein, car la théorie de la relativité restreinte, issue de l'examen des équations de Maxwell, a été établie par Henri Poincaré, Hendrick Antoon Lorentz et Einstein (sans oublier Hermann Minkowski qui en donna plus tard une élégante version géométrique) [Pen93].

Selon le physicien Hubert Krivine, la relativité générale est difficile à exposer et encore plus à comprendre. L'usage pour vulgariser est de dire qu'elle décrit un espace-temps qui est déformé par la présence des masses.

N'oublions pas que la théorie de la relativité générale d'Einstein ne peut être comprise que si l'on a bien assimilé la relativité restreinte. La relativité générale a un impact en matière de technologie dans le calcul des orbites des sondes spatiales qui tiennent compte des effets de la théorie de la relativité générale. Il existe des dispositifs capables de déterminer un point à la surface de la terre avec une précision telle (quelques dizaines de centimètres) que les effets dus à la courbure de l'espace-temps doivent être effectivement pris en compte [You05].

Reste un bémol, dans ce vaste champ d'information, de communication, d'électromagnétisme, de relativité, le canal de communication. Le discours engage tous les sens, l'ouïe, la vue, l'odorat, le goût éventuellement et le toucher (très en vogue, la tape sur l'épaule dans certains milieux). Et qu'en est-il du toucher ? 2015. Un universitaire londonien, Adrian David Cheok, s'est penché sur la question des odeurs. Et de ses recherches est finalement né le Scentee, un croisement entre un diffuseur de parfum traditionnel et votre téléphone. Hiroo Iwata de l'Université de Tsukuba au Japon et ses collègues proposent le système Food Simulator. Ce simulateur de goût se place dans la cavité buccale de l'utilisateur. Il libère une odeur, un goût, et émet des bruits de mastication qui correspondent à l'aliment prétendument mastiqué. Des chercheurs japonais ont développé RingU, un appareil en forme de bague qui permet de communiquer des sensations via internet.

Références.

[Ars87] J. Arsac. *Les machines à penser*. Seuil, 1987.

[Esc91] R. Escarpit. *L'information et la communication. Théorie générale*. Collection Hachette Université. Communication. Hachette, 1991.

[Gle11] J. Gleick. *l'information. L'histoire - La théorie - Le déluge*. Cassini, 2011.

[Lich83] A. Lichnerowicz, Fr. Perroux, G. Gadoffre. *Information et communication*. Collection "Recherches Interdisciplinaires". Maloine S.A., 1983.

[Pen93] R Penrose. *L'esprit, l'ordinateur et les lois de la physique*. InterEditions, 1993.

[Qui70] J.-Cl. Quiniou, J.-M. Font, G. Verroust, J.-M. Philippe, Cl. Marenco. *Les cerveaux non humains, 1970*.

[Wil05] W. Wildgen. *Dynamique narrative du texte, du film et de la musique*. Cahiers de Narratologie. N° 28. http ://narratologie.revues.org/, 2015.

[You05] P. Yourgrau. *Einstein/Gödel. Quand deux génies refont le monde*. Dunod, 2005.

(2016) B.I.A.A. no **103** BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
ISSN 0291-5413 (Université d'Aix-Marseille)

Mécanique quantique I : bases mathématiques

Eric OLIVIER [1] [2]

Résumé. – La formulation moderne de « *la mécanique quantique* » est présentée de manière systématique (axiomatique) par Von Neumann au tournant des années 30 : un état quantique est alors vu comme un « *point* » de la sphère unité d'un « *espace de Hilbert complexe* » de dimension infinie. Dans cette première note, nous donnons les repères d'analyse fonctionnelle nécessaires pour comprendre comment cette « *nouvelle mécanique* » s'attache à décrire le réel microscopique. En particulier la transformée de Fourier sur l'espace de Hilbert des signaux d'énergie finie nous permet d'énoncer et de démontrer l'inégalité de Heisenberg-Gabor dans le bon espace de Sobolev. Nous verrons dans les notes suivantes que cette inégalité correspond à l'inégalité de Heisenberg pour la position et l'impulsion des fonctions d'onde de la mécanique ondulatoire.

1. Introduction

Dans le milieu des années 20, *« la mécanique des quanta »* se formalise sous deux formes apparemment différentes (mais en fait équivalentes), avec d'une part *« la mécanique des matrices »* de Heisenberg, Jordan et Born et d'autre part *« la mécanique ondulatoire »* de de Broglie et Schrödinger. La synthèse de ces deux approches donnera *« la mécanique quantique »*. Dans cette série de notes, nous commençons par adopter le point de vue de de Broglie et Schrödinger [3]. Dans une lignée de savants éminents (principalement Descartes, Fermat, Newton, Huyghens, Maupertuis, Hamilton, Jacobi, Maxwell, Fresnel, Einstein,...), l'approche de de Broglie et Schrödinger marque une étape significative de plusieurs siècles de polémiques et de réflexions sur l'ambivalence entre *« l'optique géométrique »* et *« l'optique ondulatoire »* [4] (nous y reviendrons dans [Oli17]). Pour comprendre les idées mises en jeu, nous avons besoin d'un certain nombre de bases d'analyse fonctionnelle et tout spécialement de la théorie des *« espaces de Hilbert »*. Nous considérons ici qu'un espace de Hilbert est un espace de Banach (espace vectoriel normé complet) dans lequel l'identité du parallélogramme est satisfaite (théorème de Fréchet-von Neumann-Jordan) : cela entraîne que la norme est nécessairement déduite d'un produit scalaire (dont l'expression est donnée par l'identité de polarization). Dans cette présentation, *« l'inégalité de Cauchy-Schwarz »* est une conséquence directe de *« l'inégalité triangulaire »* (qui pré-existe comme propriété de la norme) : ce point est particulièrement intéressant dans la mesure où nous verrons (c.f. [Oli16a]) que *« l'inégalité de Heisenberg »* découle directement de l'inégalité de Cauchy-Schwarz et donc de l'inégalité triangulaire !

Dans cette première note, nous donnons une description rapide des principaux espaces fonctionnels classiques : les espaces de Banach avec les espaces L^p (c.f. Section 2), les espaces de Hilbert (c.f. Section 5) et enfin les espaces de Sobolev (c.f. Section 9). La

1. GDAC-I2M UMR 7373 CNRS Université d'Aix-Marseille
2. eric.olivier@univ-amu.fr
3. Nous verrons plus tard comment le point de vue de Heisenberg autorise la modélisation de phénomènes quantiques très généraux (statistique quantique) dans le cadre du formalisme des *« C^*-algèbres »* tel qu'initié par Von-Neumann.
4. On parle aussi *« d'optique physique »*.

Section 8 est consacrée à la théorie des probabilités sur la droite réelle. Nous insistons sur des points élémentaires de la théorie qui jouent un rôle en mécanique quantique : en particulier la signification *intuitive* attachée aux moments d'ordre 1 et 2 d'une distribution de probabilité (i.e. la moyenne et l'écart type) autorise une interprétation du *« théorème d'Ehrenfest »* (sur la correspondance entre les trajectoires classiques et quantiques) ainsi que du *« principe d'incertitude de Heisenberg »* dans le formalisme probabiliste de la mécanique quantique proposée par Born (nous y reviendrons dans [Oli16b]). De manière anachronique et hors du contexte de la mécanique quantique, nous utilisons les propriétés des moments probabilistes pour aborder le principe d'incertitude à partir de la théorie du signal et de la transformée de Fourier des signaux d'énergie finie (c.f. Sections 6 & 7). Plus précisément, nous démontrons *« l'inégalité de Heisenberg-Gabor »* dans le cas où les signaux en question sont dans le bon espace de Sobolev. Cette approche de la mécanique quantique par la théorie du signal est intéressante à deux titres. Premièrement, elle permet de s'abstraite d'un contexte physique où les diverses interprétations compliquent les choses et deuxièmement, elle explique – en un sens – le rôle joué par la transformée de Fourier dans *« le principe de correspondance de Schrödinger »* et la définition des couples d'opérateurs conjugués (c.f. [Oli16b, Oli17]). Typiquement les opérateurs position et impulsion existent déjà – implicitement – dans l'inégalité de Heisenberg-Gabor.

2. Espace de Banach – Algèbre de Banach

2.1. Pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, un $\mathbb{K}$-espace vectoriel normé $(\mathscr{E}, \|\cdot\|)$ est appelé un *« espace de Banach »* lorsqu'il est complet, c'est-à-dire lorsque toute suite de points de $\mathscr{E}$ qui est de Cauchy possède une limite dans $\mathscr{E}$. Rappelons qu'une série $\sum_{k=0}^{\infty} a_k$ de $\mathscr{E}$ est dite *« normalement convergente »* (resp. *« absolument convergente »*) ssi les sommes partielles $\sum_{k=0}^{n} a_k$ convergent dans $\mathscr{E}$ quand $n \to +\infty$ (resp. la série numérique $\sum_{k=0}^{\infty} \|a_k\|$ est convergente) : nous utiliserons la caractérisation suivante qui est très pratique.

Lemme 2.1. *L'espace normé $(\mathscr{E}, \|\cdot\|)$ est de Banach ssi toute série absolument convergente de $\mathscr{E}$ est normalement convergente sur $\mathscr{E}$.*

Preuve. Soit $\sum_{k=0}^{\infty} a_k$ une série absolument convergente de l'espace de Banach $(\mathscr{E}, \|\cdot\|)$; alors les $A_n := \sum_{k=0}^{n} a_k$ $(n \geq 0)$ forment une suite de points de $\mathscr{E}$ qui est de Cauchy et nous notons A sa limite : le fait que la série $\sum_{k=0}^{\infty} a_k$ soit normalement convergente sur $(\mathscr{E}, \|\cdot\|)$ découle alors de l'inégalité $\|A_n - A\| \leq \sum_{k=n+1}^{\infty} \|a_k\|$ qui assure que $\|A_n - A\| \to 0$ quand $n \to +\infty$. Réciproquement, supposons que toute série absolument convergente de $(\mathscr{E}, \|\cdot\|)$ est normalement convergente et soit $a_0, a_1, \ldots$ une suite de $\mathscr{E}$ possédant la propriété de Cauchy : il s'agit de montrer que cette suite converge dans $\mathscr{E}$. Pour cela nous partons du fait qu'il existe une sous-suite $a_{n_0}, a_{n_1}, \ldots$ t.q. $\|a_{n_k} - a_{n_{k-1}}\| \leq 1/2^k$. Si nous posons $\alpha_0 := a_{n_0}$ et $\alpha_k := a_{n_k} - a_{n_{k-1}}$, pour tout $k \geq 1$, alors

$$\sum_{i=0}^{k} \alpha_i = a_{n_0} + (a_{n_1} - a_{n_0}) + \cdots + (a_{n_k} - a_{n_{k-1}}) = a_{n_k}$$

Or, par construction $\|\alpha_k\| \leq 1/2^k$, pour tout $k \geq 1$; par suite $\sum_{k=0}^{\infty} \alpha_k$ est une série absolument convergente de $\mathscr{E}$ et donc (hypothèse) normalement convergente : en d'autres termes, $a_* := \sum_{k=0}^{\infty} \alpha_k$ est un élément de $\mathscr{E}$ t.q. $\|a_* - a_{n_k}\| = \|a_* - \sum_{i=0}^{k} \alpha_i\| \to 0$ quand

$k \to +\infty$. Nous venons de démontrer que a_* est une valeur d'adhérence de la suite des a_n : mais cette suite étant de Cauchy, elle converge nécessairement vers a_*.

□

2.2. Soit $\mathscr{A}$ une $\mathbb{K}$-algèbre[5] normée, en ce sens que $\mathscr{A}$ (considéré comme $\mathbb{K}$-espace vectoriel) est muni d'une norme $\|\cdot\|$. Alors $\mathscr{A}$ est appelée *« une algèbre de Banach »* si l'espace vectoriel normé $(\mathscr{A}, \|\cdot\|)$ est un espace de Banach et si de plus la norme est sous multiplicative en ce sens que $\|AB\| \leq \|A\|\,\|B\|$, pour tout $A, B \in \mathscr{A}$; si de plus il existe un élément I in $\mathscr{A}$ t.q. $AI = IA = A$, pour tout $A \in \mathscr{A}$, alors l'algèbre de Banach est dite unitaire. L'espace $(\mathscr{C}[0\,;1], \|\cdot\|_\infty)$ des fonctions réelles continues sur l'intervalle $[0\,;1]$ et muni de la norme uniforme est un espace de Banach ; étant données f et g deux fonctions quelconques de $\mathscr{C}[0\,;1]$, il est immédiat de vérifier $\|fg\|_\infty \leq \|f\|_\infty \|g\|_\infty$: il en découle que $\mathscr{C}[0\,;1]$ est une algèbre de Banach (unitaire et commutative) pour l'addition et la multiplication des fonctions.

2.3. Soient $\mathscr{E}$ et $\mathscr{F}$ deux $\mathbb{K}$-espaces vectoriels : dans la suite, nous noterons $L(\mathscr{E}, \mathscr{F})$ le $\mathbb{K}$-espace vectoriel des homomorphismes $\mathbb{K}$-linéaires de $\mathscr{E}$ dans $\mathscr{F}$; quand $\mathscr{E} = \mathscr{F}$ nous notons simplement $L(\mathscr{E})$ le $\mathbb{K}$-espace vectoriel des endomorphismes de $\mathscr{E}$. Lorsque $\mathscr{E}$ et $\mathscr{F}$ sont tous deux munis d'une norme, alors $\mathcal{L}(\mathscr{E}, \mathscr{F})$ (resp. $\mathcal{L}(\mathscr{E})$) désignera le sous-espace de $L(\mathscr{E}, \mathscr{F})$ (resp. $L(\mathscr{E})$) des homomorphismes (resp. endomorphismes) de $\mathscr{E}$ dans $\mathscr{F}$ (resp. de $\mathscr{E}$). Si l'axiome du choix fait partie de nos outils mathématiques, alors $\mathcal{L}(\mathscr{E}, \mathscr{F})$ est toujours un sous-espace strict de $L(\mathscr{E}, \mathscr{F})$, dès que $\mathscr{E}$ est de dimension infinie. Il est facile de vérifier que tout endomorphisme linéaire $X \mapsto AX$ de $\mathscr{E}$ est continu (i.e. $A \in \mathcal{L}(\mathscr{E})$) ssi il est borné (sur la sphère unité), c'est-à-dire si le supremum $\|A\|$ des $\|AX\|$, pour $\|X\| = 1$ est fini. L'application $A \mapsto \|A\|$ est une norme sur $\mathcal{L}(\mathscr{E})$ appelée *« norme subordonnée »* (à la norme sur $\mathscr{E}$), ou encore *« la norme opérateur »* : celle-ci possède la propriété importante d'être sous-multiplicative, en ce sens que pour tout[6] $A, B \in \mathcal{L}(\mathscr{E})$

$$\|AB\| \leq \|A\|\,\|B\| \tag{1}$$

2.4. Considérons maintenant que $(\mathscr{E}, \|\cdot\|)$ est un espace de Banach ; alors l'espace normé $(\mathcal{L}(\mathscr{E}), \|\cdot\|)$ est lui aussi un espace de Banach : l'inégalité (1) vérifiée par la norme subordonnée fait alors de $\mathcal{L}(\mathscr{E})$ une algèbre de Banach unitaire, où l'unité est l'identité de $\mathscr{E}$ que nous notons I. (Lorsque $\mathscr{E}$ est dimension au moins égale à 2, l'algèbre $\mathcal{L}(\mathscr{E})$ est non commutative.) Cette structure d'algèbre de Banach unitaire permet en particulier de définir l'image d'un endomorphisme A par une application analytique (comme généralisation des valeurs polynomiales sur A). En effet si $F(z) := \sum_{k=0}^{\infty} a_k z^k$ est une série entière de la variable complexe z (avec $a_k \in \mathbb{K}$) de rayon de convergence $0 < \rho \leq +\infty$ (i.e. $F(z)$ est holomorphe sur l'ensemble des $z \in \mathbb{C}$ t.q. $|z| < r$ ssi $r < \rho$), alors pour tout $A \in \mathcal{L}(\mathscr{E})$ t.q. $\|A\| < \rho$ (et en utilisant la convention que $A^0 = I$)

$$F(A) := \sum_{k=0}^{\infty} a_k A^k$$

5. Où nous notons AB le produite de deux éléments de $\mathscr{A}$.
6. Nous notons AB la composition $A \circ B$.

est un élément de $\mathcal{L}(\mathscr{E})$. Supposons par exemple que $\|A\| < 1$: alors l'endomorphismes $I - A$ est inversible et par *« la formule d'inversion de von Neumann »* :

$$(I - A)^{-1} = \textstyle\sum_{k=0}^{\infty} A^k$$

Un autre exemple important est celui de l'exponentielle : pour tout $A \in \mathcal{L}(\mathscr{E})$, nous pouvons définir *« l'exponentielle »* de A, soit

$$\exp(A) := \textstyle\sum_{k=0}^{\infty} 1/k!A^k$$

avec la propriété importante que pour tout $A, B \in \mathcal{L}(\mathscr{E})$

$$AB = BA \implies \exp(A + B) = \exp(A)\exp(B)$$

(en particulier $\exp(A)$ est toujours inversible avec $\exp(A)^{-1} = \exp(-A)$).

2.5. L'exponentielle permet de résoudre les *« problèmes de Cauchy linéaires »* du premier ordre sur $\mathscr{E}$, du type $dY/dt = AY(t)$ avec la condition initiale $Y(0) = Y_0 \in \mathscr{E}$ et où $A \in \mathcal{L}(\mathscr{E})$. En effet, pour tout $t, h \in \mathbb{R}$, du fait que tA commute avec hA, nous avons

$$\exp\big((t+h)A\big) = \exp(tA)\exp(hA) = \exp(tA)\big(I + hA + \cdots\big) = \exp(tA) + tA\exp(tA) + \cdots$$

de sorte que $d/dt \exp(tA) = A\exp(tA)$. Alors nous avons

$$\frac{dY}{dt} = AY(t) \iff \exp(-tA)\left(\frac{dY}{dt} - AY(t)\right) = 0 \iff \frac{d}{dt}\big(\exp(-tA)Y(t)\big) = 0$$

ce qui (avec la condition $Y(0) = Y_0$) donne finalement la solution

$$Y(t) = \exp(tA)Y_0$$

Remarque 2.2. *Lorsque l'espace de Banach $\mathscr{E}$ est de dimension finie, tout endomorphisme est continu, de sorte que $L(\mathscr{E}) = \mathcal{L}(\mathscr{E})$; quand $\mathscr{E}$ est de dimension infinie, l'axiome du choix permet d'exhiber facilement des endomorphismes non continus, faisant de $\mathcal{L}(\mathscr{E})$ un sous-espace strict de $L(\mathscr{E})$. Un théorème difficile*[7] *montre qu'en un sens, l'axiome du choix est la condition d'existence des endomorphismes non bornés dans les espaces de Banach de dimension infinie. Les « opérateurs non bornés » de la mécanique quantique (c.f.* [Oli16a]*) sont des applications linéaires définies sur un sous-espace dense de l'espace total – en l'occurence un espace de Hilbert $\mathscr{H}$ – à valeurs dans $\mathscr{H}$ mais ne possédant pas de prolongement borné à tout $\mathscr{H}$. Nous verrons (c.f.* [Oli16a]*) comment il est possible de résoudre (dans certains cas) les problèmes de Cauchy du type*

$$\frac{dY}{dt} = AY(t) \quad \textit{avec} \quad Y(0) = Y_0 \in \mathscr{D} \tag{2}$$

lorsque $(A, \mathscr{D})$ est un opérateur non borné défini sur un sous-espace $\mathscr{D}$ dense dans $\mathscr{H}$. « L'équation de Schrödinger » correspond à un problème de Cauchy du type (2) dans le cas où $(A, \mathscr{D})$ est autoadjoint : c'est « le théorème de Stone » (prototype du théorème de « Hille-Yosida ») qui permet la résolution abstraite de l'équation de Schrödinger.

7. Si $\mathscr{E}$ est un espace de Banach de dimension infinie, alors il existe une suite $e_0, e_1, \ldots$ formant une famille dénombrable et libre de vecteurs unitaires : grâce à l'axiome du choix cette famille peut être complétée en une famille $\{e_i\ ;\ I\}$ (avec $I \supset \mathbb{N}$) formant une base (algébrique) de $\mathscr{E}$: si nous posons $f(e_n) = ne_n$, pour tout $n \in \mathbb{N}$ et $f(e_i) = 0$ pour tout $i \in I \setminus \mathbb{N}$, alors f est un endomorphisme non borné. En un sens, c'est le seul moyen d'obtenir un endomorphisme non borné sur un espace de Banach : ce résultat est dû à Solovay.

2.6. Dans cette série de notes, nous supposons connue la *« théorie de la mesure »* et de *« l'intégrale de Lebesgue »* (des présentations simples et complètes sont données dans [ST89, GW90] – pour une présentation plus détaillée c.f. e.g. [KF94]). Nous utiliserons sans plus de précisions, les notions de *« fonction borélienne »* et *« d'égalité presque partout »*, les théorèmes de la *« convergence monotone »* et de la *« convergence dominée »*, le *« théorème de dérivation sous le signe intégral »*, le *« théorème de dérivation de Lebesgue »*, le *« théorème de Radon-Nikodym »* et enfin le *« théorème de Fubini »*. En particulier nous noterons dans la suite $\lambda(dx) = dx$ la mesure de Lebesgue sur $\mathbb{R}$ (qui est caractérisée par le fait que $\lambda([a\,;b]) = b - a$, pour tout $-\infty < a \leq b < +\infty$). Afin d'aborder les bases mathématiques de la mécanique quantique, nous aurons justement besoin d'un certain nombres d'espaces de Banach très importants et dont la définition dépend de l'intégrale de Lebesgue sur[8] $\mathbb{R}$. Soit I un intervalle fermé de la droite réelle et $p \geq 1$ un réel donné ; nous allons montrer (c.f. Corollaire 2.8 infra) que l'ensemble des fonctions (classes de fonctions boréliennes) $f : I \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) telle que f^p soit Lebesgue intégrable (sur l'intervalle I) est bien un espace vectoriel : cet espace est classiquement noté $L^p(I)$ (avec le cas spécial de $L^\infty(I)$ qui coïncide par convention avec l'espace des fonctions qui sont bornées presque partout sur I). Chaque espace $L^p(I)$ est associé à une norme qui en fait un espace de Banach (théorème de Riesz-Fisher : c.f. Théorème 2.9 infra). Pour voir tout cela, considérons que $f : I \to \mathbb{K}$ est une fonction borélienne ; si d'une part $f \in L^\infty(I)$, alors nous noterons $\|f\|_\infty$ l'infimum des $M > 0$ t.q. $|f(x)| \leq M$ pour presque tout $x \in I$ (par convention $\|f\|_\infty = +\infty$ si f n'est pas dans $L^\infty(I)$) ; si d'autre part $1 \leq p < +\infty$ alors nous posons par définition

$$\|f\|_p := \left(\int_I |f(x)|^p dx\right)^{1/p}$$

Définition 2.3. *Soient $0 \leq p \leq +\infty$ et I un intervalle fermé de $\mathbb{R}$; alors la fonction borélienne $f : I \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) appartient à $L^p(I)$ ssi $\|f\|_p < +\infty$.*

La proposition suivante est *« l'inégalité de Hölder »*.

Proposition 2.4 (Hölder)**.** *Soient $1 \leq p, q \leq +\infty$ t.q. $1/p + 1/q = 1$ (p et q sont dits conjugués) et I un intervalle fermé de $\mathbb{R}$; si $f, g : I \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) sont deux fonctions boréliennes, alors :*

$$\|fg\|_1 = \int_I |f(x)g(x)|dx \leq \left(\int_I |f(x)|^p dx\right)^{1/p} \left(\int_I |g(x)|^q dx\right)^{1/q} = \|f\|_p \cdot \|g\|_q$$

de plus $\|fg\|_1 = \|f\|_p \, \|g\|_q$ si et seulement si $|f|^p = |g|^q$ presque partout.

Lemme 2.5 (Inégalité de Young)**.** *Pour tous réels $a, b \geq 0$ et tous réels $p, q \geq 1$ t.q. $1/p + 1/q = 1$:*

$$ab \leq a^p/p + b^q/q$$

De plus, l'égalité a lieu si et seulement si $a^p = b^q$.

Preuve. Avec $a^p = \exp(x)$ et $b^q = \exp(y)$ la convexité de la fonction exponentielle donne

$$ab = \exp\left(\frac{x}{p} + \frac{y}{q}\right) \leq \frac{1}{p}\exp(x) + \frac{1}{q}\exp(y) = \frac{a^p}{p} + \frac{b^q}{q}$$

Le cas d'égalité est une conséquence de la stricte convexité de l'exponentielle.

□

8. Ou par extension sur $\mathbb{R}^n$.

Preuve de la Proposition 2.4. Dans le cas où $0 < p, q < +\infty$ avec $0 < \|f\|_p, \|g\|_q < +\infty$ (les autres cas étant laissés au soin du lecteur), l'inégalité de Young entraîne que :

$$\begin{aligned} \|fg\|_1 &= \|f\|_p \cdot \|g\|_q \left(\int_I \frac{|f(x)|}{\|f\|_p} \cdot \frac{|g(x)|}{\|g\|_q} dx \right) \\ &\leq \|f\|_p \cdot \|g\|_q \left(\frac{1}{p} \int_I \left(\frac{|f(x)|}{\|f\|_p} \right)^p dx + \frac{1}{q} \int_I \left(\frac{|g(x)|}{\|g\|_q} \right)^q dx \right) = \|f\|_p \cdot \|g\|_q \end{aligned}$$

Le cas d'égalité découle du cas d'égalité de l'inégalité de Young.

□

L'inégalité de Hölder généralisée suivante s'obtient par récurrence.

Corollaire 2.6. *Pour $n \geq 2$ et $1 \leq p_1, \ldots, p_n \leq +\infty$ t.q. $1/p_1 + \cdots + 1/p_n = 1$ et $f_1, \ldots, f_n$ des fonctions (complexes) boréliennes sur $\mathbb{R}$:*

$$\|f_1 \cdots f_n\|_1 \leq \|f_1\|_{p_1} \cdots \|f_n\|_{p_n}$$

Un corollaire important de l'inégalité de Hölder est *« l'inégalité de Minkowski »*.

Corollaire 2.7 (Minkowski)**.** *Si $1 \leq p \leq +\infty$ et si $f, g \in L^p(I)$, alors $f + g \in L^p(I)$ et :*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

Preuve. Nous considérons que $1 < p < +\infty$ (les deux autres cas étant laissés au soin du lecteur). Supposons donc que f et g sont deux fonctions dans $L^p(I)$, i.e. $\|f\|_p, \|g\|_p < +\infty$; alors en utilisant le fait que pour tout $a, b \in \mathbb{R}$

$$|a + b|^p \leq (|a| + |b|)^p \leq 2^p \max\{|a|, |b|\}^p \leq 2^p(|a|^p + |b|^p)$$

nous pouvons intégrer l'inégalité $|f(x) + g(x)|^p \leq 2^p(|f(x)|^p + |g(x)|^p)$ et obtenir que

$$\int_I |f(x) + g(x)|^p dx \leq 2^p \left(\int_I |f(x)|^p dx + \int_I |g(x)|^p dx \right)$$

Cela assure que $f + g$ est bien dans $L^p(I)$, c'est-à-dire que $\|f + g\|_p < +\infty$. Pour l'inégalité de Minkowski, elle est triviale lorsque $\|f + g\|_p = 0$. Lorsque $\|f + g\|_p > 0$, nous notons A l'ensemble des $x \in I$ t.q. $|f(x) + g(x)| > 0$, de sorte que $\|f + g\|_p^p = \int_A |f(x) + g(x)|^p dx$. Grâce à l'inégalité de Hölder (avec $q = p/(p - 1)$) nous pouvons écrire :

$$\begin{aligned} \|f + g\|_p^p &\leq \int_A |f(x)| \cdot |f(x) + g(x)|^{p-1} \mathrm{d}x + \int_A |g(x)| \cdot |f(x) + g(x)|^{p-1} dx \\ &\leq \left(\int_I |f(x)|^p \mathrm{d}x \right)^{1/p} \left(\int_I |f(x) + g(x)|^{(p-1)q} dx \right)^{1/q} \\ &\quad + \left(\int_I |g(x)|^p \mathrm{d}x \right)^{1/p} \left(\int_I |f(x) + g(x)|^{(p-1)q} dx \right)^{1/q} \end{aligned}$$

et donc (avec $(p - 1)q = p$ et $1/q = 1 - 1/p$) il vient : $\|f + g\|_p^p \leq (\|f\|_p + \|g\|_p) \cdot \|f + g\|_p^{p-1}$; du fait que $0 < \|f + g\|_p < +\infty$, nous pouvons conclure que $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

□

Corollaire 2.8. *Pour tout $1 \leq p \leq +\infty$ l'ensemble $L^p(I)$ est un espace vectoriel et l'application $f \mapsto \|f\|_p$ est une norme sur $L^p(I)$.*

Le Théorème de Riesz-Fisher ci-dessous est un beau résultat de la théorie de la mesure.

Théorème 2.9. *Soit I un intervalle fermé de la droite réelle et $p \geq 1$ un réel donné ; alors l'espace $L^p(I)$ des fonctions réelles (ou complexes) f définies et boréliennes sur I, telles que $f^p(x)$ soit Lebesgue-intégrable (sur I) est un espace de Banach pour la norme $\|\cdot\|_p$ telle que*

$$\|f\|_p = \Big(\int_I |f(x)|^p dx\Big)^{1/p}$$

Preuve. D'après le Lemme 2.1, il suffit de montrer que si $\sum_{n=0}^{\infty} f_k$ est une série absolument convergente de $L^p(I)$ (i.e. $f_k \in L^p(I)$ et $\sum_{k=0}^{\infty} \|f_k\|_p < \infty$), alors les sommes partielles $F_n := \sum_{k=0}^{n} f_k$ forment une suite qui converge dans $L^p(I)$ vers une fonction F de $L^p(I)$. Pour voir cela nous pouvons définir pour tout $x \in I$,

$$G(x) := \sum_{k=0}^{\infty} |f_k(x)| \quad (\in [0\,;+\infty])$$

de sorte que $x \mapsto G(x)$ est une fonction borélienne. D'après l'inégalité de Minkowski (inégalité triangulaire pour la norme $\|\cdot\|_p$) nous avons

$$\int G^p(x)dx = \int_I \Big(\sum_{k=0}^{\infty} |f_k(x)|\Big)^p dx = \Big\|\sum_{k=0}^{\infty} |f_k|\Big\|_p^p \leq \Big(\sum_{k=0}^{\infty} \|f_k\|_p\Big)^p < +\infty$$

ce qui signifie que $G \in L^p(I)$. En particulier la valeur de $G(x)$ doit être finie pour presque tout x ; par suite $\sum_{k=0}^{\infty} f_k(x)$ est une série numérique (à valeurs réelles ou complexes) absolument convergente et donc sommable pour presque tout x : il existe donc une fonction F, borélienne sur I et telle que pour presque tout $x \in I$

$$F(x) := \lim_{n\to+\infty} F_n(x) = \sum_{k=0}^{\infty} f_k(x)$$

D'une part, il est facile de vérifier que F est est une fonction de $L^p(I)$ puisque

$$\int_I |F(x)|^p dx \leq \int_I \Big(\sum_{k=0}^{\infty} |f_k(x)|\Big)^p dx = \int_I G^p(x)dx < +\infty$$

Mais d'autre part, les sommes partielles $F_n := \sum_{k=0}^{n} f_k$ sont aussi des fonctions de $L^p(I)$ et nous avons

$$\int_I \Big|F(x) - F_n(x)\Big|^p dx \leq \int_I \Big(\sum_{k=n+1}^{\infty} |f_k(x)|\Big)^p dx$$

Comme $\big(\sum_{k=n+1}^{\infty} |f_k(x)|\big)^p$ est dominée par la fonction $G^p(x)$ qui est Lebesgue intégrable et que (pour presque tout x) $\big(\sum_{k=n+1}^{\infty} |f_k(x)|\big)^p \to 0$ quand $n \to +\infty$, nous pouvons conclure (convergence dominée) que les F_n tendent vers F dans $L^p(I)$.

□

Il y a plusieurs versions du théorème dit de Riesz-Fisher (le Théorème 5.8 ci-dessous en est une version élémentaire[9]). On trouve plusieurs énoncés de ce théorème dans [RSN55, p. 58 & 69-70], dont la formulation primitive due à Riesz : c.f. [RSN55, p. 70]) : l'énoncé de la version primitive utilise implicitement la caractérisation des espaces de Banach donnée par le Lemme 2.1. Nous pointons qu'une version plus complète affirme en plus que si les f_n tendent vers f dans $L^p(I)$, alors il existe une sous-suite $f_{r_0}, f_{r_1}, \ldots$ t.q. $f_{r_0}(x), f_{r_1}(x), \ldots$ qui tend vers $f(x)$ pour presque tout $x \in \mathbb{R}$. Ce résultat est par exemple implicite dans la démonstration[10] donnée dans [RSN55, p. 58] : nous allons démontrer cette dernière proposition à partir de la complétude de l'espace $L^p(I)$.

9. Les démonstrations des Théorèmes 2.9 & 5.8 sont d'ailleurs calquées l'une sur l'autre.
10. Cette démonstration n'utilise pas le Lemme 2.1.

Proposition 2.10. *Soit I un intervalle fermé de $\mathbb{R}$ et $1 \leq p \leq +\infty$; si $f_0, f_1, \dots$ est une suite de $L^p(I)$ convergeant vers f dans $L^p(I)$, alors il existe une suite strictement croissante $r_0, r_1, \dots$ formée d'entiers pour laquelle $f_{r_0}(x), f_{r_1}(x), \dots$ tend vers $f(x)$ pour presque tout $x \in \mathbb{R}$.*

Preuve. Soit $f_0, f_1, \dots$ une suite de $L^p(I)$ convergeant vers f dans $L^p(I)$: alors il existe une sous-suite $f_{r_0}, f_{r_1}, \dots$ t.q. $\|f_{r_k} - f_{r_{k-1}}\|_p \leq 1/2^k$. Si nous posons

$$g_n := \sum_{k=1}^{n} |f_{r_k} - f_{r_{k-1}}|$$

alors il est immédiat de vérifier que les g_n forment une suite de Cauchy dans $L^p(I)$; par suite (Théorème de Riesz-Fisher) la suite des g_n converge vers la fonction g_* de $L^p(I)$ t.q $g_*(x) = \sup_n\{g(x)\}$. Mais g_* étant dans $L^p(I)$, il est nécessaire que $g_*(x) < +\infty$ pour presque tout $x \in I$; en particulier, cela signifie que la suite (croissante) $g_n(x)$ est de Cauchy (i.e. bornée), pour presque tout $x \in \mathbb{R}$. Or nous avons

$$\begin{aligned} \left|f_{r_{n+m}}(x) - f_{r_n}(x)\right| &= \left|\left(-f_{n_0}(x) + \sum_{k=1}^{n+m} f_{r_k}(x) - f_{r_{k-1}}(x)\right) - \left(-f_{n_0}(x) + \sum_{k=1}^{n} f_{r_k}(x) - f_{r_{k-1}}(x)\right)\right| \\ &\leq \sum_{k=1}^{n+m} |f_{r_k}(x) - f_{r_{k-1}}(x)| - \sum_{k=1}^{n} |f_{r_k}(x) - f_{r_{k-1}}(x)| \\ &= g_{n+m}(x) - g_n(x) \end{aligned}$$

ce qui signifie que $f_{r_n}(x)$ est une suite de Cauchy pour presque tout $x \in I$.

□

3. Convolutions

3.1. La convolution est un outil très puissant d'analyse, en particulier à cause de ses propriétés de régularisation et d'approximation. C'est aussi une opération essentielle à la théorie du signal : la transmission de signaux à travers des systèmes linéaires se traduit par l'action d'opérateurs de convolution [11]. Au lieu d'une présentation systématique, nous donnons quelques exemples d'applications qui couvrirons nos besoins ultérieurs. Pour commencer, nous dirons que deux fonctions réelles f et g définies et boréliennes sur $\mathbb{R}$ sont *« convolables »* ssi $y \mapsto f(x-y)g(y)$ est intégrable pour presque partout $x \in \mathbb{R}$; alors la convolution $f * g$ est la fonction définie pour (presque) tout $x \in \mathbb{R}$ et telle que

$$f * g(x) = \int f(x-y)g(y)dy$$

Par exemple, si f et g sont des fonctions appartenant respectivement aux espaces $L^1(\mathbb{R})$ et $L^\infty(\mathbb{R})$, alors elles sont convolables : de plus, il est immédiat que $f * g(x)$ est bien définie pour tout $x \in \mathbb{R}$ et appartient à $L^\infty(\mathbb{R})$: nous allons voir comment la convolution entraîne ici un effet de régularisation.

Proposition 3.1. *Si f est une fonction de $L^\infty(\mathbb{R})$, alors pour toute fonction $U \in L^1(\mathbb{R})$, la fonction $U * f = f * U$ est un élément uniformément continu de $L^\infty(\mathbb{R})$.*

Preuve. Par les propriétés élémentaires de l'intégrale de Lebesgue (formule de changement de variable linéaire), il est immédiat que $U * f(x) = f * U(x)$ est bien défini pour tout x et que la fonction $U * f = f * U$ est bornée sur $\mathbb{R}$. Pour démontrer l'uniforme continuité

11. Nous n'abordons pas ici ce point très important.

de $U * f$, nous commençons par supposer que U est une fonction de $L^1(\mathbb{R})$ qui est uniformément continue sur $\mathbb{R}$. Pour $\varepsilon > 0$ arbitraire, nous notons $\eta_U(\varepsilon)$ le module de continuité de U pour la jauge ε (i.e. $\eta_U(\varepsilon) := \max\{|U(x) - U(y)|\ ;\ |x - y| \leq \varepsilon\}$), de sorte que :

$$\big|U * f(x) - U * f(x+\varepsilon)\big| = \left|\int \big(U(x-t) - U(x+\varepsilon-t)\big) f(t)dt\right| \leq \eta_U(\varepsilon)\|f\|_\infty$$

Cela assure l'uniforme continuité de $U * f$. Nous utilisons maintenant la densité [12] dans $L^1(\mathbb{R})$ de l'espace $\mathscr{C}_C(\mathbb{R})$ de fonctions réelles continues sur $\mathbb{R}$ et à support compact : ainsi, pour $U \in L^1(\mathbb{R})$ arbitrairement donnée, nous pouvons considérer une suite $U_1, U_2, \ldots$ de fonctions dans $\mathscr{C}_C(\mathbb{R})$ t.q. $\|U - U_k\|_1 \to 0$ quand $k \to +\infty$. Pour tout $x \in \mathbb{R}$, nous avons alors

$$\begin{aligned} \big|U * f(x) - U_k * f(x)\big| &= \left|\int \big(U(x-t) - U_k(x-t)\big) f(t)dt\right| \\ &\leq \|f\|_\infty \int \big|U(x-t) - U_k(x-t)\big|dt = \|f\|_\infty \|U - U_k\|_1 \end{aligned}$$

Les U_k étant étant uniformément continues sur $\mathbb{R}$, nous savons d'après la première partie de la démonstrations, que chacune des $U_k * f$ est uniformément continue sur $\mathbb{R}$: l'inégalité $\|U * f - U_k * f\|_\infty \leq \|f\|_\infty \|U - U_k\|_1$ assure alors que la fonction $U * f$ est elle aussi uniformément continue sur $\mathbb{R}$

□

Les propriétés d'approximation liées à la convolution sont basées sur la notion de *« noyau de sommabilité »* que nous définissons maintenant.

Définition 3.2. *Une suite $(U_0, U_1, \ldots)$ de fonctions de $L^1(\mathbb{R})$ est appelée un noyau de sommabilité lorsque les trois propriétés suivantes sont vérifiées, soient :*

(i) : $\int U_n(x)dx = 1$;
(ii) : $\sup_n\{\|U_n\|_1\} < +\infty$;
(iii) : pour tout $\varepsilon > 0$, $\int_{|x|\geq\varepsilon} |U_n(x)|dx = 0 \to 0$ quand $\varepsilon \to 0$.

(dans le cas où les fonctions U_k sont positives où nulles, la condition (ii) est redondante).

La proposition suivante affirme que toute fonction uniformément continue $f \in L^\infty(\mathbb{R})$ peut être uniformément approchée par une suite de fonctions uniformément continues de la forme $U_k * f$, où $(U_1, U_2, \ldots)$ est un noyau de sommabilité.

Proposition 3.3. *Soit $f : \mathbb{R} \to \mathbb{R}$ uniformément continue et bornée sur $\mathbb{R}$: si $(U_1, U_2, \ldots)$ est un noyau de sommabilité, alors $U_n * f$ converge vers f uniformément sur $\mathbb{R}$ quand $n \to +\infty$.*

Preuve. Soit $\varepsilon > 0$ arbitrairement donné : alors par uniforme continuité de f, il existe $\eta > 0$ t.q. $|f(x) - f(y)| \leq \varepsilon/2$ dés que $|x - y| \leq \eta$. Pour tout $x \in \mathbb{R}$ nous avons

$$\begin{aligned} \big|U_n * f(x) - f(x)\big| &= \left|\int U_n(t)\big(f(x-t) - f(x)\big)dt\right| \\ &\leq \int_{|t|\leq\eta} U_n(t)\big|f(x-t) - f(x)\big|dt + \int_{|t|>\eta} U_n(t)\big|f(x-t) - f(x)\big|dt \end{aligned}$$

Par définition du module de continuité η et en utilisant le fait que f est bornée, il vient

$$\big|U_n * f(x) - f(x)\big| \leq \frac{\varepsilon}{2} \int U_n(t)dt + 2\|f\|_\infty \int_{|t|>\eta} U_n(t)dt$$

12. Résultat de la théorie de la mesure que nous admettons ici.

La conclusion découle de l'existence d'un N t.q. , pour tout $n \geq N$

$$\int_{|t|>\eta} U_n(t)dt \leq \frac{\varepsilon}{4\|f\|_\infty}$$

□

Remarque 3.4. *L'intérêt de la Proposition 3.3 peut sembler (à priori) très discutable du fait (c.f. Proposition 3.1) que nous approximons une fonction bornée et uniformément continue sur $\mathbb{R}$ (soit f) par des fonctions bornées et uniformément continues sur $\mathbb{R}$ (soient $U_k * f$, pour $k = 1, 2, \dots$) ! Le point important (et que nous utiliserons par exemple pour démontrer le théorème d'approximation de Weierstrass : c.f. § 4.1) est que les fonctions $U_k * f$ ont la même régularité que les fonctions U_k (l'idée générale est que la convolée de deux fonctions possède la régularité de la plus régulière des deux fonctions). Pour illustrer cela, notons par exemple que si U est une une fonction dans l'espace $\mathscr{C}^n_c(\mathbb{R})$ des fonctions n $(= 1, 2, \dots, +\infty)$ fois continûment dérivables sur $\mathbb{R}$ et à support compact, alors (c.f.* [GW90, § 21.2]*) pour tout $f \in L^1(\mathbb{R})$ la fonction $U * f$ est un élément de $L^1(\mathbb{R}) \cap \mathscr{C}^n(\mathbb{R})$ et que de plus, pour tout $1 \leq k \leq n$:*

$$\frac{d^k}{dx^k} U * f = \left(\frac{d^k}{dx^k} U\right) * f$$

3.2. Considérons maintenant que f et g sont toutes deux des fonctions de $L^1(\mathbb{R})$; alors d'après le théorème de Fubini-Tonelli (et par changement de variable) nous avons

$$\iint |f(x-y)g(y)|dxdy = \int \left(\int |f(x-y)|dx\right)|g(y)|dy = \|f\|_1 \, \|g\|_1$$

de sorte que f et g sont convolables (i.e. $y \mapsto |f(x-y)g(y)|$ est dans $L^1(\mathbb{R})$ pour presque tout $x \in \mathbb{R}$) avec

$$\int |f * g(x)|dx = \int \left|\int f(x-y)g(y)dy\right| dx \leq \iint |f(x-y)g(y)|dxdy \leq \|f\|_1 \, \|g\|_1 < +\infty$$

Proposition 3.5. *Si f et g sont deux fonctions de $L^1(\mathbb{R})$ alors la convolution $f * g = g * f$ est une fonction de $L^1(\mathbb{R})$ telle que*

$$\|f * g\|_1 \leq \|f\|_1 \, \|g\|_1$$

La Proposition 3.5 possède plusieurs extensions possibles. Ainsi notons que pour $1 \leq p, q \leq +\infty$ t.q. $1/p + 1/q = 1$, l'inégalité de Hölder nous donne

$$\begin{aligned} |f| * |g|(x) &\leq \int |f|(x-y)|g|(y)dy \\ &\leq \left(\int |f|^p(x-y)dy\right)^{1/p} \left(\int |g|^q(y)dy\right)^{1/q} \leq \left(\int |f|^p(z)dz\right)^{1/p} \left(\int |g|^q(y)dy\right)^{1/q} \end{aligned}$$

Par suite, lorsque $f \in L^p(\mathbb{R})$ et $g \in L^q(\mathbb{R})$ nous avons $|f| * |g|(x) \leq \|f\|_p \cdot \|g\|_q$: en d'autres termes $f * g$ est une fonction de $L^\infty(\mathbb{R})$. Les deux résultats précédents sur l'existence de la convolution ne sont apparemment pas compatibles (puisque le conjugué de $p = 1$ est $q = +\infty$) : ils sont en fait des cas de *« l'inégalité de Young »* que nous allons démontrer.

Théorème 3.6 (Inégalité de Young). *Soit $0 \leq p, q, r \leq +\infty$ t.q.*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$$

Si $f \in L^p(\mathbb{R})$ et $g \in L^q(\mathbb{R})$, alors f et g sont convolables et de plus $f * g \in L^r(\mathbb{R})$ avec

$$\|f * g\|_r \leq \|f\|_p \, \|g\|_q$$

Preuve. Il suffit de considérer que f et g sont positives ou nulles en presque tout point. Le cas où $r = +\infty$ se traitant directement grâce à l'inégalité de Hölder classique, nous supposons que $r < +\infty$. Alors il est nécessaire que $\max\{p, q\} \leq r < +\infty$. Nous pouvons appliquer l'inégalité de Hölder à trois facteurs (c.f. Corollaire 2.6) avec le triplet d'exposants

$$(\alpha, \beta, \gamma) = \left(r, \frac{rp}{r-p}, \frac{rq}{r-q}\right)$$

En effet on vérifie immédiatement que $1/\alpha + 1/\beta + 1/\gamma = 1$ et par suite :

$$\begin{aligned} f * g(x) &= \int f^{\frac{p}{r}}(x-y) g^{\frac{q}{r}}(y) f^{1-\frac{p}{r}}(x-y) g^{1-\frac{q}{r}}(y) dy \\ &\leq \left(\int \left(f^{\frac{p}{\alpha}}(x-y) g^{\frac{q}{\alpha}}(y)\right)^{\alpha} dy\right)^{\frac{1}{\alpha}} \left(\int \left(f^{\frac{p}{\beta}}(x-y)\right)^{\beta} dy\right)^{\frac{1}{\beta}} \left(\int \left(g^{\frac{q}{\gamma}}(y)\right)^{\gamma} dy\right)^{\frac{1}{\gamma}} \\ &= \left(\int f^{p(x-y)} g^{q}(y) dy\right)^{\frac{1}{r}} \|f\|_p^{1-\frac{p}{r}} \|g\|_q^{1-\frac{q}{r}} \end{aligned}$$

En élevant à la puissance $r = \alpha$ puis en intégrant, il vient (Fubini-Tonelli)

$$\begin{aligned} \int \left(f * g(x)\right)^r dx &\leq \left(\iint f(x-y)^p g(y)^q dy dx\right) \|f\|_p^{r-p} \|g\|^{r-q} \\ &= \left(\int \left(\int f(x-y)^p dx\right) g(y)^q dy\right) \|f\|_p^{r-p} \|g\|^{r-q} \\ &= \left(\int f(z)^p dz\right) \left(\int g(y)^q dy\right) \|f\|_p^{r-p} \|g\|^{r-q} = \|f\|_p^r \|g\|^r \end{aligned}$$

□

3.3. Comme d'après le théorème de Riesz-Fisher nous savons que l'espace $(L^1(\mathbb{R}), \|\cdot\|)$ est un espace de Banach, nous pouvons déduire de la Proposition 3.5 que :

Corollaire 3.7. *$L^1(\mathbb{R})$ muni de la convolution est une $\mathbb{R}$-algèbre de Banach commutative.*

Notons au passage que $L^1(\mathbb{R})$ est un exemple d'algèbre de Banach (commutative) non unitaire (son unité devrait être l'impulsion de Dirac en 0). Il existe une notion naturelle permettant de palier l'absence d'unité d'une algèbre de Banach : ainsi, une suite $(E_0, E_1, \dots)$ d'éléments d'une algèbre de Banach $\mathscr{A}$ est appelée *« une approximation de l'unité à gauche (resp. à droite) »* lorsque pour tout $A \in \mathscr{A}$ nous avons $\|E_n A - A\| \to 0$ (resp. $\|A E_n - A\| \to 0$) quand $n \to +\infty$. Le théorème suivant est l'analogue de la Proposition 3.3 pour la convergence $L^p(\mathbb{R})$: dans le cas $p = 1$ il signifie que les noyaux de sommabilité tels que donnés dans la Définition 3.2 sont en fait des approximations de l'unité de l'algèbre de Banach $L^1(\mathbb{R})$ pour la convolution.

Théorème 3.8. *Si $(U_0, U_1, \dots)$ est un noyau de sommabilité, alors pour tout $1 \leq p < +\infty$ et $f \in L^p(\mathbb{R})$ nous avons $\|U_n * f - f\|_p \to 0$ quand $n \to +\infty$. (En particulier pour $p = 1$ un noyau de sommabilité est une approximation de l'unité de l'algèbre $L^1(\mathbb{R})$ pour la convolution.)*

Lemme 3.9. *Pour tout $a \in \mathbb{R}$, nous notons $T_a f(x) = f(x-a)$ – opérateur de translation spatiale ; si $f \in L^p(\mathbb{R})$ pour tout $1 \leq p < +\infty$ alors $\|T_a f - f\|_p \to 0$ quad $a \to 0$.*

Preuve. Par uniforme continuité, le résultat est vrai dans l'espace $\mathscr{C}_c(\mathbb{R})$ des fonctions continues sur $\mathbb{R}$ à support compact. Maintenant, considérons que $f \in L^p(\mathbb{R})$ et soit $\varepsilon > 0$ arbitrairement donné. Puisque [13] $\mathscr{C}_c(\mathbb{R})$ est dense dans $L^p(\mathbb{R})$, il existe $g \in \mathscr{C}_c(\mathbb{R})$ telle que $\|f - g\| \leq \varepsilon/4$. Or nous avons vu qu'il existe $\alpha > 0$ tel que $\|T_a g - g\|_p \leq \varepsilon/2$, pour tout $a \in \mathbb{R}$ t.q. $\|a\| \leq \alpha$; pour un tel a nous pouvons alors écrire

$$\|T_a f - f\|_p \leq \|T_a f - T_a g\|_p + \|T_a g - g\|_p + \|g - f\|_p = 2\|f - g\|_p + \|T_a g - g\|_p \leq \varepsilon$$

□

Preuve du Théorème 3.8. Soit $f \in L^p$. Comme $1/1 + 1/p = 1/p + 1$ le fait que U_n soit dans $L^1(\mathbb{R})$ assure (c.f. Théorème 3.6) que $f * U_n$ est une fonction définie presque partout et appartenant à $L^p(\mathbb{R})$. Par définition d'une approximation de l'unité nous avons $f(x) = \int f(x) U_n(y) dy$; par suite, si nous notons q le conjugué de p (i.e. $1/p + 1/q = 1$), alors pour tout rang n et presque tout x nous avons (inégalité de Hölder.)

$$\begin{aligned}
|f * U_n(x) - f(x)| &\leq \int |f(x-y) - f(x)|\,|U_n(y)| dy \\
&= \int \left(|f(x-y) - f(x)|\,|U_n(y)|^{1/p}\right) |U_n(y)|^{1/q} dy \\
&= \left(\int |f(x-y) - f(x)|^p\,|U_n(y)| dy\right)^{1/p} \left(\int |U_n(y)| dy\right)^{1/q}
\end{aligned}$$

En intégrant la puissance p-ème, le théorème de Fubini-Tonelli nous donne :

$$\begin{aligned}
\|f * U_n - f\|_p^p &\leq \left(\iint |f(x-y) - f(x)|^p\,|U_n(y)| dy dx\right) \|U_n\|_1^{p/q} \\
&\leq K^{p/q} \int \left(\int |f(x-y) - f(x)|^p dx\right) |U_n(y)| dy
\end{aligned}$$

où nous avons noté $K := \sup_n \{\|U_n\|_1\}$ (constante finie par définition d'une approximation de l'unité). En utilisant l'opérateur de translation nous avons alors :

$$\|f * U_n - f\|_p^p \leq K^{p/q} \int \|T_y f - f\|_p^p |U_n(y)| dy \tag{3}$$

Etant donné $\varepsilon > 0$, nous savons (c.f. Lemme 3.9), qu'il existe $\eta > 0$ tel que $\|T_y f - f\|_p \leq \varepsilon$ dès que $\|y\| \leq \eta$. Nous pouvons alors transformer (3) et obtenir pour tout rang n :

$$\begin{aligned}
\|f * U_n - f\|_p^p &\leq K^{p/q} \left(\int_{|y|\leq\eta} \|T_y f - f\|_p^p |U_n(y)| dy + \int_{|y|>\eta} \|T_y f - f\|_p^p |U_n(y)| dy\right) \\
&\leq K^{p/q} \left(\varepsilon^p \int_{|y|\leq\eta} |U_n(y)| dy + 2^p \|f\|_p^p \int_{|y|>\eta} |U_n(y)| dy\right) \\
&\leq K^{p/q} \left(\varepsilon^p K + 2^p \|f\|_p^p \int_{|y|>\eta} |U_n(y)| dy\right)
\end{aligned}$$

En prenant la limite-sup, quand $n \to +\infty$, la propriété (iii) d'une approximation de l'unité nous permet de conclure que $\limsup_n \|f * U_n - f\|_p^p \leq \varepsilon^p K^{p/q+1}$.

□

3.4. Nous terminons cette courte introduction aux techniques de convolution par une notion de convolution adaptée aux fonctions 2π-périodiques [14]. Deux fonctions réelles f

13. Résultat de la Théorie de la mesure que nous admettons ici.

14. Le choix de la période 2π parmi toutes les périodes possibles est arbitraire ; il est cependant motivé par le fait que la longueur du circle unité vaut 2π ; le terme circulaire est systématiquement utilisé pour préciser (souvent implicitement) que les fonctions considérées sont 2π-périodiques.

et g définies sur $\mathbb{R}$, boréliennes et 2π-périodiques sont dites *« circulairement convolables »* si la fonction $y \mapsto f(x-y)g(y)$ est localement intégrable pour presque tout $x \in \mathbb{R}$: dans ce cas *« la convolution circulaire »* $x \mapsto f \circledast g(x)$ de ces deux fonctions est par définition la fonction 2π-périodique [15]

$$(4) \qquad f \circledast g(x) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)g(t)dt = \left(\frac{f}{2\pi}\right) * \left(g \cdot \mathbf{1}_{[-\pi\,;\pi]}\right)$$

Dans ce paragraphe, nous identifions l'espace de Banach $L^1[-\pi\,;\pi]$ à l'espace des fonctions réelles définies sur $\mathbb{R}$ boréliennes, 2π-périodiques et localement intégrables. La définition de la convolution circulaire sur les différents espaces $L^p[-\pi\,;\pi]$ (avec $1 \leq p \leq +\infty$) est simplifiée par le fait que $L^p[-\pi\,;\pi]$ est un sous-espace de $L^1[-\pi\,;\pi]$. La proposition suivante est une conséquence directe de la Proposition 3.5.

Proposition 3.10. *Si f et g sont deux fonctions 2π-périodiques de $L^1[-\pi\,;\pi]$ alors la convolution circulaire $f \circledast g(x)$ de f et g est définie pour presque tout $x \in \mathbb{R}$; de plus $f \circledast g = g \circledast f$ est une fonction 2π-périodique de $L^1[-\pi\,;\pi]$.*

De manière analogue à la notion d'approximation de l'unité sur $\mathbb{R}$ (c.f. Théorème 3.8), nous avons la notion *« un noyau de sommabilité circulaire »*.

Définition 3.11. *Une suite $(U_1(x), U_2(x), \dots)$ de fonctions boréliennes supposées 2π-périodiques et localement intégrables sur $\mathbb{R}$ est appelée un noyau de sommabilité circulaire lorsque* [16]

- (i) : $\frac{1}{2\pi}\int_{-\pi}^{\pi} U_n(x)dx = 1$;
- (ii) : $\sup_n\{\|U_n\|_1\} < +\infty$;
- (iii) : *pour tout* $\varepsilon > 0$, $\int_{\varepsilon}^{2\pi-\varepsilon} |U_n(x)|dx \to 0$ *quand* $n \to +\infty$.

(dans le cas où les fonctions U_n sont positives où nulles, la condition (ii) est redondante).

Le théorème suivant est l'analogue *circulaire* du Théorème 3.8 (dans le cas $p = 1$).

Théorème 3.12. *Si $(U_0, U_1, \dots)$ est un noyau de sommabilité circulaire, alors pour tout $f \in L^1[-\pi\,;\pi]$ nous avons $\|f \circledast U_n - f\|_1 \to 0$ quand $n \to +\infty$: tout noyau de sommabilité circulaire est une approximation de l'unité de l'algèbre $L^1[-\pi\,;\pi]$ pour la convolution circulaire.*

Le proposition suivante est la version circulaire de la Proposition 3.3 pour la convolution circulaire (elle permet de démontrer facilement la version trigonométrique du théorème d'approximation de Weierstrass grâce au théorème de Féjer : c.f. § 4.4).

15. Le facteur $\frac{1}{2\pi}$ présent en (4) assure que la transformée de Fourier transforme la convolution circulaire des fonctions en produits des transformées de Fourier. Plus précisément, si nous définissons la transformée de Fourier $\mathfrak{F}[f]$ d'une fonction $f \in L^1[-\pi\,;\pi]$ comme la suite $\mathfrak{F}[f] = (c_n(f))_{n\in\mathbb{Z}}$, où

$$c_n(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-2i\pi x}dx$$

alors $\mathfrak{F}[\cdot]$ réalise une application de $L^1[-\pi\,;\pi]$ dans l'espace des suites bilatérales de nombre complexes. D'autre part, si $f, g \in L^1[-\pi\,;\pi]$, alors $f \circledast g$ est dans $L^1[-\pi\,;\pi]$ et (avec les précautions d'usages) :

$$\begin{aligned} c_n(f \circledast g) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f \circledast g(x)e^{-inx}dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)e^{-in(x-y)}dx\right)g(y)e^{-iny}dy = \frac{1}{2\pi}\int_{-\pi}^{\pi} c_n(f)g(y)e^{-iny}dy \end{aligned}$$

En d'autres termes, $c_n(f \circledast g) = c_n(f)c_n(g)$, soit encore $\mathfrak{F}[f \circledast g] = \mathfrak{F}[f] \cdot \mathfrak{F}[g]$.

16. La présence du facteur $\frac{1}{2\pi}$ dans la condition de normalisation des U_n est due à la définition de la convolution circulaire en (4).

Proposition 3.13. *Soit $f(x)$ 2π-périodique continue sur $\mathbb{R}$ et soit $(U_0(x), U_1(x), \dots)$ une approximation circulaire de l'unité : alors les fonctions $U_n \circledast f$ $(n \geq 0)$ sont des fonctions 2π-périodiques continues sur $\mathbb{R}$ qui convergent uniformément sur $\mathbb{R}$ quand $n \to +\infty$.*

4. Théorème de Stone-Weierstrass

4.1. Le fait que les espaces de Hilbert séparables possèdent tous une base hilbertienne dénombrable et un point particulièrement important : cela prouve en un sens qu'il n'existe qu'un seul espace Hilbert séparable : nous reviendrons sur ce point dans le paragraphe suivant entièrement consacré aux espaces de Hilbert (c.f. Remarque 5.9 infra). Nous commençons par traiter ici les *« théorèmes d'approximation de Weierstrass et de Stone-Weierstrass »*.

Théorème 4.1 (Weierstrass classique)**.** *Si f est une fonction réelle continue sur $[0\,;1]$ alors pour tout $\varepsilon > 0$, il existe un polynôme $P(X) \in \mathbb{R}[X]$ t.q. $|f(x) - P(x)| \leq \varepsilon$, pour tout $x \in [0\,;1]$.*

Nous allons donner plusieurs preuves de ce théorème. La première est basée sur la Proposition 3.3 qui assure l'approximation uniforme des fonctions continues sur la droite réelle par convolution avec *« un noyau de sommabilité »*.

Preuve du Théorème 4.1 (n°1). Etant donnée $f(x)$ une fonction (réelle) définie et continue sur $[0\,;1]$, nous définissons la fonction $g(x)$ telle que $g(x) = 0$ pour $x \in \mathbb{R} \setminus [0\,;1]$ et $g(x) = f(x) - A(x)$ – avec $A(x) = f(0) + \big(f(1) - f(0)\big)x$ – pour $x \in [0\,;1]$. Par définition $g(x)$ est une fonction bornée et uniformément continue sur $\mathbb{R}$. Si pour $n \geq 1$ nous posons

$$U_n(x) := \frac{\max\left\{0\,;(1-x^2)^n\right\}}{\int_{-1}^{+1}(1-y^2)^n dy}$$

alors $(U_1, U_2, \dots)$ est une approximation de l'unité et la Proposition 3.3, assure que

$$U_n * g(x) = \int U_n(y)g(x-y)dy = \int_0^1 U(x-y)g(y)dy$$

converge uniformément vers $g(x)$. Or il existe $2n+1$ fonctions polynomiales en y, soient $a_0(y), \dots, a_{2n}(y)$, telles que

$$\forall x, y \in \mathbb{R},\ x - y \in [-1\,;1] \Longrightarrow U_n(x-y) = \sum_{k=0}^{2n} a_k(y)x^k$$

de sorte que

$$U_n * g(x) = \sum_{k=0}^{2n}\left(\int_0^1 a_k(y)g(y)dy\right)x^k =: B_n(x)$$

Nous avons ainsi obtenu une suite $B_1(x), B_2(x), \dots$ de fonctions polynomiales définies pour tout $x \in [0\,;1]$ qui convergent uniformément vers $g(x)$ sur $[0\,;1]$: cela assure que $A(x) + B_n(x)$ converge uniformément vers $f(x)$ sur $[0\,;1]$ quand $n \to +\infty$.

□

4.2. Une autre démonstration constructive du théorème de Weierstrass et celle proposée par Bernstein : ainsi, toute fonction continue $f : [0\,;1] \to \mathbb{R}$ est-elle associée au polynôme

$$P_n f(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) B_{n,k}(x)$$

où les $B_{n,k}$ sont *« les polynômes de Bernstein* définis en posant

$$B_{n,k}(x) = \complement_n^k x^k (1-x)^{n-k}$$

Preuve du Théorème 4.1 (n°2). Comme f est continue sur un intervalle compact, c'est une fonction bornée et uniformément continue. Ainsi $\max\{|f(x)| \ ; \ 0 \le x \le 1\} = \|f\|_\infty < +\infty$ et pour tout $\varepsilon > 0$ arbitrairement donné, il existe $\eta > 0$ t.q. $|f(x) - f(y)| < \varepsilon/2$ dès que $|x-y| < \eta$. D'après *« la formule du binôme de Newton »* (et du fait que $0 \le x \le 1$) il est clair que $\sum_{k=0}^n B_{n,k}(x) = 1$: en fait $B_{n,k}(x)$ s'interprète comme la probability d'avoir k réalisations d'un événement au cours de n expériences indépendantes et où x est la probabilité de l'événement au cours d'une seule expérience. Dans ce modèle probabiliste, nous notons $(\Omega, \mathbb{P})$ l'espace probabilisé associé à la variable aléatoire $X_n : \Omega \to \mathbb{R}$ telle que $\mathbb{P}\{X_n = k\} = B_{n,k}(x)$ pour tout $0 \le k \le n$. En utilisant le fait que l'espérance et la variance de X_n valent respectivement nx et $nx(1-x)$ (c.f. Remarque 8.16 infra), nous pouvons déduire de *« l'inégalité de Tchebychev »* (c.f. § 8.14 infra) que

$$\sum_{|k-nx|\ge\lambda} B_{n,k}(x) = \mathbb{P}\big\{|X_n - nx| \ge \lambda\big\} \le \frac{nx(1-x)}{\lambda^2} \le \frac{n}{\lambda^2} \tag{5}$$

Maintenant nous pouvons écrire que

$$\begin{aligned} |f(x) - P_n f(x)| &= \sum_{|x-k/n|<\eta} \left| f(x) - f\left(\frac{k}{n}\right)\right| B_{n,k}(x) + \sum_{|x-k/n|\ge\eta} \left| f(x) - f\left(\frac{k}{n}\right)\right| B_{n,k}(x) \\ &\le \frac{\varepsilon}{2} + 2\|f\|_\infty \sum_{|nx-k|\ge n\eta} B_{n,k}(x) \end{aligned}$$

soit encore, d'après (5)

$$|f(x) - P_n f(x)| \le \frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{n\eta^2}$$

Lorsque $n > 4\|f\|_\infty/(\varepsilon\eta^2)$, nous obtenons bien que $|f(x) - P_n f(x)| \le \varepsilon$.

□

4.3. Un polynôme trigonométrique est une fonction complexe $T(x)$ de la variable réelle x qui peut s'écrire sous la forme exponentielle

$$T(x) = \sum_{k=-n}^{n} c_k e^{ikx}$$

où les coefficients c_k sont les nombres complexes tels que

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} T(x) e^{-ikx} dx$$

(le polynôme trigonométrique $T(x)$ est dit réel si $T(x) \in \mathbb{R}$, pour tout $x \in \mathbb{R}$).

Théorème 4.2 (Weierstrass trigonométrique). *Si f est une fonction réelle (resp. complexe) définie sur $\mathbb{R}$, continue et 2π-périodique, alors pour tout $\varepsilon > 0$, il existe un polynôme trigonométrique réel (resp. complexe) $T(x)$ telle que $|f(x) - T(x)| \le \varepsilon$, pour tout $x \in \mathbb{R}$.*

Preuve. Le cas complexe se déduit du cas réel ; il suffit donc d'établir le résultat pour $f(x)$ une fonction 2π-périodique, continue sur $\mathbb{R}$ et à valeur réelle. Pour une telle fonction nous pouvons toujours écrire $f(x) = f_P(x) + f_I(x)$ où $f_P(x)$ (resp. $f_I(x)$) est une fonction continue 2π-périodique paire (resp. impaire) : pour cela il suffit de prendre

$$f_P(x) := \frac{f(x)+f(-x)}{2} \quad \text{et} \quad f_I(x) := \frac{f(x)-f(-x)}{2}$$

Nous pouvons alors écrire $f_P(x) = F_P(\cos(x))$, où $F_P(x)$ est une fonction définie et continue sur $[-1\,;1]$ et $f_I(x) = F_I(\sin(x))$, où $F_I(x)$ est une fonction définie, continue et impaire sur $[-1\,;1]$. D'après la version classique du théorème d'approximation de Weierstrass (c.f. Théorème 4.1), pour $\varepsilon > 0$ arbitrairement donné, il existe un polynôme $A(X) = \sum_{k=0}^{a} \alpha_k X^k \in \mathbb{R}[X]$ tel que $|F_P(x) - A(x)| \le \varepsilon/2$, pour tout $-1 \le x \le +1$: par suite $|f_P(x) - A(\cos x)| \le \varepsilon/2$, pour tout $x \in \mathbb{R}$. De même, il existe un polynôme $B(X) = \sum_{k=0}^{b} \beta_k X^k \in \mathbb{R}[X]$ tel que $|f_I(x) - B(\sin x)| \le \varepsilon/2$, pour tout $x \in \mathbb{R}$. Finalement en posant $T(x) := A(\cos x) + B(\sin x)$, il vient pour tout $x \in \mathbb{R}$

$$\left|f(x) - T(x)\right| \le \left|f_P(x) - A(\cos x)\right| + \left|f_I(x) - B(\sin x)\right| \le \varepsilon$$

Or grâce aux *« formules d'Euler »* nous avons

$$T(x) = A(\cos x) + B(\sin x) = \sum_{p=0}^{a} \alpha_p \left(\frac{e^{ix}+e^{-ix}}{2}\right)^p + \sum_{p=0}^{b} \beta_p \left(\frac{e^{ix}-e^{-ix}}{2i}\right)^p$$

où nous reconnaissons un polynôme trigonométrique (nécessairement réel).

□

Nous venons de déduire la version trigonométrique du théorème d'approximation de Weierstrass à partir de sa version classique ; réciproquement il est possible de déduire la version classique de la version trigonométrique. Pour cela, supposons (pour simplifier) que $f : [-\pi\,;\pi] \to \mathbb{R}$ soit une fonction continue ; quitte à retrancher une fonction affine (i.e. polynomiale de degré 1), nous pouvons supposer que $f(-\pi) = f(\pi) = 0$ et noter $f_P : \mathbb{R} \to \mathbb{R}$ la fonction continue et 2π-périodique t.q. $f_P(x) = f(x)$ pour tout $-\pi \le x \le \pi$. D'après le théorème de Weierstrass trigonométrique, pour tout $\varepsilon > 0$ il existe alors un polynôme trigonométrique à valeurs réelles, soit $T(x) = \sum_{k=0}^{n} c_k e^{ikx}$, t.q. pour tout $x \in \mathbb{R}$.

$$|f_P(x) - T(x)| \le \varepsilon/2 \tag{6}$$

Nous utilisons ici le fait que le polynôme trigonométrique $T(x)$ peut être considéré comme la restriction à l'axe réel d'une fonction $T(z)$ de la variable complexe z, définie et analytique sur tout le plan complexe. En d'autre termes, $T(z)$ coïncide avec la somme $\sum_{k=0}^{\infty} \xi_k z^k$ d'une série entière de rayon de convergence infini : mais par définition [17] le rayon de convergence est le supremum des nombres réels R de $[0\,;+\infty]$ tels que la suite des sommes partielle $S_r(z) := \sum_{k=0}^{r} \xi_k z^k$ soient uniformément convergentes dans le disque ouvert $\{|z| < R\}$ (lorsque $p \to +\infty$). Mais le rayon de convergence de $T(z) = \sum_{k=0}^{\infty} \xi_k z^k$ étant infini, les fonctions polynomiales $S_r(z)$ convergent uniformément vers $T(z)$ sur le disque fermé $\{|z| \le \pi\}$ quand $r \to +\infty$: par suite, il existe un rang r t.q. $|T(x) - S_r(x)| \le \varepsilon/2$ pour tout $x \in [-\pi\,;\pi]$, de sorte qu'avec (6) il vient : $|f(x) - S_p(x)| = |f_P(x) - T(x)| + |T(x) - S_p(x)| \le \varepsilon$.

17. Usuellement c'est plutôt une propriété du rayon de convergence (défini par exemple par la formule d'Hadamard) : nous avons choisi de prendre cette propriété comme définition afin de simplifier notre présentation (pour une présentation classique sur les propriétés de convergence uniforme des séries entières, on pourra consulter [Gam01, Chap. V, § 3]).

4.4. Une autre manière d'aborder le théorème d'approximation de Weierstrass est de considérer directement la question de l'approximation uniforme des fonctions 2π-périodiques par des polynômes trigonométriques. Cette approche remonte (au moins) à la polémique entre D'Alembert Euler et Daniel Bernoulli sur la représentations des fonctions permettant une résolution de l'équation de la corde vibrante (voir les mémoires de D'Alembert de 1747 [D'A47b, D'A47a]), puis avec le mémoire [Fou22] publié en 1822 par Fourier sur l'équation de la chaleur (voir aussi le magnifique texte de Jean-Pierre Kahane [18] sur les séries de Fourier dans [KLR98, p. 3 - 279]). Concernant le théorème d'approximation de Weierstrass, nous nous limiterons aux deux piliers de la théorie que sont *« le théorème de Dirichlet »* et *« le théorème de Féjer »*. Pour cela, commençons par noter qu'à toute fonction complexe f localement intégrable et 2π-périodique est associée le polynôme trigonométrique

$$S_n f(x) := \sum_{k=-n}^{n} c_k(f) e^{ikx} \quad \text{où} \quad c_k(f) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) e^{-ikt}\, dt$$

Avec une petite transformation nous pouvons écrire

$$S_n f(x) := \sum_{k=-n}^{n}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}\,dt\right)e^{ikx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\left(\sum_{k=-n}^{n} e^{ik(x-t)}\right)dt$$

soit encore, (en utilisant la *« convolution circulaire »* : c.f. § 3.4) $S_n f(x) = D_n \circledast f(x)$ où

$$D_n(x) := \sum_{k=-n}^{n} e^{ikx} = \frac{\sin\left((2n+1)\frac{x}{2}\right)}{\sin\left(\frac{x}{2}\right)}$$

est le *« le noyau de Dirichlet »* (d'ordre n). Une version primitive du *« théorème de Dirichlet »* s'énonce comme suit :

Théorème 4.3 (Dirichlet). *Soit f une fonction complexe localement intégrable sur $\mathbb{R}$ et 2π-périodique ; si f est continue par morceau et de classe C^1 par morceau, alors pour tout $x \in \mathbb{R}$*

$$\lim_{n\to+\infty} S_n f(x) = \frac{1}{2}\left(f(x^+) + f(x^-)\right)$$

Preuve. Voir [GW90, Théorème 5.2.4 & 5.2.5, p. 40] ou encore [ST90, § 3.3 p. 26]

□

En un sens, *« les moyennes de Cesàro » améliorent la convergence des suites numériques*. En partant du théorème de Dirichlet nous pouvons définir les moyennes de Cesàro successives de la suite $S_0 f(x), S_1 f(x), \ldots$ en posant pour tout $n \geq 1$:

$$\sigma_n f(x) = \frac{1}{n}\sum_{k=0}^{n-1} S_k f(x) = \left(\frac{1}{n}\sum_{k=0}^{n-1} D_k\right) \circledast f(x) = F_n \circledast f(x) \tag{7}$$

où (après calcul) nous obtenons *« le noyau de Féjer »* (fonction continue et 2π-périodique) :

$$F_n(x) = \frac{1}{n}\left(\frac{\sin\left(n\frac{x}{2}\right)}{\sin\left(\frac{x}{2}\right)}\right)^2$$

18. Jean-Pierre Kahane nous a quitté en 2017 : c'est une grande peine pour tous ceux qui ont eu la chance de le croiser. Il a été le premier contact avec l'Académie des Sciences de beaucoup de jeunes mathématiciens, pour la publication de leurs résultats de thèse *aux notes CRAS*. Je me souviens de ce sentiment intimidant et stimulant que j'ai pu ressentir lors de mes échanges avec ce grand mathématicien du 20-ème siècle : ces échanges était toujours empreint d'une sorte de rigueur à visage humain que je ne pourrai pas oublier.

La grande différence entre le noyau de Dirichlet $D_n(x)$ et le noyau de Féjer $F_n(x)$ vient du fait que les $F_1(x), F_2(x), \ldots$ forment une approximation circulaire de l'unité (au sens de la Définition 3.11) : Le *« théorème de Féjer »*[19] découle de la Proposition 3.13.

Théorème 4.4 (Fejér – version uniforme)**.** *Si $f(x)$ est une fonction complexe 2π-périodique et continue sur $\mathbb{R}$, alors la série de fonctions $\sigma_n f(x)$ converge uniformément vers $f(x)$.*

Les fonctions $\sigma_n f(x)$ associées en (7) à une fonction f continue et 2π-périodique étant des polynômes trigonométriques, le théorème de Féjer établit de fait la version trigonométrique du théorème d'approximation de Weierstrass.

4.5. *« Le théorème de Stone-Weierstrass »* est une extension abstraite (et significativement plus générale) du théorème d'approximation de Weierstrass. Remarquons d'abord que l'ensemble des fonctions polynômiales (e.g. réelles) sur l'intervalle $[0\,;1]$ forme une sous-algèbre de $\mathscr{C}([0\,,1],\mathbb{R})$ (fonctions réelles) qui est unitaire (i.e. les fonctions constantes sur $[0\,;1]$ sont polynômiales) et qui séparent les points (on dit aussi séparante) en ce sens que pour $x \neq y$, il existe une fonction polynomiale $P(X) \in \mathbb{R}[X]$ t.q. $P(x) \neq P(y)$. Toutes ces remarques sont faciles à noter : il est remarquable qu'elles constituent les bonnes hypothèses d'un théorème très général[20]. Afin d'énoncer la version classique du théorème de Stone-Weierstrass, considérons que (X,d) est un espace métrique compact : alors, muni de la norme uniforme $f \mapsto \|f\|_\infty = \max\{|f(x)|\ ;\ x \in X\}$, l'espace $\mathscr{C}(X,\mathbb{K})$ des fonctions continues sur X et à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est un espace de Banach. La multiplication des fonctions est une loi de composition interne sur $\mathscr{C}(X,\mathbb{K})$ qui en fait une $\mathbb{K}$-algèbre[21]. On dira qu'une sous-algèbre $\mathscr{A}$ de $\mathscr{C}(X,\mathbb{K})$ sépare les points de X lorsque pour tout couple x, y de points distincts de X il existe $f \in \mathscr{A}$ tel que $f(x) \neq f(y)$.

Théorème 4.5 (Stone-Weierstrass)**.** *Soient (X,d) un espace métrique compact et $\mathscr{A} \subset \mathscr{C}(X,\mathbb{R})$ une sous-algèbre unitaire séparante ; alors $\mathscr{A}$ est uniformément dense dans $\mathscr{C}(X,\mathbb{R})$.*

Preuve. Voir [Rud64, § 7.24 p. 146]

□

L'énoncé du théorème de Stone-Weierstrass sous sa forme réelle donné dans le Théorème 4.5, ne peut s'étendre tel quel au cas complexe. Par exemple si nous considérons $D := \{z \in \mathbb{C}\ ;\ |z| \le 1\}$ le disque unité fermé du plan complexe et $D^\circ := \{z \in \mathbb{C} : |z| < 1\}$ son intérieur, alors l'ensemble $\mathscr{A}$ formé des polynômes complexes $P(z) = \sum_{k=0}^{n} \alpha_k z^k$ (avec $n \ge 0$ et $\alpha_k \in \mathbb{C}$) est une algèbre unitaire qui sépare les points de D. Il est cependant assez facile de vérifier que $\mathscr{A}$ n'est pas dense dans $\mathscr{C}(D,\mathbb{C})$. En effet, notons que $\mathscr{A}$ est une sous-algèbre de l'algèbre $\mathscr{H}$ formée des fonctions $f(z)$ qui sont analytiques (holomorphes) sur D° et continues sur D. Or $\mathscr{H}$ est un sous-ensemble propre de $\mathscr{C}(D,\mathbb{C})$ qui est fermé dans l'espace de Banach $(\mathscr{C}(D,\mathbb{C}), \|\cdot\|_\infty)$; en effet si $f_1(z), f_2(z), \ldots$ est une suite de fonctions de $\mathscr{H}$ qui converge uniformément vers $f(z)$ dans $\mathscr{C}(D,\mathbb{C})$, alors $f(z)$ est par hypothèse continue sur D, la convergence uniforme sur D° assurant assurant

19. Féjer était âgé de 19 ans lorsqu'il a établi ce résultat.

20. Notons que l'hypothèse de séparation des points est importante : en effet l'algèbre unitaire des fonctions polynômiales $P : [0\,;1] \to \mathbb{R}$ t.q. $P(0) = P(1)$ n'est pas dense dans $\mathscr{C}([0\,,1],\mathbb{R})$ – la clôture uniforme de ces fonctions polynomiales est l'algèbre $\mathscr{C}_P([0\,,1],\mathbb{R})$ des fonctions continues $f : [0\,;1] \to \mathbb{R}$ t.q. $f(0) = f(1)$.

21. De plus, comme $\|fg\|_\infty \le \|f\|_\infty \|g\|_\infty$, nous pouvons dire que $\mathscr{C}(X,\mathbb{K})$ est une algèbre de Banach.

aussi que $f(z)$ est analytique sur D° (c.f. [Gam01, Chap. V § 3]). La version complexe du théorème de Stone-Weirstrass est un corollaire immédiat de sa version réelle.

Corollaire 4.6 (Stone-Weierstrass complexe)**.** *Soient (X,d) un espace métrique compact et $\mathscr{A} \subset \mathscr{C}(X,\mathbb{C})$ une sous-algèbre unitaire, séparant les points de X et stable par conjugaison (i.e. $\bar{f} \in \mathscr{A}$ dès que $f \in \mathscr{A}$) ; alors $\mathscr{A}$ est uniformément dense dans $\mathscr{C}(X,\mathbb{C})$.*

A partir du théorème de Stone-Weierstrass complexe, il est possible de retrouver le théorème d'approximation de Weierstrass dans sa version trigonométrique pour les fonctions complexes continues 2π-périodiques. En effet si nous notons $\mathbb{T} = \{z \in \mathbb{C}\ ;\ |z| = 1\}$ le 1-tore complexe, alors l'algèbre des polynômes complexes sur $\mathbb{T}$ n'est pas stable par conjugaison complexe (effet $\bar{z}$ n'est pas un polynôme complexe). Par contre en notant $\mathscr{A}$ l'ensemble des fonctions de la forme $g(z) = 1/z^N P(z)$, où $P(X)$ est un polynôme de $\mathbb{C}[X]$ et où N est un entier ≥ 0, alors il est facile de vérifier que $\mathscr{A}$ est une sous-algèbre de $\mathscr{C}(\mathbb{T},\mathbb{C})$ unitaire séparante et stable par conjugaison. Par une application directe du théorème de Stone-Weierstrass complexe nous retrouvons l'analogue complexe du théorème d'Approximation de Weierstrass affirmant que si $g(z)$ est une fonction complexe définie et continue sur $\mathbb{T}$, alors pour tout $\varepsilon > 0$, il existe une polynôme $P(X) \in \mathbb{C}[X]$ et un entier $N \geq 0$ tels que $\left|g(z) - P(z)/z^N\right| \leq \varepsilon$. Le Théorème 4.2 s'en déduit, puisque toute fonction complexe $f(x)$ continue et 2π-périodique sur $\mathbb{R}$ peut s'écrire $f(x) = g(e^{ix})$, pour $g(z)$ une fonction complexe continue sur $\mathbb{T}$.

5. Espace de Hilbert

5.1. L'utilisation des espaces de Hilbert pour la formulation abstraite de la mécanique quantique s'introduit dans la deuxième moitiés des années 20 [22]. Plusieurs mathématiciens et physiciens sont impliqués dans ce travail de formalisation : mentionnons en particulier Wiener, Born, Jordan, Von Neumann et Stone. La présentation systématique/axiomatique de la mécanique quantique est exposée dans le livre de Von Neumann [VN55].

5.2. Dans la suite, nous supposons que les scalaires sont pris dans le corps $\mathbb{C}$ des nombres complexes ; *« le théorème de Fréchet-von Neumann-Jordan »* (c.f. [JvN35]) nous permet de donner un définition très générale des espaces de Hilbert comme cas particuliers des espaces de Banach. Ainsi, nous dirons que le $\mathbb{C}$-espace de Banach $(\mathscr{H}, \|\cdot\|)$ est un *« espace de Hilbert »* (complexe), lorsque *« l'identité du parallélogramme »*

$$\|X\|^2 + \|Y\|^2 = \frac{1}{2}\Big(\|X+Y\|^2 + \|X-Y\|^2\Big) \tag{8}$$

22. On lit dans [Loc94] : « *Ce sont les séries de Fourier qui sont à l'origine de l'espace de Hilbert. En effet, en étudiant les vibrations de membranes et d'enceintes quelconques, les mathématiciens mirent en évidences des vibrations propres de formes variées. Or, au début de notre siècle, surgit d'un autre horizon un nouveau venu en mathématique – l'équation intégrale de Fredholm – qui introduisait sans le dire la notion d'opérateur et exprimait sous une forme nouvelle le problème des vibrations. Grâce à cette équation, Hilbert remarqua que la recherche des vibrations d'une membrane ou d'une cavité possède une signification géométrique remarquable : elle est analogue à la recherche des axes de symétrie d'une conique (ellipse, hyperbole ou parabole).[...]*

La théorie des espaces de Hilbert, déjà très élaborée dans les années 20, était prête à accueillir en son sein la mécanique quantique, de même que les espaces de Riemann semblaient attendre la relativité générale. C'est à John von Neumann que revint le mérite de construire l'algèbre des opérateurs dans cet espace et de créer, dans un ouvrage célèbre [VN55] *la forme mathématique de la mécanique quantique.* »

est satisfaite ; alors (8) assure que l'application $(X, Y) \mapsto \langle X|Y \rangle$ définie par *« l'identité de polarisation (complexe) »*, en posant

$$\langle X|Y \rangle = \frac{1}{4}\big(\|X+Y\|^2 - \|X-Y\|^2 + \boldsymbol{i}\|X+\boldsymbol{i}Y\|^2 - \boldsymbol{i}\|X-\boldsymbol{i}Y\|^2\big) \tag{9}$$

est *« un produit scalaire hermitien (semi-linéaire à droite) sur $\mathscr{H}$ »*, la norme $\|\cdot\|$ étant liée au produit scalaire par la relation $\|X\|^2 = \langle X|X \rangle$. Dans la suite nous travaillons dans un espace de Hilbert (complexe) supposé séparable (i.e. qui possède un sous-ensemble dénombrable dense).

5.3. Partant de l'inégalité triangulaire, nous avons $(\|X\| + \|Y\|)^2 \geq \|X+Y\|^2$ et donc,

$$\|X\|^2 + \|Y\|^2 + 2\|X\|\|Y\| \geq \|X\|^2 + \|Y\|^2 + \langle X|Y \rangle + \langle Y|X \rangle$$

soit encore, après simplification :

$$\|X\|\|Y\| \geq \mathfrak{re}\,\langle X|Y \rangle \tag{10}$$

Par suite, si $\langle X|Y \rangle \neq 0$ et si θ est un argument de $\langle X|Y \rangle$, alors

$$\|X\|\|Y\| = \|X\|\|e^{\boldsymbol{i}\theta}Y\| \geq \mathfrak{re}\big(e^{-\boldsymbol{i}\theta}\langle X|Y \rangle\big) = |\langle X|Y \rangle| \tag{11}$$

Cette inégalité est connue sous plusieurs noms ; en France elle est appelée *« l'inégalité de Cauchy-Schwarz »* (Lusternik et Sobolev l'appelle *« l'inégalité de Bouniakovski-Schwarz »* : c.f. [LS89]). Il peut sembler étrange de démontrer l'inégalité de Cauchy-Schwarz à partir de l'inégalité triangulaire : de fait, nous avons choisi ici de définir un espace de Hilbert comme un espace de Banach spécial dans lequel le produit scalaire se définit à partir de la norme. Dans la pratique (et de manière plus classique : voir par exemple les §§ 5.6 & 5.7) on a souvent affaire à un espace pré-hilbertien, i.e. un espace vectoriel normé dont la norme est déduite d'un produit scalaire $\langle\cdot|\cdot\rangle$: cet espace sera hilbertien si et seulement si il est complet. Dans cette dernière présentation, les propriétés du produit scalaire permettent de déduire successivement l'inégalité de Cauchy-Schwarz puis le fait que l'application $X \mapsto \|X\| := \langle X|X \rangle^{1/2}$ est bien une norme : en particulier, l'inégalité triangulaire s'obtient à partir de l'inégalité de Cauchy-Schwarz en écrivant :

$$\begin{aligned}\|X+Y\|^2 = \|X\|^2 + \|Y\|^2 + 2\mathfrak{re}\,\langle X|Y \rangle &\leq \|X\|^2 + \|Y\|^2 + 2|\langle X|Y \rangle| \\ &\leq \|X\|^2 + \|Y\|^2 + 2\|X\|\,\|Y\| = (\|X\| + \|Y\|)^2\end{aligned}$$

On obtient ainsi un espace pré-hilbertien, dont la complétude doit être vérifiée à postériori (voir par exemple les versions du théorème de Riesz-Fisher données par les Théorèmes 2.9 & 5.8). Dans un espace de Hilbert, l'inégalité triangulaire et l'inégalité de Cauchy Schwarz, sont donc deux formes équivalentes d'une même propriété. Pour conclure ce paragraphe, nous démontrons ci-dessous la forme complète de l'inégalité de Cauchy-Schwarz à partir des propriétés du produit scalaire (et où figure le cas d'égalité).

Théorème 5.1. *Soit $\langle\cdot|\cdot\rangle$ le produit scalaire d'un espace de Hilbert $\mathscr{H}$ et $X \mapsto \|X\| := \langle X|X \rangle^{1/2}$ la norme associée ; alors l'inégalité de Cauchy-Schwarz est satisfaite, i.e. pour tout $X, Y \in \mathscr{H}$:*

$$|\langle X|Y \rangle| \leq \|X\|\,\|Y\| \tag{12}$$

De plus l'égalité en (12) a lieu ssi X et Y sont proportionnels.

Preuve. Le cas $Y = 0$ étant trivial, nous pouvons supposer que $Y \neq 0$; de plus, quitte à multiplier Y par $e^{i\theta}$, pour un réel θ adéquat, nous pouvons supposer sans perte de généralité que $\langle X|Y\rangle$ est un nombre réel. Pour tout λ réel nous posons

$$P(\lambda) = \|X + \lambda Y\|^2 = \|X\|^2 + 2\lambda\langle X|Y\rangle + \lambda^2\|Y\|^2.$$

qui (dans le cas considéré) est un polynôme du second degré en λ (car $\|Y\|^2 > 0$) et dont les coefficients sont réels (car $\langle X|Y\rangle$ est réel). Mais par définition $P(\lambda) \geq 0$, pour tout $\lambda > 0$, ce qui signifie que son discriminant est négatif ou nul, soit encore que

$$4\langle X|Y\rangle^2 - 4\|X\|^2\|Y\|^2 \leq 0,$$

L'inégalité de Cauchy-Schwarz en découle directement. Pour le cas d'égalité, il est immédiat que si X et Y sont proportionnels, alors $|\langle X|Y\rangle| = \|X\|\,\|Y\|$. Réciproquement, si $|\langle X|Y\rangle| = \|X\|\,\|Y\|$ alors le discriminant ci-dessus est nul et donc $P(X)$ admet une racine réelle (double) λ_0, ce qui signifie que $\|X + \lambda_0 Y\| = 0$.

□

Le paragraphe § 10 est consacré au principe d'incertitude de Heisenberg-Gabor, dans le cadre de la théorie des signaux d'énergie finie. Ce résultat est une forme déguisée du principe d'incertitude pour le couple d'observable position-impulsion de la mécanique ondulatoire (c.f. [Oli16b, Théorèmes 5.1 & 5.5]). Cependant, il existe un principe d'incertitude plus général que nous énoncerons dans le cadre abstrait de la mécanique ondulatoire (c.f. [Oli16b, Théorème 6.4]) qui se trouve être une conséquence directe de l'inégalité de Cauchy-Schwarz telle qu'écrite en (10).

5.4. Deux vecteurs X et Y d'un espace de Hilbert $\mathscr{H}$ sont dits orthogonaux lorsque $\langle X|Y\rangle = 0$. L'orthogonal d'un sous-ensemble $K \subset \mathscr{H}$ (non vide) est l'ensemble noté $K^\perp$ formé des $Y \in \mathscr{H}$ qui sont orthogonaux à tous les vecteurs de K, c'est-à-dire t.q. $\langle X|Y\rangle = 0$, pour tout $X \in K$.

Théorème 5.2. *L'orthogonal $K^\perp$ de tout sous-ensemble $K \subset \mathscr{H}$ (non vide) est un sous-espace vectoriel de $\mathscr{H}$ qui est fermé.*

Preuve. Il est immédiat que $K^\perp$ est un sous-espace de $\mathscr{H}$. Pour la fermeture de $K^\perp$, considérons que $X_0, X_1, \ldots$ est une suite de $K^\perp$ qui converge vers une limite X_* (dans $\mathscr{H}$). Si Y est un point arbitraire de K alors, par la continuité de la forme linéaire $\langle Y|\cdot\rangle$, nous savons que $\langle Y|X_n\rangle \to \langle Y|X_*\rangle$ quand $n \to +\infty$; or $\langle Y|X_n\rangle = 0$, pour tout n et donc $\langle Y|X_*\rangle = 0$: la conclusion que $X_* \in K^\perp$ vient du fait que Y est arbitrairement dans K.

□

Le lemme suivant est très important.

Lemme 5.3. *Si K est un sous ensemble dense de $\mathscr{H}$, alors $K^\perp = \{0\}$.*

Preuve. Soit $X \in K^\perp$ et $X_0, X_1, \ldots$ une suite de points de K t.q. $X_n \to X$ quand $n \to +\infty$; alors $\langle X_n|X\rangle = 0$ pour tout n. La conclusion vient de la continuité de l'application $Y \mapsto \langle Y|X_n\rangle$, de sorte que $0 = \langle X_n|X\rangle \to \langle X|X\rangle = \|X\|^2$ quand $n \to +\infty$.

□

5.5. Rappelons qu'une famille $\{X_i\}_{i=M}^N$ $(-\infty \leq M \leq N \leq +\infty)$ est *« orthonormée »* ssi $\langle X_i|X_j\rangle = \delta_{ij}$ (symbole de Kchronecker) et que c'est une *« base hilbertienne de $\mathscr{H}$ »*, lorsque le sous-espace vectoriel engendré par les X_n est dense dans $\mathscr{H}$. Dans la suite nous prendrons $(M,N) = (0,+\infty)$ (nous rencontrerons d'autre cas – en particulier le cas $(M,N) = (-\infty,+\infty)$ – qui se traitent de manière analogue).

Proposition 5.4. (i) : *Etant donnée $\{X_n\}_{n=0}^N$ (avec $0 \leq N \leq +\infty$) une famille orthonormée de $\mathscr{H}$, nous avons pour tout $Y \in \mathscr{H}$:*

$$\textstyle\sum_{k=0}^{\infty} |\langle Y|X_k\rangle|^2 \leq \|X\|^2 \quad \text{(« Inégalité de Bessel »)}$$

(ii) : *si $\{X_n\}_{n=0}^N$ est une base hilbertienne de $\mathscr{H}$ alors tout $Y \in \mathscr{H}$ s'écrit $Y = \sum_{k=1}^N \langle Y|X_k\rangle X_k$ i.e. :*

$$\lim_{n\to N} \left\|Y - \textstyle\sum_{k=0}^n \langle Y|X_k\rangle X_k\right\| = 0$$

de plus, $n \mapsto \langle Y|X_n\rangle$ est une suite de nombres complexes de carré sommables t.q. :

$$\textstyle\sum_{n=0}^{N} |\langle Y|X_n\rangle|^2 = \|X\|^2 \quad \text{(« Identité de Pythagore-Parseval »)}$$

Preuve. (i) : Il suffit de montrer l'inégalité de Bessel lorsque $N < +\infty$; dans ce cas, nous associons à tout $Y \in \mathscr{H}$ donné le vecteur $Y' = \sum_{k=0}^N \langle Y|X_k\rangle X_k$ de sorte que $\|Y'\|^2 = \sum_{k=0}^N |\langle Y|X_k\rangle|^2$; alors, en posant $Z = Y - Y'$, il vient, pour tout $i = 1,\ldots,N$

$$\langle Z|X_i\rangle = \langle Y|X_i\rangle - \textstyle\sum_j \langle Y|X_j\rangle \langle X_j|X_i\rangle = \langle Y|X_i\rangle - \sum_j \langle Y|X_j\rangle \delta_{ji} = 0$$

Par suite Z étant orthogonal à tous les X_i, il est aussi orthogonal à Y' et donc

$$\|Y\|^2 = \|Y'+Z\|^2 = \|Y'\|^2 + \|Z\|^2 + 2\mathfrak{re}\left\langle Y'|Z\right\rangle = \|Y'\|^2 + \|Z\|^2 \geq \|Y'\|^2$$

(ii) : Nous considérons ici que $N = +\infty$ ($N < +\infty$ étant classique). Pour Y arbitrairement pris dans $\mathscr{H}$, nous savons d'après (i) que $n \mapsto \langle Y|X_n\rangle$ est une suite de nombres complexes de carré sommable : cela entraîne que la suite $n \mapsto \sum_{k=0}^n \langle Y|X_k\rangle X_k$ est de Cauchy dans $\mathscr{H}$ (puisque $\|X_{p+q} - X_p\|^2 = \sum_{k=1}^q |\left\langle Y|X_{p+k}\right\rangle|^2$) : par suite nous pouvons considérer $Y' := \sum_{k=0}^\infty \langle Y|X_k\rangle X_k$ (i.e. la limite dans $\mathscr{H}$ de $\sum_{k=0}^n \langle Y|X_k\rangle X_k$ quand $n \to +\infty$). En posant $Z := Y - Y'$, nous pouvons montrer par un calcul analogue à celui effectué en (i) que Z est orthogonal à tous les X_n ; par suite Z est orthogonal à l'espace vectoriel engendré par les X_n : or cet espace étant dense dans $\mathscr{H}$ (définition d'une base hilbertienne), nous pouvons en conclure que $Z = 0$, i.e. $Y = Y'$. L'identité de Pythagore-Parseval découle alors de la continuité de la norme et du fait que pour tout $k \geq 0$

$$\left\|\textstyle\sum_{k=0}^n \langle Y|X_k\rangle X_k\right\|^2 = \sum_{k=0}^n |\langle Y|X_k\rangle|^2$$

□

Théorème 5.5. *Un espace de Hilbert est séparable (de dimension infinie) si et seulement si il possède une base hilbertienne dénombrable.*

Preuve. D'après la partie (ii) de la Proposition 5.4, il est immédiat que si l'espace de Hilbert $\mathscr{H}$ possède une base hilbertienne dénombrable, alors il est nécessairement séparable. Réciproquement si $\mathscr{H}$ est séparable alors il existe une famille dénombrable $\{X_0, X_1, \ldots\} \subset \mathscr{H}$ qui est libre et qui engendre un sous-espace vectoriel dense dans $\mathscr{H}$: en d'autre termes, si $\mathscr{H}_n$ est le sous-espace engendré par $\{X_0,\ldots,X_n\}$, alors $\bigcup_{n=0}^\infty \mathscr{H}_n$ est dense dans $\mathscr{H}$. En appliquant l'algorithme d'ortho-normalisation de Gramm-Schmidt, on peut construire une famille dénombrable $\{Y_0, Y_1, \ldots\}$ orthonormée de sorte que pour tout $n \geq 0$, le sous-espace $\mathscr{H}_n$ coïncide avec le sous-espace engendré par $\{Y_0, \ldots, Y_n\}$: cela entraîne que $\{Y_0, Y_1, \ldots\}$ est une base hilbertienne de $\mathscr{H}$.

□

Théorème 5.6. *Soit $\mathscr{H}$ un espace de Hilbert séparable ; si $\mathscr{K}$ est un sous-espace fermé de $\mathscr{H}$ alors $\mathscr{K}^\perp$ est un sous-espace fermé, supplémentaire de $\mathscr{K}$ en ce sens que $\mathscr{H} = \mathscr{K} \oplus \mathscr{K}^\perp$.*

Preuve. Comme $\mathscr{K} \cap \mathscr{K}^\perp = \{0\}$, il suffit de démontrer que tout $X \in \mathscr{H}$ peut s'écrire $X = Y + Z$, avec $Y \in \mathscr{K}$ et $Z \in \mathscr{K}^\perp$. Si $\mathscr{K} = \{0\}$ le résultat est immédiat ; nous supposons donc maintenant que $\mathscr{K} \neq \{0\}$. Comme $\mathscr{K}$ est un sous-espace fermé de $\mathscr{H}$ (espace de Hilbert séparable), c'est lui-même un espace de Hilbert séparable : il possède donc une base hilbertienne (finie ou dénombrable : c.f. Théorème 5.5) soit $(X_i)_{i=0}^N$ (avec $1 \leq N \leq +\infty$). D'après *« l'inégalité de Bessel »* la série $\sum_{i=0}^N |\langle X|X_i\rangle|^2$ est convergente, ce qui entraîne que $Y := \sum_{i=0}^N \langle X|X_i\rangle X_i$ est un élément bien défini de $\mathscr{K}$: or pour tout $1 \leq i_0 \leq N$

$$\langle X - Y|X_{i_0}\rangle = \langle X|X_{i_0}\rangle - \textstyle\sum_i \langle X|X_i\rangle \langle X_i|X_{i_0}\rangle = \langle X|X_{i_0}\rangle - \sum_i \langle X|X_i\rangle \delta^i_{i_0} = 0$$

ce qui entraîne que $Z = X - Y \in \mathscr{K}^\perp$.

□

Lorsque $\mathscr{K}$ est un sous-espace vectoriel fermé de $\mathscr{H}$, on dit aussi que $\mathscr{K}^\perp$ est *« le supplémentaire orthogonal de $\mathscr{K}$ »* ; nous pourrons aussi noter $\mathscr{H} = \mathscr{K} \boxplus \mathscr{G}$ une somme directe orthogonale, ce qui sous-entend que $\mathscr{G} = \mathscr{K}^\perp$.

Corollaire 5.7. *Si $\{0\} \neq \mathscr{K}$ est un sous-espace fermé de $\mathscr{H}$, alors la projection orthogonale sur $\mathscr{K}$ (i.e. la projection sur $\mathscr{K}$ parallèlement à $\mathscr{K}^\perp$) est un endomorphisme de $\mathcal{L}(\mathscr{H})$ de norme 1.*

Preuve. Si $(X_i)_{i=1}^N$ (avec $1 \leq N \leq +\infty$) est une base hilbertienne (finie ou dénombrable) $\mathscr{K}$, alors $PX = \sum_{i=1}^N \langle X|X_i\rangle X_i$ est la projection orthogonale sur $\mathscr{K}$: par suite $\|P\| = 1$, car d'une part (inégalité de Bessel) $\|PX\| \leq \|X\|$ et d'autre part $PX = X$ dès que $X \in \mathscr{K}$.

□

5.6. Les suites de scalaires pris dans le corps $\mathbb{K} = \mathbb{R}$ où $\mathbb{C}$ qui sont de carré sommable, forment un $\mathbb{K}$-espace vectoriel possédant une structure naturelle d'espace de Hilbert : cet espace représente une forme canonique des espaces de Hilbert séparables (c.f. Remarque 5.9). Pour préciser cela, considérons que $a : \mathbb{N} \to \mathbb{K}$ et $b : \mathbb{N} \to \mathbb{K}$ sont deux suites de scalaires ; alors pour tout entier $n \geq 0$, l'inégalité de Minkowski, i.e. l'inégalité triangulaire pour la norme euclidienne (resp. hermitienne) sur $\mathbb{R}^{n+1}$ (resp. $\mathbb{C}^{n+1}$), affirme que

$$\left(\sum_{k=0}^n |a(k) + b(k)|^2\right)^{1/2} \leq \left(\sum_{k=0}^n |a(k)|^2\right)^{1/2} + \left(\sum_{k=0}^n |b(k)|^2\right)^{1/2} \tag{13}$$

Nous en déduisons que l'ensemble $\ell^2 = \ell^2(\mathbb{N}, \mathbb{K})$ formé des suites $a : \mathbb{N} \to \mathbb{K}$ de nombres complexes qui sont de carré sommable (i.e. $\sum_{k=0}^\infty |a(k)|^2 < +\infty$) est un $\mathbb{K}$-espace vectoriel. De même (extension de l'inégalité de Chauchy-Schwarz sur $\mathbb{K}^{n+1}$) l'application $\langle \cdot|\cdot\rangle$ de $\ell^2 \times \ell^2 \to \mathbb{K}$ et t.q. $\langle a|b\rangle := \sum_{k=0}^\infty a(k)\overline{b(k)}$ est bien définie : elle définit un produit scalaire sur ℓ^2 associée à la norme $\|a\|_2 := \langle a|a\rangle^{1/2}$. Nous venons de vérifier que $(\ell^2, \|\cdot\|_2)$ est *« un $\mathbb{K}$-espace préhilbertien »*. Le théorème suivant est le prototype du *« théorème de Riesz-Fisher »*.

Théorème 5.8. *$\ell^2 = \ell^2(\mathbb{N}, \mathbb{K})$ est un $\mathbb{K}$-espace de Hilbert.*

Preuve. Nous devons démontrer que l'espace préhilbertien $(\ell^2, \|\cdot\|_2)$ est complet. D'après le Lemme 2.1, il s'agit de montrer que si $a_0, a_1, \ldots$ est une suite de ℓ^2 telle que la série $\sum_{q=0}^{\infty} a_q$ soit absolument convergente dans ℓ^2 (i.e. $\sum_{q=0}^{\infty} \|a_q\|_2 < \infty$), alors les sommes partielles $\alpha_n := \sum_{q=0}^{n} a_q$ forment une suite qui converge dans ℓ^2. Nous devons donc d'abord vérifier que pour tout p la série scalaire $\sum_{q=0}^{\infty} a_q(p)$ est sommable, puis en notant $\alpha(p)$ sa somme, vérifier que $p \mapsto \alpha(p)$ appartient à ℓ^2. Commençons par définir

$$A(p) := \sum_{q=0}^{\infty} |a_q(p)| \quad (\in [0\,; +\infty])$$

D'après l'inégalité de Minkowski (inégalité triangulaire pour la norme $\|\cdot\|_2$) nous avons

$$\sum_{p=0}^{\infty} A^2(p) = \lim_{n\to+\infty} \sum_{p=0}^{\infty} \left(\sum_{q=0}^{n} |a_q(p)| \right)^2 = \lim_{n\to+\infty} \left\| \sum_{q=0}^{n} |a_q| \right\|_2^2 \leq \lim_{n\to+\infty} \left(\sum_{q=0}^{n} \|a_q\|_2 \right)^2$$

soit encore

$$\sum_{p=0}^{\infty} A^2(p) \leq \left(\sum_{q=0}^{\infty} \|a_q\|_2 \right)^2 < \infty \tag{14}$$

Cela signifie que $A^2(p)$ est sommable, ou encore que $A \in \ell^2$; en particulier $A(p)$ est nécessairement fini pour tout p, ce qui assure à la série $\sum_{q=0}^{\infty} a_q(p)$ d'être absolument sommable pour tout p. Ainsi $p \mapsto \alpha(p) := \sum_{q=0}^{\infty} a_q(p)$ est une suite de scalaires bien définie : or en utilisant (14) nous avons

$$\sum_{p=0}^{\infty} |\alpha(p)|^2 = \sum_{p=0}^{\infty} \left| \sum_{q=0}^{\infty} a_q(p) \right|^2 \leq \sum_{p=0}^{\infty} \left(\sum_{q=0}^{\infty} |a_q(p)| \right)^2 = \sum_{p=0}^{\infty} A^2(p) < +\infty$$

ce qui signifie que $\alpha \in \ell^2$. Comme combinaison linéaire finie d'élément de ℓ^2, la suite $\alpha_n := \sum_{q=0}^{n} a_q$ est aussi dans ℓ^2 et de plus

$$\sum_{p=0}^{\infty} \left| \alpha(p) - \alpha_n(p) \right|^2 \leq \sum_{p=0}^{\infty} \left(\sum_{q=n+1}^{\infty} |a_q(p)| \right)^2$$

Or $p \mapsto \left(\sum_{q=n+1}^{\infty} |a_q(p)| \right)^2$ est dominée par $p \mapsto A^2(p)$ qui d'après (14) est sommable ; d'autre part, pour tout p, nous savons que $0 \leq A(p) < +\infty$ de sorte que $\sum_{q=n+1}^{\infty} |a_q(p)| \to 0$ quand $n \to +\infty$: par suite (convergence dominée) les α_n tendent vers α dans ℓ^2.

□

Remarque 5.9. *En un sens, l'espace ℓ^2 (sur $\mathbb{K} = \mathbb{R}$ où $\mathbb{C}$) est une version canonique des $\mathbb{K}$-espaces de Hilbert séparables. En effet, notons que si ϵ_n est l'élément de ℓ^2 t.q. $\epsilon_n(n) = 1$ et $\epsilon_n(k) = 0$ si $n \neq k$, alors $\mathcal{E} := \{\epsilon_n\,;\, n = 0, 1, \ldots\}$ est une base hilbertienne de ℓ^2 ; mais $\mathcal{E}$ étant dénombrable, nous en déduisons que ℓ^2 est séparable. Si maintenant $(\mathscr{H}, \|\cdot\|)$ est un $\mathbb{K}$-espace de Hilbert séparable alors (algorithme de Gramm-Scmidt) $\mathscr{H}$ possède une base hilbertienne dénombrable, soit $\mathcal{F} := \{\varepsilon_n\,;\, n = 0, 1, \ldots\}$. Le $\mathbb{K}$-espace vectoriel ℓ_0^2 engendré par $\mathcal{E}$ est un sous-espace dense de ℓ^2 ; il est immédiat que l'application linéaire $\Phi : \ell_0^2 \to \mathscr{H}$ t.q. $\Phi(\epsilon_n) = \varepsilon_n$ est isométrique (i.e. $\|\Phi(a)\| = \|a\|_2$, pour tout $a \in \ell_0^2$) : par suite Φ s'étend en une unique application linéaire de ℓ^2 sur $\mathscr{H}$ qui est une isométrie de ℓ^2 sur $\mathscr{H}$.*

5.7. Nous considérons maintenant le cas particulier du théorème de Riesz-Fisher sur les espaces $L^p(I)$ (c.f. Théorème 2.9 supra) avec $I = [-\pi\,;\pi]$ et $p = 2$ qui nous permet d'affirmer que $L^2[-\pi,\pi]$ est un espace de Banach. Il est alors immédiat de vérifier que $L^2[-\pi,\pi]$ est un espace de Hilbert pour le norme $\|f\|_2 := \langle f|f\rangle^{1/2}$ avec le produit scalaire

$$(f,g) \mapsto \langle f|g\rangle := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx$$

Pour tout entier $n \in \mathbb{Z}$ nous notons

$$\mathbf{e}_n(x) := \cos(nx) + \boldsymbol{i}\sin(nx) = e^{\boldsymbol{i}nx}$$

de sorte que $\mathcal{E} := \{\mathbf{e}_n\,;\ n \in \mathbb{Z}\}$ est une famille orthonormée de $L^2[-\pi,\pi]$.

Théorème 5.10. *La famille orthonormée $\mathcal{E}$ est totale en ce sens que pour $f \in L^2[-\pi,\pi]$, si $\langle f\,|\,\mathbf{e}_n\rangle = 0$ pour tout $n \in \mathbb{Z}$, alors $f = 0$ presque partout : c'est une base hilbertienne de $L^2[-\pi,\pi]$.*

Une première démonstration du Théorème 5.10 est basée sur le fait [23] que $\mathscr{C}([-\pi\,;\pi],\mathbb{C})$ est dense dans $L^2[-\pi,\pi]$ (pour la topologie de la norme $\|\cdot\|_2$) combiné à la version trigonométrique du théorème d'approximation de Weierstrass affirmant que l'espace des polynômes trigonométriques est uniformément dense dans l'espace $\mathscr{C}_P([-\pi\,;\pi],\mathbb{C})$ des fonctions $f \in \mathscr{C}([-\pi\,;\pi],\mathbb{C})$ t.q. $f(-\pi) = f(\pi)$.

Lemme 5.11. *Soit S_nf la n-ième somme de Dirichlet de la fonction 2π-périodique f t.q. $f(x) = x$, pour tout $-\pi < x < \pi$ et $f(\pi) = 0$: alors $\|S_nf - f\|_2 \to 0$ quand $n \to +\infty$.*

Preuve. La fonction $f(x)$ étant impaire, nous avons $S_nf(x) = \sum_{k=1}^{n} b_k\sin(kx)$ où le coefficient en sinus d'ordre k (pour $k \geq 1$) est

$$b_k := \frac{1}{\pi}\int_{-\pi}^{\pi} x\sin(kx)dx = \frac{1}{\pi}\left[-\frac{x\cos(kx)}{k}\right]_{-\pi}^{\pi} + \frac{1}{k\pi}\int_{-\pi}^{\pi}\cos(kx)dx = 2\frac{(-1)^{k+1}}{k}$$

De plus d'après le théorème de Dirichlet (c.f. Théorème 4.3) nous avons pour tout $x \in \mathbb{R}$

$$f(x) = 2\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\sin(kx) = 2\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\left(\frac{e^{\boldsymbol{i}kx} - e^{-\boldsymbol{i}kX}}{2\boldsymbol{i}}\right) \tag{15}$$

La famille $\mathcal{E} := \{e^{\boldsymbol{i}kx}\,;\ k \in \mathbb{Z}\}$ étant orthonormée, c'est une base hilbertienne de la clôture $\mathscr{H}_0$ du $\mathbb{C}$-espace vectoriel engendré par $\mathcal{E}$. Or d'après (15), nous savons que f et S_nf ($n \geq 1$) sont des éléments de $\mathscr{H}_0$ tels que $\|f - S_n(f)\|_2^2 = 2\sum_{k=n+1}^{\infty} 1/k^2$: cela entraîne [24] que f est la limite dans $\mathscr{H}_0$ (et donc dans $L^2[-\pi\,;\pi]$) de S_nf quand $n \to +\infty$.

□

Preuve du Théorème 5.10 n°1. Nous devons démontrer que l'espace des polynômes trigonométriques est dense dans $L^2[-\pi\,;\pi]$ (muni de la norme $\|\cdot\|_2$). Considérons pour

---

23. Résultat de la théorie de la mesure que nous admettons (voir par exemple [GW90, Théorème 15.3.3]).
24. Notons au passage que l'identité de *« Pythagore-Parseval »* dans $\mathscr{H}_0$ appliquée à

$$f(x) = \sum_{k=1}^{+\infty}\frac{(-1)^{k+1}}{\boldsymbol{i}k}(e^{\boldsymbol{i}kx} - e^{-\boldsymbol{i}kx})$$

nous donne

$$2\sum_{k=1}^{\infty}\frac{1}{k^2} = \|f\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} x^2dx = \frac{1}{2\pi}\left[\frac{x^3}{3}\right]_{-\pi}^{\pi} = \frac{\pi^2}{3}$$

d'où *« la formule d'Euler »* $\sum_{k=1}^{\infty} 1/k^2 = \pi^2/6$.

cela $f \in L^2[-\pi\,;\pi]$ et soit $\varepsilon > 0$ arbitrairement donné. Nous partons du fait qu'il existe $g(x) \in \mathscr{C}([-\pi\,;\pi],\mathbb{C})$ t.q. $\|f-g\|_2 \le \varepsilon/3$. Si nous posons pour tout $-\pi < x < \pi$

$$h(x) = \frac{g(\pi)-g(-\pi)}{2\pi}x$$

alors d'une part (c.f. Lemme 5.11) il existe un polynôme trigonométrique $S(x)$ t.q. $\|h - S\|_2 \le \varepsilon/3$. Mais d'autre part $g-h$ étant une fonction de $\mathscr{C}_P([-\pi\,;\pi],\mathbb{C})$, nous pouvons appliquer la version trigonométrique du théorème d'approximation de Weierstrass qui nous donne un polynôme trigonométrique $T(x)$ t.q. $\|(g-h)-T\|_\infty \le 2\pi\varepsilon/3$. En particulier, nous avons $\|(g-h)-T\|_2 \le \varepsilon/3$ et finalement nous avons

$$\|f-(S+T)\|_2 \le \|f-g\|_2 + \|(g-h)-T\|_2 + \|h-S\|_2 \le \varepsilon$$

□

Remarque 5.12. *Le fait que la famille $\mathcal{E}$ formée des exponentielles complexes $\mathbf{e}_k(x) = e^{ikx}$ pour $k \in \mathbb{Z}$ soit une base hilbertienne de $L^2[-\pi\,;\pi]$ entraîne que pour toute fonction $f \in L^2[-\pi\,;\pi]$ la série de Fourier $\sum_{k=-\infty}^{+\infty}\langle f|\mathbf{e}_k\rangle\,\mathbf{e}_k$ converge vers f dans $L^2[-\pi\,;\pi]$: en d'autres termes, si nous notons $c_k := \langle f|\mathbf{e}_k\rangle$, alors le polynôme trigonométrique $S_nf(x) := \sum_{k=-n}^{n} c_k e^{ikx}$ converge vers f dans $L^2[-\pi\,;\pi]$. Nous retrouvons ainsi le résultat du Lemme 5.11, puisque la fonction 2π-périodique $f(x)$ t.q. $f(x) = x$, pour tout $-\pi < x < \pi$ avec $f(\pi) = 0$ appartient à $L^2[-\pi\,;+\infty]$.*

5.8. Pour démontrer le Théorème 5.10, nous avons utilisé que $\mathscr{C}([-\pi\,;\pi],\mathbb{C})$ était un sous-espace dense de $L^2[-\pi\,;\pi]$: c'est un résultat classique de la théorie de la mesure que nous n'avons pas justifié. Nous proposons maintenant une autre approche du Théorème 5.10 en suivant l'argument donné par J. Bass dans [Bas71, § II-3-4, p. 146 - 149]. Il s'agit de vérifier d'abord que la famille $\mathcal{E}$ *sépare* les points du sous-espace $\mathscr{C}_P([-\pi\,;\pi],\mathbb{C})$ de $L^2[-\pi\,;\pi]$ formé des fonctions φ qui sont continues et prolongeables en une fonction continue et 2π-périodique sur $\mathbb{R}$ (i.e. $\varphi \in \mathscr{C}([-\pi\,;\pi],\mathbb{C})$ avec $\varphi(-\pi) = \varphi(\pi)$).

Proposition 5.13. *Si $\varphi \in \mathscr{C}_P([-\pi\,;\pi],\mathbb{C})$, alors $\varphi \equiv 0$ dès que $\langle\varphi\,|\,\mathbf{e}_n\rangle = 0$, pour tout $n \in \mathbb{Z}$.*

Preuve. D'après le théorème de Dirichlet (c.f. Théorème 4.3).

□

Pour $f \in L^2[-\pi\,;\pi]$ nous commençons par remarquer que (inégalité de Cauchy-Schwarz)

$$\int_{-\pi}^{\pi}|f(x)|dx \le \sqrt{2\pi}\left(\int_{-\pi}^{\pi}|f(x)|^2dx\right)^{1/2} = 2\pi\|f\|_2$$

ce qui assure que $f \in L^1[-\pi\,;\pi]$; par suite, pour tout $x \in [-\pi,\pi]$

$$F(x) := \int_{-\pi}^{x} f(y)dy$$

est un nombre complexe bien définie. De plus, pour tout $0 \le a \le b \le 1$, nous avons (deuxième application de l'inégalité de Cauchy-Schwarz) :

$$|F(b)-F(a)| = \left|\int_a^b f(y)dy\right| \le \sqrt{(b-a)}\,\left(\int_a^b |f(y)|^2dy\right)^{1/2} \le \sqrt{2\pi(b-a)}\,\|f\|_2$$

de sorte que $x \mapsto F(x)$ est une fonction de $\mathscr{C}([-\pi\,;\pi],\mathbb{C})$.

Proposition 5.14. *Soit f une fonction de $L^2[-\pi\,;\pi]$ soit $F : [-\pi\,;\pi] \to \mathbb{C}$ la fonction telle que*

$$F(x) = \int_{-\pi}^{x} f(y)dy$$

Si $F \equiv 0$ alors f est nécessairement nulle presque partout.

Pour démontrer la Proposition 5.14, nous avons besoin du fait que si f est une fonction intégrable positive ou nulle sur $[-\pi\,;\pi]$ et telle que $\int_{-\pi}^{\pi} f(x)dx = 0$ alors $f(x) = 0$ pour presque tout $x \in [-\pi\,;\pi]$. Nous utiliserons aussi que la mesure de Lebesgue λ (sur la tribu $\mathcal{B}$ des boréliens inclus dans $[-\pi\,;\pi]$) est *« régulière »*, en ce sens que pour tout $A \in \mathcal{B}$

(16) $$\lambda(A) = \inf\left\{\lambda(U)\ ;\ A \subset U,\ U \text{ ouvert}\right\}$$

Preuve de la Proposition 5.14. Nous pouvons considérer séparément la partie réelles et la partie imaginaire de f : pour la démonstration, cela revient à supposer que f et F sont à valeurs réelles. Si $F(x) = 0$ pour tout $x \in [-\pi\,;\pi]$, alors nous avons

(17) $$\int_A f(y)dy = 0$$

pour tout intervalle A ouvert de $[-\pi\,;\pi]$: les ouverts de $[-\pi\,;\pi]$ étant réunions dénombrables d'intervalles ouverts disjoints deux à deux, nous en déduisons (convergence monotone) que (17) est vraie pour tout ouvert de $[-\pi\,;\pi]$. Enfin, d'après (16), l'identité (17) demeure vraie pour tout borelien $A \in \mathcal{B}$: en prenant pour A les deux ensembles boréliens $\{f > 0\}$ et $\{f < 0\}$, nous pouvons conclure que f est nulle presque partout.

$\square$

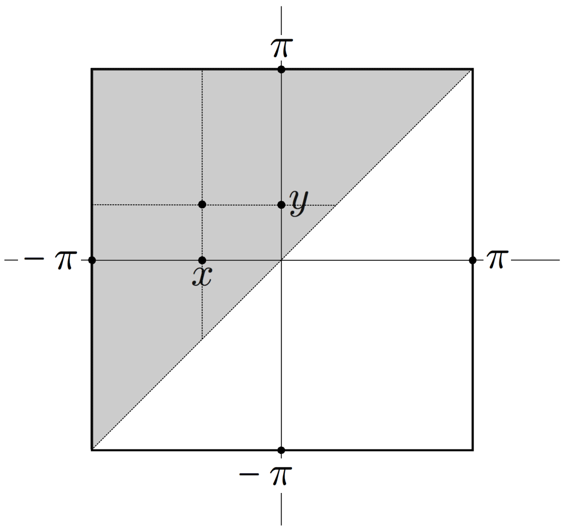


FIGURE 1. *Changement de sommation sur un domaine triangulaire du plan.*

Preuve du Théorème 5.10 n°2. Soit $f \in L^2[-\pi\,;\pi]$ telle que $\langle f\,|\,\mathrm{e}_n\rangle = 0$, pour tout $n \in \mathbb{Z}$ et soit $F : [-\pi\,;\pi] \to \mathbb{C}$ la fonction telle que $F(x) := \int_{-\pi}^{x} f(t)dt$. Nous savons d'une part que $F \in \mathscr{C}([-\pi\,;\pi], \mathbb{C})$; mais d'autre part $F(\pi) = 2\pi\,\langle f|\mathrm{e}_0\rangle = 0$ et donc $F(\pi) = F(-\pi) = 0$:

par suite F appartient à l'espace $\mathscr{C}_P([-\pi\,;\pi],\mathbb{C})$ des fonctions continues prolongeables en une fonction continue 2π-périodique sur $\mathbb{R}$. Nous affirmons que

$$(\forall n \in \mathbb{Z} \smallsetminus \{0\}) \qquad \int_{-\pi}^{\pi} F(x)e^{inx}dx = 0 \tag{18}$$

Si nous admettons (18) et si nous notons $C := \int_{-\pi}^{\pi} F(x)dx$, alors nous avons

$$(\forall n \in \mathbb{Z}) \qquad \int_{-\pi}^{\pi} (F(x) - C)e^{inx}dx = 0$$

Comme F est continue, cela assure que $F(x) = C$ pour tout $x \in [-\pi\,;\pi]$; or par définition $F(-\pi) = 0$ et donc $F(x) = C = 0$ pour tout $x \in [-\pi\,;\pi]$. Or d'après les Propositions 5.13 et 5.14, nous pouvons conclure successivement que F est nulle et que f s'annule presque partout. Il nous reste à établir (18) : en utilisant le fait que $1_{[-\pi\,;x]}(y) = \mathbf{1}_{[y\,;\pi]}(x)$ (voir Fig. 1), nous pouvons écrire que pour tout $n \in \mathbb{Z} \smallsetminus \{0\}$

$$\begin{aligned}
\int_{-\pi}^{\pi} F(x)e^{inx}dx &= \int_{-\pi}^{\pi} \left(\int_{-\pi}^{x} f(y)dy \right) e^{inx}dx \\
&= \int_{-\pi}^{\pi} f(y) \left(\int_{y}^{\pi} e^{inx}dx \right) dy \\
&= \int_{-\pi}^{\pi} f(y) \left[\frac{1}{in} e^{inx} \right]_y^{\pi} dy = \frac{1}{in} \left((-1)^n \int_{-\pi}^{\pi} f(y)dy - \int_{-\pi}^{\pi} f(y)e^{iny}dy \right)
\end{aligned}$$

soit encore

$$\int_{-\pi}^{\pi} F(x)e^{inx}dx = \frac{2\pi}{in} \Big((-1)^n \langle f|\mathbf{e}_0\rangle - \langle f|\mathbf{e}_{-n}\rangle \Big) = 0$$

□

6. Transformée de Fourier L^1

6.1. La transformée de Fourier permet d'analyser les phénomènes de nature vibratoire. En pratique on se ramène à un signal $f(t)$, i.e. une fonction de la variable temporelle t à valeurs complexes. La transformée de Fourier $f \mapsto \hat{f}$ permettant alors d'obtenir une représentation $\hat{f}(\nu)$ du signal dans la variable de fréquence ν. La *« dualité »* portant sur le couple *« temps-fréquence »* n'est pas la seule possible ; nous verrons qu'en mécanique quantique *« le principe de correspondance de Schrödinger »* est basé sur la transformée de Fourier de sorte que *« les variables conjuguées »* (comme par exemple la position et l'impulsion) sont liées par la transformée de Fourier : ici nous utiliserons simplement le couple (x, y) pour désigner les variables conjuguées. Les transformées de Fourier $\mathfrak{F}_-[f]$ (directe i.e. en analyse) et $\mathfrak{F}_+[f]$ (inverse i.e. en synthèse) d'une fonction complexe $f \in L^1(\mathbb{R})$ (i.e. Lebesgue-intégrable sur $\mathbb{R}$) sont définies en posant, pour tout $p \in \mathbb{R}$ (resp. $q \in \mathbb{R}$) :

$$\mathfrak{F}_-[f](y) = \frac{1}{\sqrt{2\pi}} \int f(x)e^{-ixy}dx \quad \text{et} \quad \mathfrak{F}_+[f](x) = \frac{1}{\sqrt{2\pi}} \int f(y)e^{ixy}dy \tag{19}$$

Il est immédiat de vérifier que la transformée de Fourier $\mathfrak{F}_\pm[\cdot]$ est $\mathbb{C}$-linéaire et que pour toute fonction $f \in L^1(\mathbb{R})$, la fonction $\mathfrak{F}_\pm[f]$ est borélienne avec

$$\|\mathfrak{F}_\pm[f]\|_\infty \leq \|f\|_1$$

Ainsi $\mathfrak{F}_\pm[\cdot]$ est une application linéaire bornée de $L^1(\mathbb{R})$ dans $L^\infty(\mathbb{R})$. Cependant $\mathfrak{F}_\pm[\cdot]$ n'est pas un endomorphisme de $L^1(\mathbb{R})$. En effet, la transformée de Fourier de la fonction

« porte » $\Pi(x) := \mathbf{1}_{[-1/2\,;1/2]}(x)$ s'écrit

$$\mathfrak{F}\left[\Pi\right](y) = \frac{1}{\sqrt{2\pi}}\int_{-1/2}^{+1/2} e^{-\boldsymbol{i}xy}dx = \frac{1}{\sqrt{2\pi}}\left[-\frac{1}{y/2}\,\frac{e^{-\boldsymbol{i}xy}}{2\boldsymbol{i}}\right]_{x=-1/2}^{x=+1/2} = \frac{1}{\sqrt{2\pi}}\,\frac{\sin(y/2)}{y/2} = \frac{1}{\sqrt{2\pi}}\mathrm{sinc}\left(\frac{y}{2\pi}\right)$$

où nous avons introduit *« la fonction sinus cardinal »* [25]

$$\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x} = \prod_{n=1}^{+\infty}\left(1-\frac{x^2}{n^2}\right)$$

Il est facile de vérifier que le sinus cardinal est une fonction qui n'est pas Lebesgue intégrable [26]. Le *« lemme de Riemann-Lebesgue »* donne une information importante sur les images des fonctions Lebesgue intégrables par la transformée de Fourier [27].

Théorème 6.1 (Riemann-Lebesgue)**.** *Si $f(x)$ est une fonction de $L^1(\mathbb{R})$ alors sa transformée de Fourier $\hat{f}(y) = \mathfrak{F}[f](y)$ est continue sur $\mathbb{R}$ et tend vers 0 quand $p \to \pm\infty$.*

Preuve. Voir par exemple [GW90, Théorèmes 5.1.1 & 17.1.3] ou [ST90, § 2.3, p. 96]. □

La proposition suivante résume les propriétés basiques de la transformée de Fourier.

Proposition 6.2. *Si f est une fonction dans $L^1(\mathbb{R})$, alors pour tout $\alpha \in \mathbb{R}\smallsetminus\{0\}$:*

$$\begin{array}{rcll} (i) & : & \mathfrak{F}[f(\alpha x)](y) = 1/|\alpha|\,\mathfrak{F}[f]\,(y/\alpha) & \quad (\textit{changement d'échelle}) \\ (ii) & : & \mathfrak{F}[f(x-\alpha)](y) = e^{-\boldsymbol{i}\alpha y}\,\mathfrak{F}[f](y) & \quad (\textit{effet de translation}) \\ (iii) & : & \mathfrak{F}[e^{\boldsymbol{i}\alpha x}f(x)](y) = \mathfrak{F}[f](y-\alpha) & \quad (\textit{effet de modulation}) \end{array}$$

6.2. Les relations entre la transformée de Fourier et la dérivation des fonctions peuvent être exprimées à l'aide de *« l'opérateur position »* $f \mapsto Qf$ t.q. $Qf(x) := xf(x)$. Notons au passage que Q ne peut définir un endomorphisme de $L^1(\mathbb{R})$ (ni de $L^2(\mathbb{R})$: cette remarque deviendra déterminante pour la mécanique ondulatoire de de Broglie-Schrödinger : c.f. [Oli16a]). Par application du théorème de dérivation sous le signe intégral nous pouvons déduire la proposition suivante.

Proposition 6.3. *Soit $f \in L^1(\mathbb{R})$ et $k \geq 1$; (i) : si $x^k f(x) \in L^1(\mathbb{R})$ alors $\mathfrak{F}[f](y) \in \mathscr{C}^k(\mathbb{R})$ et*

$$\partial^k\mathfrak{F}[f](y) = \mathfrak{F}\left[(-\boldsymbol{i}Q)^k f\right](y)$$

(ii) : si $f \in \mathscr{C}^k(\mathbb{R})$ et si $\partial^k f \in L^1(\mathbb{R})$ alors $Q^k\mathfrak{F}[f] \in L^1(\mathbb{R})$ et

$$\mathfrak{F}\left[\partial^k f\right] = (\boldsymbol{i}Q)^k\mathfrak{F}[f]$$

25. L'expression de la fonction sinus cardinal en produit eulérien a été introduite et utilisée par Euler dans sa preuve heuristique géniale de la formule $\zeta(2) = \pi^2/6$; rappelons aussi que la fonction sinus cardinal est reliée à la fonction gamma par la formule des compléments d'Euler avec (c.f. [ST89, p. 155-156]) :

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

26. Il est aussi facile de vérifier que $\mathrm{sinc}(x)$ est de carré intégrable :. Il est alors possible de démontrer que la famille $(\mathrm{sinc}(x-n)\ ;\ n \in \mathbb{Z})$ est une base hilbertienne de $L^2(\mathbb{R})$: ce résultat important est lié au *« théorème de l'échantillonnage de Shannon »* (c.f. [GW90, Leçon n° 38]).

27. Bien que très fondamental, nous ne développons pas ici la question du lemme de Riemann-Lebesgue. Nous y reviendrons plus en détail à propos de l'intégrale de chemin de Feynamnn (c.f. [Oli17]) et de l'interprétation du phénomène *« des interférences destructrices »*.

Preuve. Voir par exemple [GW90, Théorème 17.2.1, p. 130] ou [ST90, § 2.4, p. 97].

□

6.3. Les fonctions gaussiennes jouent un rôle important pour la transformée de Fourier (nous les retrouverons en particulier dans le cas d'égalité de l'inégalité de Heisenberg-Gabor : c.f. § 10). Le résultat suivant sur les *« fonctions gaussiennes à écart type complexe »* est classique (dans le cas réel).

Proposition 6.4. *Soit $\xi = a + \boldsymbol{i}b$ avec $a, b \in \mathbb{R}$ et $a > 0$; alors*

$$(i)\ :\ \int e^{-\xi x^2}dx = \sqrt{\frac{\pi}{\xi}} \quad et \quad (ii)\ :\ \int e^{-\xi x^2}e^{\boldsymbol{i}yx}dx = \sqrt{\frac{\pi}{\xi}}\, e^{-\frac{1}{4}\frac{y^2}{\xi}}$$

(ici $\sqrt{\xi}$ est le nombre complexe η t.q. $\mathfrak{re}(\eta) > 0$ et $\eta^2 = \xi$).

Preuve. (i) : Pour calculer $\gamma := \int e^{-\xi x^2}dx$, nous remarquons que

$$\begin{aligned}\gamma^2 &= \iint e^{-\xi(x^2+y^2)}dxdy\\ &= \frac{1}{2\xi}\int_{-\pi}^{\pi}\int_0^{\infty} 2\xi\rho e^{-\xi\rho^2}d\rho d\theta = \frac{1}{2\xi}\int_{-\pi}^{\pi}\left[-e^{-\xi\rho^2}\right]_{\rho=0}^{\infty}d\theta = \frac{\pi}{\xi}\lim_{\rho\to+\infty}\left[1 - e^{-a\rho^2}e^{-\boldsymbol{i}b\rho^2}\right]\end{aligned}$$

Or nous avons supposé $a > 0$ et par suite $\gamma^2 = \pi/\xi$.

(ii) : Pour tout $y \in \mathbb{R}$ nous notons

$$\varphi(y) := \int e^{-\xi x^2}e^{\boldsymbol{i}yx}dx$$

alors (dérivation sous le signe intégral puis, intégration par parties) il vient :

$$\begin{aligned}\frac{d\varphi}{dy} &= -\frac{1}{2\xi}\int\left(-2\xi x e^{-\xi x^2}\right)\left(\boldsymbol{i}e^{\boldsymbol{i}xy}\right)dx\\ &= -\frac{1}{2\xi}\left[e^{-\xi x^2}\left(\boldsymbol{i}e^{\boldsymbol{i}xy}\right)\right]_{-\infty}^{+\infty} + \frac{1}{2\xi}\int e^{-\xi x^2}\left(-ye^{\boldsymbol{i}xy}\right)dx\end{aligned}$$

L'hypothèse $a > 0$ assurant que le terme intégré est nul, nous obtenons

$$\frac{d\varphi}{dy} + \frac{y\varphi(y)}{2\xi} = 0 \tag{20}$$

ce qui équivaut à la quadrature, $d/dy\left(e^{\frac{1}{4}y^2/\xi}\varphi(y)\right) = 0$. Par suite $\varphi(y) = \varphi(0)e^{-\frac{1}{4}y^2/\xi}$ et la conclusion vient de (i) du fait que $\varphi(0) = \sqrt{\pi/\xi}$.

□

Le cas particulier de la partie (ii) de la Proposition 6.4 énoncé dans le corollaire suivante est particulièrement important en pratique.

Corollaire 6.5. *Soit $G_\xi(x) := e^{-\xi x^2/2}$, où $\xi = a + \boldsymbol{i}b$ avec $a, b \in \mathbb{R}$ et $a > 0$; alors*

$$\mathfrak{F}[G_\xi](y) = \frac{1}{\sqrt{2\pi}}\int e^{-\xi x^2/2}e^{\boldsymbol{i}yx}dx = \frac{1}{\sqrt{\xi}}\,e^{-\frac{1}{2}\frac{y^2}{\xi}}$$

(ici $\sqrt{\xi}$ est le nombre complexe η t.q. $\mathfrak{re}(\eta) > 0$ et $\eta^2 = \xi$). En particulier

$$\mathfrak{F}[G_1] = G_1$$

6.4. Dans ce paragraphe nous allons démontrer *« le théorème d'inversion de la transformée de Fourier L^1 »* (c.f. Théorème 6.6) affirmant que si f et $\mathfrak{F}_-[f]$ sont toutes deux des fonctions de $L^1(\mathbb{R})$ alors $\mathfrak{F}_+[\mathfrak{F}_-[f]] = f$ (dans $L^1(\mathbb{R})$). Noter au passage (lemme de Riemann-Lebesgue) que dans cette situation les deux fonctions f et $\mathfrak{F}_-[f]$ sont dans $\mathscr{C}_0(\mathbb{R})$ (espace des fonctions continues sur $\mathbb{R}$ et nulle en $\pm\infty$).

Théorème 6.6 (Formule d'inversion). *Si f et $\mathfrak{F}_-[f]$ sont dans $L^1(\mathbb{R})$, alors $\mathfrak{F}_+[\mathfrak{F}_-[f]](x) = f(x)$ pour presque tout $x \in \mathbb{R}$. (De même, si f et $\mathfrak{F}_+[f]$ sont dans $L^1(\mathbb{R})$, alors $\mathfrak{F}_-[\mathfrak{F}_+[f]] = f$ p.p..)*

Preuve. Soit $f \in L^1(\mathbb{R})$ t.q. $\mathfrak{F}_-[f] \in L^1(\mathbb{R})$ et notons $U(x) := 1/\sqrt{2\pi}e^{-x^2/2}$: le facteur $1/\sqrt{2\pi}$ est introduit pour assurer que $\int U(x)dx = 1$. La première étape de l'argument consiste à remarquer que $U(0) = 1/\sqrt{2\pi}$, de sorte que pour tout $x, y \in \mathbb{R}$:

$$\lim_{n\to+\infty} e^{\boldsymbol{i}xy}U(y/n)\,\mathfrak{F}_-[f](y) = \frac{1}{\sqrt{2\pi}}e^{\boldsymbol{i}xy}\mathfrak{F}_-[f](y)$$

Mais d'autre part $|e^{\boldsymbol{i}xy}U(y/n)\mathfrak{F}_-[f](y)| \le 1/\sqrt{2\pi}\,|\mathfrak{F}_-[f](y)|$: or (hypothèse) $\mathfrak{F}_-[f] \in L^1(\mathbb{R})$, de sorte que par le *« théorème de la convergence dominée »*, nous avons pour tout $x \in \mathbb{R}$:

$$\lim_{n\to+\infty} \mathfrak{F}_+\Big[U(\cdot/n)\mathfrak{F}_-[f]\Big](x) = \frac{1}{\sqrt{2\pi}}\mathfrak{F}_+[\mathfrak{F}_-[f]](x) \tag{21}$$

La deuxième étape utilise le fait que les $U_n(x) := nU(nx)$ $(n \ge 1)$ définissent une approximation de l'unité (dite gaussienne). Comme U est une fonction bornée et que $\mathfrak{F}_-[f]$ est (par hypothèse) dans $L^1(\mathbb{R})$, il est évident que $U(\cdot/n)\mathfrak{F}_-[f]$ est aussi dans $L^1(\mathbb{R})$, pour tout $n \ge 1$ entier : en utilisant *« le théorème de Fubini »* nous pouvons écrire

$$\begin{aligned}\mathfrak{F}_+\Big[U(\cdot/n)\mathfrak{F}_-[f]\Big](x) &= \frac{1}{\sqrt{2\pi}}\int e^{\boldsymbol{i}xy}U(y/n)\left(\frac{1}{\sqrt{2\pi}}\int e^{-\boldsymbol{i}zy}f(z)dz\right)dy\\ &= \frac{1}{\sqrt{2\pi}}\int\left(\frac{1}{\sqrt{2\pi}}\int e^{-\boldsymbol{i}y(z-x)}U(y/n)dy\right)f(z)dz\end{aligned}$$

soit encore

$$\mathfrak{F}_+\Big[U(\cdot/n)\mathfrak{F}_-[f]\Big](x) = \frac{1}{\sqrt{2\pi}}\int \mathfrak{F}_-\big[U(\cdot/n)\big](z-x)f(z)dz \tag{22}$$

Du fait que $\mathfrak{F}_-[U] = U$ (c.f. Corollaire 6.5) et par définition de U_n, nous avons

$$\mathfrak{F}_-\big[U(\cdot/n)\big](z-x) = n\mathfrak{F}_-[U](n(z-x)) = nU\big(n(z-x)\big) = U_n(z-x)$$

et comme $U_n(-t) = U_n(t)$, nous pouvons récrire (22) sous la forme

$$\mathfrak{F}_+\Big[U(\cdot/n)\mathfrak{F}_-[f]\Big](x) = \frac{1}{\sqrt{2\pi}}\int U_n(x-z)f(z)dz = \frac{1}{\sqrt{2\pi}}U_n * f(x) \tag{23}$$

Le Théorème 3.8 nous assurant que $U_n * f \to f$ dans $L^1(\mathbb{R})$, nous déduisons du corollaire du *« théorème de Riesz-Fisher »* (c.f. Proposition 2.10) qu'il existe une suite strictement croissante d'entiers $n_0, n_1, \ldots$ telle que pour presque tout $x \in \mathbb{R}$

$$\lim_{k\to+\infty} U_{n_k} * f(x) = f(x) \tag{24}$$

La conclusion vient grâce à l'identité (23) en comparant (24) et (21).

□

Nous pouvons reformuler la formule d'inversion de la transformée de Fourier tel qu'énoncée dans le Théorème 6.6 en introduisant *« l'espace de Wiener »* $\mathscr{A}(\mathbb{R})$ c'est-à-dire le sous-espace de $L^1(\mathbb{R})$ constitué des fonctions dont la transformée de Fourier est encore dans $L^1(\mathbb{R})$ (i.e. $f \in \mathscr{A}(\mathbb{R})$ ssi f et $\hat{f} = \mathfrak{F}_-[f]$ sont dans $L^1(\mathbb{R})$). En effet, le Théorème 6.6 assure que $\mathscr{A}(\mathbb{R})$ est stable par transformée de Fourier (directe et inverse) et que la restriction $\mathfrak{F}_\pm[\cdot]$ à $\mathscr{A}(\mathbb{R})$ est un endomorphisme bijectif continu de norme 1 avec $\mathfrak{F}_-[\cdot]^{-1} = \mathfrak{F}_+[\cdot]$. Notons aussi que par le lemme de Riemann-Lebesgue, $\mathscr{A}(\mathbb{R})$ s'identifie à un sous-espace[28] de $\mathscr{C}_0(\mathbb{R})$ (espace des fonctions continues sur $\mathbb{R}$ et nulle en $\pm\infty$)[29] : en d'autres termes, cela signifie que chaque classe de fonctions de $\mathscr{A}(\mathbb{R})$ est identifiée à son unique représentant continu. Du fait que deux fonctions continues qui sont égales presque partout sont nécessairement égales *partout*, nous pouvons récrire le Théorème 6.6 comme suit.

Théorème 6.7 (Formule d'inversion bis)**.** *Si f et appartient à l'espace de Wiener $\mathscr{A}(\mathbb{R})$ alors $\mathfrak{F}_+[\mathfrak{F}_-[f]](x) = f(x)$ pour tout $x \in \mathbb{R}$: ainsi la transformée de Fourier $\mathfrak{F}_-[\cdot]$ réalise un isomorphisme de $\mathscr{A}(\mathbb{R})$ dont l'inverse est la transformée de Fourier inverse $\mathfrak{F}_+[\cdot]$.*

Remarque 6.8. *(1) : Remarquons que le Théorème 3.6 combiné à la Proposition 6.9 entraînent que si f et g sont dans l'espace de Wiener $\mathscr{A}(\mathbb{R})$ alors la convolée $f * g$ est aussi dans $\mathscr{A}(\mathbb{R})$ avec de plus $\|f * g\|_1 \le \|f\|_1 \cdot \|g\|_1$. Ainsi $(\mathscr{A}(\mathbb{R}), *)$ est une algèbre de convolution (ce n'est pas une algèbre de Banach car $(\mathscr{A}(\mathbb{R}), \|\cdot\|_1)$ n'est pas complet).*

(2) : L'espace de Wiener $\mathscr{A}(\mathbb{R})$ est un sous-espace strict de $\mathscr{C}_0(\mathbb{R})$; il est difficile de cerner les contours de $\mathscr{A}(\mathbb{R})$: notons par exemple (c.f. [GW90, Proposition 18.1.2, p. 137]*) que si f appartient à $\mathscr{C}^2(\mathbb{R}) \cap L^1(\mathbb{R})$ et que les dérivées f' et f'' sont toutes les deux dans $L^1(\mathbb{R})$ alors $\mathfrak{F}_-[f]$ est dans $L^1(\mathbb{R})$ et donc $f \in \mathscr{A}(\mathbb{R})$.*

(3) : La transformée de Fourier de la fonction porte $\Pi(x) = \mathbf{1}_{[-1/2\,;1/2]}(x)$ n'est pas une fonction de $L^1(\mathbb{R})$ (c.f. § 6.1) : cela pose la question de la formule d'inversion pour des fonctions de $L^1(\mathbb{R})$ continues par morceau. Mentionnons une version possible de la formule d'inversion qui est – en un sens – l'analogue du théorème de Dirichlet pour les séries de Fourier (c.f. Théorème 4.3) : elle est connue sous le nom de « formule d'inversion en valeurs principales ». On trouve un énoncé de ce résultat dans [GW90, Théorème 18.3.1, p. 139] *affirmant que si f est une fonction de $L^1(\mathbb{R})$ continûment dérivable sur chaque intervalle $]-\infty\,; a_1[,]a_2\,; a_3[, \dots,]a_n\,; +\infty[$ (avec $-\infty < a_1 < a_2 < \cdots < a_N < +\infty$) et que la fonction dérivée f' (définie partout sauf possiblement en un nombre fini de points) est elle-même dans $L^1(\mathbb{R})$, alors*

$$\lim_{a\to+\infty} \frac{1}{\sqrt{2\pi}} \int_{-a}^{+a} e^{ixy} \mathfrak{F}_-[f](y)dy = \frac{1}{2}\big(f(x^+) + f(x^-)\big)$$

Proposition 6.9. *Soient $f, g \in L^1(\mathbb{R})$, de transformée de Fourier $\hat{f} = \mathfrak{F}_-[f]$ et $\hat{g} = \mathfrak{F}_-[g]$; alors*

(i) : $f * g \in L^1(\mathbb{R})$ et $\mathfrak{F}_-[f * g] = \hat{f}\hat{g}$;

(ii) : si de plus $\hat{f}$ et $\hat{g}$ sont aussi dans $L^1(\mathbb{R})$, alors $fg \in L^1(\mathbb{R})$ et $\mathfrak{F}_-[fg] = \hat{f} * \hat{g}$.

28. C'est un sous-espace strict de $\mathscr{C}_0(\mathbb{R})$.
29. Notons aussi que $\mathscr{A}(\mathbb{R})$ est aussi sous-espace de $L^p(\mathbb{R})$ pour tout $1 \le p \le +\infty$.

Preuve. (i) : D'une part, le fait que $f * g$ soit dans $L^1(\mathbb{R})$ est inclus dans l'inégalité de Young (c.f. Théorème 3.6) ; d'autre part, d'après le théorème de Fubini, nous avons

$$\begin{aligned}
\mathfrak{F}_-[f*g](y) &= \int\left(\int f(x-z)g(z)dz\right)e^{-ixy}dx \\
&= \int\left(\int f(x-z)e^{-ixy}dx\right)g(z)dz \\
&= \int\left(\int f(u)e^{-i(u+z)y}du\right)g(z)dz = \left(\int f(u)e^{-iuy}du\right)\left(\int g(z)e^{-izy}dz\right)
\end{aligned}$$

(ii) : Comme $\hat{f}$ et $\hat{g}$ sont dans $L^1(\mathbb{R})$ nous pouvons appliquer la partie (i) avec $\mathfrak{F}_+[\cdot]$ (au lieu de $\mathfrak{F}_-[\cdot]$), de sorte que $\hat{f} * \hat{g}$ est une fonction de $L^1(\mathbb{R})$ t.q.

$$\mathfrak{F}_+[\hat{f} * \hat{g}] = \mathfrak{F}_+[\hat{f}] \cdot \mathfrak{F}_+[\hat{g}] = f\,g \tag{25}$$

où la dernière égalité découle du théorème de l'inversion (c.f. Théorème 6.6). Mais le fait que $f = \mathfrak{F}_+[\hat{f}]$ et $\hat{f} = \mathfrak{F}_-[f]$ (resp. $g = \mathfrak{F}_+[\hat{g}]$ et $\hat{g} = \mathfrak{F}_-[g]$) soient toutes deux des fonctions de $L^1(\mathbb{R})$ assure qu'elles sont aussi des fonctions continues nulles à l'infini (c.f. Lemme de Riemann-Lebesgue). Comme f et g sont des fonctions dans $L^1(\mathbb{R})$, le fait qu'elles soient aussi continues sur $\mathbb{R}$ et nulles en $\pm\infty$ entraîne que $f\,g \in L^1(\mathbb{R})$ (en effet il existe $\alpha > 0$ t.q. $f\,g$ soit bornée sur $[-\alpha\,;\alpha]$ et $|f\,g| \le |f(x)|$ pour $|x| \ge \alpha$). Nous savons donc que d'une part que $\hat{f} * \hat{g} \in L^1(\mathbb{R})$ et comme d'autre part $\mathfrak{F}_+[\hat{f} * \hat{g}] = f\,g$ est aussi une fonction de $L^1(\mathbb{R})$, en partant de l'identité (25) et en appliquant $\mathfrak{F}_-[\cdot]$, nous déduisons (nouvelle application du théorème de l'inversion) que $\hat{f} * \hat{g} = \mathfrak{F}_-[f\,g]$.

□

7. Transformée de Fourier L^2

7.1. Nous nous intéressons maintenant à la transformée de Fourier des *« signaux d'énergie finie »*, c'est-à-dire à la transformée de Fourier (L^2) comme endomorphisme de l'espace de Hilbert (complexe) $L^2(\mathbb{R})$ des fonctions de carré sommable. Une première approche est basée sur l'utilisation de *« l'espace de Schwartz »* (sur $\mathbb{R}$) traditionnellement noté[30] $\mathscr{S}(\mathbb{R})$, c'est-à-dire l'espace constitué des fonctions $\varphi : \mathbb{R} \to \mathbb{R}$ qui sont C^∞ sur $\mathbb{R}$ pour lesquelles chaque dérivée $\partial^p f$ (pour $p \ge 0$) est négligeable en $\pm\infty$ devant toute fonction polynômiale ; cette dernière condition est équivalente à dire que pour tout entier $p, q \ge 0$, il existe une constante $K(p,q)$ t.q.

$$|\partial^p f(x)| \le \frac{K(p,q)}{(1+x^2)^{q/2}} \tag{26}$$

Nous introduisons maintenant *« l'opérateur position »* noté Q (important dans la suite) et t.q. pour toute fonction $x \mapsto f(x)$ définie sur $\mathbb{R}$ et à valeurs dans $\mathbb{C}$,

$$Qf(x) = xf(x) \tag{27}$$

Alors, pour tout entier $n \ge 0$, nous notons

$$\mathcal{N}_n(f) := \sum_{i+j=n} \|Q^i \partial^j f\|_\infty$$

30. Laurent Schwartz utilisait cette notation, appelant ces fonctions tests des fonctions *sphériques*.

La famille $\{\mathcal{N}_n\ ;\ n \geq 0\}$ de semi-normes séparantes ($\mathcal{N}_0(f)$ coïncide à la norme infinie de f) : il est alors facile de vérifier qu'en posant :

(28)
$$d(f,g) = \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{\mathcal{N}_n(f-g)}{1+\mathcal{N}_n(f-g)}$$

nous définissons une métrique $d(\cdot,\cdot)$ sur $\mathscr{S}(\mathbb{R})$ (pour l'inégalité triangulaire il suffit de constater que la fonction $f(x) = x/(1+x)$ vérifie $f(a+b) \leq f(a) + f(b)$).

Proposition 7.1. *L'espace $\mathscr{S}(\mathbb{R})$ est complet* [31] *pour la métrique $d(\cdot,\cdot)$ définie en (28).*

Preuve. Il s'agit de montrer que si $\varphi_1, \varphi_2, \ldots$ est une suite de Cauchy de $\mathscr{S}(\mathbb{R})$, alors elle converge dans $\mathscr{S}(\mathbb{R})$ vers une fonction φ. Par définition de la métrique $d(\cdot,\cdot)$ nous savons que (pour tout $p \geq 0$) les fonctions $\partial^p\varphi_1, \partial^p\varphi_2, \ldots$ forment une suite de Cauchy de l'espace de Banach $(\mathscr{C}(\mathbb{R}), \|\cdot\|_\infty)$ – espace des fonctions continues pour la norme sup – et que donc il existe une fonction ϕ_p continue sur $\mathbb{R}$ t.q. $\|\partial^p\varphi_n - \phi_p\|_\infty \to 0$ quand $n \to +\infty$. En particulier, l'ensemble de ces convergences entraînent que pour tout $k \geq 1$ la fonction ϕ_p est continûment dérivable sur $\mathbb{R}$, avec $\partial\phi_p = \phi_{p+1}$: si nous notons $\varphi := \phi_0$, alors il vient par induction que φ est C^∞ sur $\mathbb{R}$ avec $\partial^p\varphi = \phi_p$. Pour conclure, considérons $p, q \geq 0$ (entiers) et $\varepsilon > 0$ (réel) arbitrairement donnés : comme $x^q\partial^p\varphi_n(x) \to x^q\partial^p\varphi(x)$ uniformément sur $\mathbb{R}$, il existe un rang N tel que $|x^q\partial^p\varphi(x)| \leq \varepsilon + |x^q\partial^p\varphi_n(x)|$, pour tout $x \in \mathbb{R}$ et tout $n \geq N$; finalement φ_n étant dans $\mathscr{S}(\mathbb{R})$, nous savons que $x^q\partial^p\varphi_n(x) \to 0$ quand $x \to \pm\infty$, ce qui assure que $|x^q\partial^p\varphi(x)| \leq \varepsilon$ quand $x \to \pm\infty$.

□

Proposition 7.2. *$\mathscr{S}(\mathbb{R})$ est stable par transformée de Fourier (L^1).*

Preuve. En utilisant l'opérateur position $f \mapsto Qf$ t.q. $Qf(x) = xf(x)$ (c.f. § 6.2) nous savons que si $\varphi \in \mathscr{S}(\mathbb{R})$, alors $Q^k\varphi \in L^1(\mathbb{R})$, pour tout $k \geq 0$: la partie (i) de la Proposition 6.3 assure alors que $\hat{\varphi} := \mathfrak{F}_-[\varphi] \in \mathscr{C}^k(\mathbb{R})$. Il nous reste à montrer que $Q^k\hat{\varphi}(x) \to 0$ quand $x \to \pm\infty$: mais comme $\partial^k\varphi \in \mathscr{S}(\mathbb{R}) \subset L^1(\mathbb{R})$, la partie (ii) de la Proposition 6.3 assure que $|Q^k\hat{\varphi}(x)| = |\mathfrak{F}_-[\partial^k\varphi](x)|$: la conclusion vient du lemme de Riemann-Lebesgue.

□

Il découle du théorème d'inversion (c.f. Théorème 6.6) que les restrictions bilatérales $\mathfrak{F}_\pm[\cdot] : \mathscr{S}(\mathbb{R}) \to \mathscr{S}(\mathbb{R})$ sont inverses l'une de l'autre, en ce sens que pour tout f dans $\mathscr{S}(\mathbb{R})$,

(29)
$$\mathfrak{F}_+[\mathfrak{F}_-[f]] = \mathfrak{F}_-[\mathfrak{F}_+[f]] = f$$

D'autre part, si g est une autre fonction de $\mathscr{S}(\mathbb{R})$, alors (en utilisant le théorème de Fubini)

$$\int \mathfrak{F}_-[f](y)\overline{g(y)}dy = \int \left(\frac{1}{\sqrt{2\pi}} \int f(x)e^{-ixy}dx\right)\overline{g(y)}dy = \int f(x)\left(\frac{1}{\sqrt{2\pi}} \int \overline{g(y)e^{-ixy}}dy\right)dx$$

soit encore

(30)
$$\int \mathfrak{F}_-[f](y)\overline{g(y)}dy = \int f(x)\overline{\mathfrak{F}_+[g](x)}dx$$

31. On peut montrer que c'est « *un espace de Fréchet* »

7.2. L'intégrale de $|f(x)|^2$ est appelée *« l'énergie du signal »* f : les signaux d'énergie finie forment donc l'espace $L^2(\mathbb{R})$ des fonctions de carré intégrable. D'après le *« théorème de Riesz-Fisher »* c'est un *« espace de Hilbert »* pour le produit scalaire $\langle\cdot|\cdot\rangle$ t.q.

$$\langle f|g\rangle = \int f(x)\overline{g(x)}dx$$

Nous notons $\|f\| := \langle f|f\rangle^{1/2}$ la norme hilbertienne sur $L^2(\mathbb{R})$. Une fonction de carré intégrable sur $\mathbb{R}$ n'étant pas nécessairement intégrable, la transformée de Fourier ne peut se définir directement sur $L^2(\mathbb{R})$. Nous allons utiliser le fait que l'espace de Schwarz $\mathscr{S}(\mathbb{R})$ est un sous-espace dense de $L^2(\mathbb{R})$ t.q. $\mathfrak{F}_{\pm}[\mathscr{S}(\mathbb{R})] \subset \mathscr{S}(\mathbb{R})$. Soient $f, g \in \mathscr{S}(\mathbb{R})$; alors d'après (30), nous pouvons écrire que :

$$\left\langle \mathfrak{F}_{-}[f]\middle|g\right\rangle = \left\langle f\middle|\mathfrak{F}_{+}[g]\right\rangle \tag{31}$$

En combinant (29) et (31) il vient aussi

$$\left\langle \mathfrak{F}_{-}[f]\middle|\mathfrak{F}_{-}[g]\right\rangle = \left\langle f\middle|\mathfrak{F}_{+}[\mathfrak{F}_{-}[g]]\right\rangle = \langle f|g\rangle \tag{32}$$

En particulier $\|\mathfrak{F}_{-}[f](x)\| = \|f\|$, pour tout $f \in \mathscr{S}(\mathbb{R})$: l'espace de Schwarz étant dense dans $L^2(\mathbb{R})$, l'endomorphisme $\mathfrak{F}_{\pm}[\cdot] : \mathscr{S}(\mathbb{R}) \to \mathscr{S}(\mathbb{R})$ se prolonge de manière unique (c.f. *« théorème de prolongement des opérateurs »*) en un endomorphisme unitaire (isométrie inversible [32]) de $L^2(\mathbb{R})$, que nous continuons à noter $\mathfrak{F}_{\pm}[\cdot]$. Ce résultat peut s'énoncer (et se démontrer) sous plusieurs formes : il est généralement connu comme de *« théorème de Plancherel »* (ou encore comme le *« théorème de Parseval-Plancherel »* [33]) pour la transformée de Fourier sur $L^2(\mathbb{R})$: on pourra consulter à ce sujet [Wie04, Ch. I, p. 46-71]. Dans la suite nous nous référerons à l'énoncé suivant :

Théorème 7.3. *La transformée de Fourier directe $\mathfrak{F}_{-}[\cdot]$ qui prolonge par continuité à $L^2(\mathbb{R})$ la transformée de Fourier directe sur $\mathscr{S}(\mathbb{R})$, est un endomorphisme unitaire de $L^2(\mathbb{R})$ dont l'inverse est le prolongement par continuité $\mathfrak{F}_{+}[\cdot]$ de la transformée de Fourier inverse sur $\mathscr{S}(\mathbb{R})$.*

7.3. Nous allons voir la définition directe de la transformée de Fourier L^2 telle qu'introduite par Wiener en 1933 (c.f. [Wie04]) : outre l'élégance de cette présentation, elle a d'autant plus d'intérêt (relativement aux questions de la mécanique quantique), qu'elle est directement liée à la résolution de l'équation de Schrödinger de l'oscillateur harmonique (c.f. [Oli16b]). En dérivant successivement la fonction e^{-x^2}, nous obtenons

$$\frac{d^n}{dx^n}e^{-x^2} = (-1)^n C_n(x) e^{-x^2}$$

où pour tout entier $n \geq 0$

$$C_n(x) = (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2}$$

32. Il découle directement de (31) que *l'adjoint* $\mathfrak{F}_{-}[\cdot]$ coïncide avec $\mathfrak{F}_{+}[\cdot]$: or (29) entraîne que $\mathfrak{F}_{-}[\cdot]^{-1} = \mathfrak{F}_{+}[\cdot]$, ce qui signifie que $\mathfrak{F}_{-}[\cdot]$ est un endomorphisme unitaire de $L^2(\mathbb{R})$.

33. En fait le résultat de Plancherel (1910) est l'analogue pour la transformée de Fourier dans $L^2(\mathbb{R})$ du résultat de Parseval (1799) pour les séries de Fourier dont les coefficients sont de carré sommable.

le n-ème polynôme d'Hermite (physique [34]). Par des intégrations par parties successives, il est facile de vérifier que $(C_0, C_1, \dots)$ est une famille orthogonale de $L^2(\mu(dx))$ où $\mu(dx) = e^{-x^2}dx$, avec plus précisément :

$$\int C_n(x)C_m(x)e^{-x^2}dx = 2^n n!\sqrt{\pi}\delta^n{}_m$$

Par suite, en définissant les fonctions d'Hermite [35]

$$\psi_n(x) := \frac{1}{\sqrt{2^n n!\sqrt{\pi}}}C_n(x)e^{-x^2/2}$$

nous obtenons une famille $(\psi_0, \psi_1, \dots)$ qui est orthonormée dans $L^2(\mathbb{R}) = L^2(dx)$. Le point important ici est que $(\psi_0, \psi_1, \dots)$ constinue en fait une base hilbertienne de $L^2(\mathbb{R})$ (c.f. [Wie04, § 8, p. 64]). La définition de la transformée de Fourier L^2 donnée par Wiener se justifie alors par la proposition suivante :

Proposition 7.4. $\mathfrak{F}[\psi_n] = (\boldsymbol{i})^n\psi_n$*, pour tout entier* $n \geq 0$.

Preuve. Nous utilisons le fait que la fonction génératrice des polynômes d'Hermite est [36]

$$G(x,t) := \sum_{n=0}^{\infty} C_n(x)\frac{t^n}{n!} = e^{2tx-t^2}$$

Ainsi, pour calculer la transformée de Fourier des fonctions $\psi_n(x) := c_n e^{-x^2/2}C_n(x)$, il est assez naturel de considérer la fonction

$$e^{-x^2/2}G(x,t) = \sum_{n=0}^{\infty}\left(e^{-x^2/2}C_n(x)\right)\frac{t^n}{n!} = e^{-x^2/2+2tx-t^2} = e^{-1/2(x^2-4tx+4t^2)+t^2} = e^{-1/2(x-2t)^2}e^{t^2}$$

de sorte qu'alors [37] (en calculant de la transformée de Fourier relativement à la variable x et t considéré comme paramètre) il vient :

$$\mathfrak{F}\left[e^{-x^2/2}G(x,t)\right](y) = \mathfrak{F}\left[e^{-1/2(x-2t)^2}e^{t^2}\right](y) = e^{2\boldsymbol{i}yt}e^{-y^2/2}e^{t^2} = e^{-y^2/2-2\boldsymbol{i}yt+t^2}$$

---

34. Mis à part la notation, c'est cette définition qu'utilise Wiener : la lettre « H » pouvant désigner un espace de Hilbert ou encore le hamiltonien d'un système mécanique, nous avons utilisé la lettre « C » (au lieu de la notation traditionnelle avec la lettre « H ») en référence au prénom de l'illustre mathématicien. *« Les polynômes d'Hermite probabilistes »* notés $\mathcal{C}_n(x)$ sont définis pour tout entier $n \geq 0$ par l'identité

$$\mathcal{C}_n(x) := (-1)^n e^{x^2/2}\frac{d^n}{dx^n}(e^{-x^2/2})$$

ce qui donne $C_n(x) = \sqrt{2^n}\,\mathcal{C}_n(x/\sqrt{2})$. Si nous $\mathbf{1}$ est la fonction t.q. $\mathbf{1}(x) = 1$, pour tout $x \in \mathbb{R}$, alors

$$\begin{aligned} C_{n+1}(x) &= (-1)^{n+1}e^{x^2/2}\frac{d}{dx}\left(\frac{d^n}{dx^n}(e^{-x^2/2})\right) \\ &= (-1)^{n+1}e^{x^2/2}\frac{d}{dx}\left((-1)^n e^{-x^2/2}C_n(x)\right) = \left(x - \frac{d}{dx}\right)C_n(x) \end{aligned}$$

de sorte que par récurrence il vient pour tout $n \geq 0$:

$$C_n(x) = \left(x - \frac{d}{dx}\right)^n \mathbf{1}$$

35. Wiener les appelle les fonctions d'Hermite : elles sont aussi appelées les fonctions de Gauss-Hermite.

36. Par développement de Taylor de e^{-x^2} en x (pour l'accroissement $-t$), nous avons

$$e^{-(x-t)^2} = \sum_{n=0}^{\infty}(-1)^n e^{-x^2}C_n(x)\frac{(-t)^n}{n!} = \left(\sum_{n=0}^{\infty}C_n(x)\frac{t^n}{n!}\right)e^{-x^2}$$

37. Nous utilisons le fait que $\mathfrak{F}[e^{-x^2/2}](y) = e^{-y^2/2}$.

soit encore :

$$\mathfrak{F}_{-}\Big[e^{-x^2/2}G(x,t)\Big](y) = e^{-y^2/2}e^{2y(\boldsymbol{i}t)-(it)^2} = G(y,\boldsymbol{i}t)$$

Mais nous pouvons identifier ce résultat à l'expression de la transformée de Fourier de $e^{-x^2/2}G(x,t)$ obtenue à partir de la série, ce qui nous donne (et en justifiant les opérations)

$$\sum_{n=0}^{\infty}\Big(e^{-y^2/2}C_n(y)\Big)\frac{(\boldsymbol{i}t^n)}{n!} = \sum_{n=0}^{\infty}\Big((\boldsymbol{i})^n e^{-y^2/2}C_n(y)\Big)\frac{t^n}{n!} = \sum_{n=0}^{\infty}\Big(\mathfrak{F}_{-}\Big[e^{-x^2/2}C_n(x)\Big](y)\Big)\frac{t^n}{n!}$$

□

7.4. Pour terminer cette section, nous allons définir un produit sur les fonctions très analogue à la convolution. Si f et g sont deux fonctions de la variable réelle à valeurs complexes, telles que la fonction $y \mapsto f(x+y)\overline{g(y)}$ soit intégrable pour presque partout $x \in \mathbb{R}$ (par exemple si f et g sont dans $L^2(\mathbb{R})$), alors *« la corrélation croisée »* $f \circledast g$ est la fonction définie pour tout $x \in \mathbb{R}$ et telle que

$$f \circledast g(x) := \int f(x+y)\overline{g(y)}dy = \int f(y)\overline{g(y-x)}dy$$

Si $\varphi \in L^2(\mathbb{R})$, alors $|\hat{\varphi}(y)|^2 = \hat{\varphi}\,\overline{\hat{\varphi}}$ est une fonction de $L^1(\mathbb{R})$ qui s'appelle *« la densité spectrale d'énergie »* ; comme $\mathfrak{F}_{+}[\hat{\varphi}](x) = \varphi(x)$ et que $\mathfrak{F}_{+}\left[\overline{\hat{\varphi}}\right](x) = \overline{\varphi(-x)}$ nous pouvons écrire :

$$\mathfrak{F}_{+}\left[|\hat{\varphi}|^2\right](x) = \mathfrak{F}_{+}\left[\hat{\varphi}\,\overline{\hat{\varphi}}\right](x) = \mathfrak{F}_{+}[\hat{\varphi}] * \mathfrak{F}_{+}\left[\overline{\hat{\varphi}}\right](x) = \int \varphi(y)\overline{\varphi(y-x)}dy = \varphi \circledast \varphi(x) \tag{33}$$

La fonction $\mathsf{C}(x) := \varphi \circledast \varphi(x)$ est généralement appelée la *« fonction d'autocorrélation »* du signal : $\mathsf{C}(x)$ étant dans l'image de $L^1(\mathbb{R})$ par $\mathfrak{F}_{+}[\cdot]$, c'est une fonction continue qui s'annule en $\pm\infty$; remarquons aussi que

$$|\mathsf{C}(x)| = \frac{1}{\sqrt{2\pi}}\left|\int |\hat{\varphi}(y)|^2 e^{\boldsymbol{i}xy}dy\right| \le \frac{1}{\sqrt{2\pi}}\int |\hat{\varphi}(y)|^2 dy = \mathsf{C}(0)$$

La *« fonction d'autocohérence »* du signal φ est par définition $\mathsf{C}(x)/\mathsf{C}(0)$. Lorsque le signal est normalisé – i.e. appartient à la sphère unité de $L^2(\mathbb{R})$ – alors (théorème de Plancherel) $\mathsf{C}(0) = 1$ et donc les fonctions d'autocorrélation et d'autocohérence sont égales.

Remarque 7.5. *L'identité en (33) est une version du « théorème de Wiener-Khintchine ». Nous avons vu que l'autocorellation $\varphi\circledast\varphi(y)$ est une fonction continue qui est maximale à l'origine. Par un argument analogue à l'argument classique donnant le cas d'égalité de l'inégalité de Cauchy-Schwarz, nous pouvons montrer que $x = 0$ est le seul point où le maximum de $\mathsf{C}(x)$ est atteint. Pour voir cela, considérons que ε est un réel arbitrairement donné. Alors*

$$\begin{aligned}
0 &\le \int \Big(\varphi(\tau) + \varepsilon\varphi(\tau+x)\Big)^2 d\tau > 0 \\
&= \int \varphi^2(\tau)d\tau + 2\varepsilon\int \varphi(\tau)\varphi(\tau+x)d\tau + \varepsilon^2\int \varphi^2(\tau+x)d\tau \\
&= \int \varphi^2(\tau)d\tau + 2\varepsilon\int \varphi(\tau)\varphi(\tau+x)d\tau + \varepsilon^2\int \varphi^2(\tau)d\tau
\end{aligned}$$

Nous obtenons alors $a\varepsilon^2 + b\varepsilon + a > 0$, où nous avons posé $a = \|\varphi\|^2$ et $b = 2\varphi\circledast\varphi(\tau)$: en particulier, nous avons $b^2 - 4ac \le 0$, i.e. $b/2 \le a$, ce qui donne le résultat annoncé.

8. Mesures de Radon, probabilités et variables aléatoires

8.1. Le *« théorème de représentation de Riesz-Markov »*, permet une simplification de la notion de mesure. Notons $\mathscr{M}(\mathbb{R})$ l'ensemble des *« mesures de Radon sur $\mathbb{R}$ »*, c'est-à-dire le demi-cône des formes linéaires positives sur le $\mathbb{R}$-espace vectoriel $\mathscr{C}_c(\mathbb{R})$ des fonctions réelles continues sur $\mathbb{R}$ et de support compact : en d'autres termes, $\mu \in \mathscr{M}(\mathbb{R})$ ssi c'est une forme linéaire sur $\mathscr{C}_c(\mathbb{R})$ telle que

$$\forall f \in \mathscr{C}_c(\mathbb{R}),\ f \geq 0 \Longrightarrow \mu(f) \geq 0$$

Le théorème de représentation de Riesz-Markov affirme que μ est une mesure de Radon sur $\mathbb{R}$ ssi il existe une mesure borélienne – notée abusivement μ – qui est *« positive »*, *« sigma-finie »*, *« régulière »*[38] et t.q.

$$\forall f \in \mathscr{C}_c(\mathbb{R}), \quad \mu(f) = \int f(x)\mu(dx)$$

(l'intégrale de gauche correspondant à la notation utilisée en théorie de la mesure).

8.2. Dans la suite nous identifions la mesure de Radon μ avec la mesure borélienne $\mu(dx)$ associée ; ainsi μ sera dite discrète s'il existe une suite injective de points $x_1, x_2, \ldots$ dans $\mathbb{R}$, t.q. pour tout intervalle compact I la mesure $\mu(I)$ coïncide avec $\sum_{x_k \in I} \mu\{x_k\}$: dans la suite $\mathscr{M}_{\mathrm{d}}(\mathbb{R})$ désignera l'ensemble des mesures de Radon discrètes. Lorsque $\mu\{x\} = 0$, pour tout $x \in \mathbb{R}$, la mesure μ est dite *« continue (ou sans atomes) »* : nous notons $\mathscr{M}_{\mathrm{c}}(\mathbb{R})$ l'ensemble des mesures de Radon continues. La mesure de Lebesgue $\lambda = \lambda(dx)$ – caractérisée par le fait que $\lambda([a\,;b]) = b - a$ pour tout $-\infty < a \leq b < +\infty$ – est une mesure de Radon continue ; $\mu \in \mathscr{M}_{\mathrm{c}}(\mathbb{R})$ sera dite *« absolument continue »* si pour tout borélien A, le fait que $\lambda(A) = 0$ entraine que $\mu(A) = 0$. Le *« théorème de Radon-Nikodym »* affirme alors que μ est absolument continue – et nous notons $\mu \in \mathscr{M}_{\mathrm{ac}}(\mathbb{R})$ – ssi il existe une fonction $f_\mu(x)$ dans $L^1(\mathbb{R})$ – la densité de μ – telle que pour tout $-\infty < a \leq b < +\infty$

$$\mu([a\,;b]) = \int_a^b f_\mu(x)dx$$

Enfin, $\mu \in \mathscr{M}_{\mathrm{c}}(\mathbb{R})$ sera dite *« purement singulière »* – et nous notons $\mu \in \mathscr{M}_{\mathrm{ps}}(\mathbb{R})$ – si elle est portée par un borélien qui est de mesure de Lebesgue nulle. Le théorème suivant est une conséquence du *« théorème de décomposition de Lebesgue »*.

Théorème 8.1. *Pour toute mesure $\mu \in \mathscr{M}(\mathbb{R})$ il existe un unique triplet de mesures de Radon $(\mu_{\mathrm{d}}, \mu_{\mathrm{ac}}, \mu_{\mathrm{ps}}) \in \mathscr{M}_{\mathrm{d}}(\mathbb{R}) \times \mathscr{M}_{\mathrm{ac}}(\mathbb{R}) \times \mathscr{M}_{\mathrm{ps}}(\mathbb{R})$ tel que $\mu = \mu_{\mathrm{d}} + \mu_{\mathrm{ac}} + \mu_{\mathrm{ps}}$.*

« La mesure de Hausdorff de l'ensemble triadique de Cantor » est un exemple de mesure purement singulière ; d'autres exemples classiques importants ont été introduits par Riesz (*« les produits de Riesz »*) afin de donner des exemples de probabilités continues, portées par l'intervalle unité et qui sont purement singulières[39].

---

38. Une mesure borélienne μ sur $\mathbb{R}$ est positive si $\mu(I) \geq 0$, pour tout intervalle I, elle est sigma-finie lorsque $\mu(I)$ est fini quand l'intervalle I est compact ; enfin μ est régulière si pour tout borélien A, la valeur de $\mu(A)$ coïncide à la fois avec l'infimum des $\mu(U)$, pour U ouvert contenant A, mais aussi au supremum des $\mu(K)$, pour K compact contenu dans A.

39. La pure singularité des mesures construites par Riesz est une conséquence du lemme de Riemann-Lebesgue affirmant que la transformée de Fourier d'une fonction Lebesgue intégrable est nulle en $\pm\infty$ (Erdös à utilisé cette propriété pour montrer que la *« la convolution de Bernoulli »* associée au nombre d'or est une mesure purement singulière).

8.3. Il existe aussi la notion de *« fonction absolument continue »* [40]. Pour voir cela, notons qu'il découle du *« le théorème de dérivation de Lebesgue »* que pour toute fonction réelle $f \in L^1(\mathbb{R})$ et tout $a \in \mathbb{R}$, en posant

$$F(x) := F(a) + \int_a^x f(u)du$$

nous définissons une fonction réelle $F \in \mathscr{C}(\mathbb{R})$ dérivable presque partout avec $F'(x) = f(x)$ pour presque tout $x \in \mathbb{R}$: la fonction F est appelée *« une primitive (au sens de Lebesgue) »* de la fonction f. La réciproque n'est pas vraie en général, en ce sens qu'il existe des fonctions réelles F continues sur $\mathbb{R}$ dérivables presque partout, de dérivée F' localement intégrable, mais ne coïncidant pas avec la primitive de leur dérivée (i.e. $F(a) + \int_a^x F'(u)du$ diffère de $F(x)$ sur un ensemble de mesure de Lebesgue strictement positif). L'exemple classique s'obtient à partir de *« la mesure de Hausdorff de l'ensemble triadique de Cantor »* : si nous notons $\mu(dx)$ cette mesure et que pour tout $x \in \mathbb{R}$, nous posons $F(x) := \int \mathbf{1}_{[0\,;x]}(u)\mu(du)$, alors F (i.e. *« l'escalier du diable »*) est une fonction de $\mathscr{C}(\mathbb{R})$ qui est dérivable presque partout avec $F'(x) = 0$ pour presque tout $x \in \mathbb{R}$. Ce préambule justifie la définition suivante.

Définition 8.2. (i) *: Une fonction réelle $f \in \mathscr{C}(\mathbb{R})$ dont la dérivée $f'(x)$ existe presque partout est dite « absolument continue » lorsque sa dérivée f' est identifiable à une fonction de $L^1_{\text{loc}}(\mathbb{R})$ et si de plus f est une primitive de sa dérivée en ce sens pour tout $-\infty < a \le b < +\infty$*

$$\int_a^b f'(u)du = f(b) - f(a)$$

(ii) *: La fonction complexe $f \in \mathscr{C}(\mathbb{R})$ sera dite « absolument continue » ssi sa partie réelle et sa partie imaginaire sont absolument continues.*

Notons $\Theta(x) := \mathbf{1}_{[0\,;+\infty]}(x)$ la fonction de Heaviside. Alors $x\Theta(x)$ est une fonction absolument continue dont la dérivée s'identifie à la fonction de Heaviside $\Theta(x)$ – ou plus rigoureusement à la classe de $\Theta(x)$ dans $L^1_{\text{loc}}(\mathbb{R})$. Noter que la fonction $\Theta(x)$ est dérivable presque partout et que sa dérivée s'identifie à la fonction nulle. Ainsi la fonction de Heaviside (qui est continue par morceaux) n'est clairement pas une primitive de sa dérivée.

8.4. Nous notons $\mathscr{P}(\mathbb{R})$ l'ensemble convexe des *« probabilités »* sur $\mathbb{R}$, c'est-à-dire des mesures $\mu \in \mathscr{M}(\mathbb{R})$ t.q. $\mu(\mathbb{R}) = 1$. Pour $n = 0, 1, \dots$ *« le moment d'ordre n »* d'une probabilité $\mu \in \mathscr{P}(\mathbb{R})$ est défini dès que la fonction $x \mapsto x^n$ est μ-intégrable comme la quantité

$$\boldsymbol{m}_n(\mu) := \int x^n \mu(dx)$$

Le moment d'ordre 0 de μ vaut toujours $\boldsymbol{m}_0(\mu) = 1$ et lorsqu'il existe, le moment $\boldsymbol{m}_1(\mu)$ représentant *« la moyenne de μ »*. Les moments d'ordre pair sont toujours définis dans $[0\,;+\infty]$: notons alors que si la moyenne $\boldsymbol{m}_1(\mu)$ existe, alors la variance

$$\operatorname{Var}(\mu) := \int \big(x - \boldsymbol{m}_1(x)\big)^2 \mu(dx) = \int \big(x^2 - 2x\boldsymbol{m}_1(x) + \boldsymbol{m}_1(\mu)^2\big)\mu(dx) = \boldsymbol{m}_2(\mu) - \boldsymbol{m}_1(\mu)^2$$

40. Sans entrer dans les détails, mentionnons qu'une fonction (réelle) de $L^1_{\text{loc}}(\mathbb{R})$ est absolument continue ssi sa dérivée au sens des distributions est identifiable à une mesure de Radon absolument continue (voir aussi la Proposition 8.5).

est une quantité toujours définie [41] dans $[0\,;+\infty]$; par définition *« l'écart type »* de μ est

$$\sigma(\mu)=\sqrt{\mathrm{Var}(\mu)}$$

8.5. Nous définissons maintenant *« la fonction génératrice des moments »* de $\mu\in\mathscr{P}(\mathbb{R})$ comme la fonction $\Psi_\mu(z)$ dépendant de la variable complexe z et telle que

$$\Psi_\mu(z)=\int e^{xz}\mu(dx)$$

Alors, $\Psi_\mu(z)$ est définie sur un sous-ensemble du plan complexe contenant $\boldsymbol{i}\mathbb{R}$. En écrivant le développement de Taylor de e^{zx} en $z=0$, il vient (formellement)

$$\Psi_\mu(z)=\int\left(\sum_{n=0}^{\infty}\frac{(zx)^n}{n!}\right)\mu(dx)=\sum_{n=0}^{\infty}\left(\int x^n\mu(dx)\right)\frac{z^n}{n!}=\sum_{n=0}^{\infty}\frac{\boldsymbol{m}_n(\mu)}{n!}z^n$$

Par suite, si le rayon de convergence ρ de la série entière $\sum_{n=0}^{\infty}\boldsymbol{m}_n(\mu)z^n/n!$ est positif, alors la fonction $\Psi_\mu(z)$ est holomorphe en tout point z t.q. $|z|<\rho$ avec

$$\Psi_\mu(z)=\sum_{n=0}^{\infty}\frac{\boldsymbol{m}_n(\mu)}{n!}z^n$$

L'existence [42] des moments $\boldsymbol{m}_1(\mu),\boldsymbol{m}_2(\mu),\dots$ (dans $\mathbb{R}$) est ainsi une condition nécessaire (mais non suffisante) assurant que $\Psi_\mu(z)$ soit holomorphe au voisinage de $z=0$.

8.6. Nous avons vu que la fonction génératrice des moments $\Psi_\mu(z)$ associée à une probabilité μ est définie sur un sous-ensemble du plan complexe contenant $\boldsymbol{i}\mathbb{R}$: *« la fonction caractéristique »* $\Phi_\mu(t)$ est alors définie pour tout $t\in\mathbb{R}$ de sorte que

$$\Phi_\mu(t)=\Psi_\mu(\boldsymbol{i}t)=\int e^{\boldsymbol{i}tx}\mu(dx)$$

Pour $\mu\in\mathscr{P}(\mathbb{R})$ et pour tout $t\in\mathbb{R}$, la fonction $x\mapsto e^{\pm\boldsymbol{i}tx}$ est μ-intégrable : *« la transformée de Fourier de μ »* (en analyse ou en synthèse) est notée $\mathfrak{F}_\pm[\mu]$ et définie en posant

$$\mathfrak{F}_\pm[\mu](t)=\frac{1}{\sqrt{2\pi}}\int e^{\pm\boldsymbol{i}tx}\mu(dx)\tag{34}$$

Nous utilisons ici les mêmes notations que pour la transformée de Fourier des fonctions de $L^1(\mathbb{R})$ (c.f. § 6 supra) du fait de la compatibilité qui existe entre le deux notions : en effet, si la probabilité μ est absolument continue, c'est-à-dire s'il existe $f\in L^1(\mathbb{R})$ t.q. $\mu(dx)=f(x)dx$, alors $\mathfrak{F}_\pm[\mu]=\mathfrak{F}_\pm[f]$. Notons alors que la fonction caractéristique de μ peut être exprimée en termes de transformée de Fourier, c'est-à-dire en écrivant

$$\Phi_\mu(t)=\sqrt{2\pi}\,\mathfrak{F}_-[\mu](-t)=\sqrt{2\pi}\,\mathfrak{F}_+[\mu](t)$$

Proposition 8.3. *La fonction caractéristique Φ_μ de $\mu\in\mathscr{P}(\mathbb{R})$ vérifie :*

(i) : $|\Phi_\mu(t)|\leq 1$ pour tout $t\in\mathbb{R}$ et [43] $\Phi_\mu(0)=1$ *;*

(ii) : Φ_μ est une fonction uniformément continue sur $\mathbb{R}$;

---

41. L'inégalité $\boldsymbol{m}_2(\mu)\geq\boldsymbol{m}_1(\mu)^2$ est aussi un cas particulier de l'inégalité de Jensen, affirmant que si $\mu\in\mathscr{P}(\mathbb{R})$ possède un moment d'ordre 1 et si $f(x)$ est une fonction convexe μ-intégrable, alors $f\left(\int x\mu(dx)\right)\leq\int f(x)\mu(dx)$.

42. Si $\mu(dx)=\frac{1}{\pi}dx/(1+x^2)$ est *« la probabilité de Cauchy »* alors $\boldsymbol{m}_1(\mu)$ n'est pas définie : par suite $\Psi_\mu(z)$ n'est pas holomorphe au voisinage de $z=0$; cependant sa fonction caractéristique Φ_μ est bien définie sur l'ensemble de la droite réelle avec $\Phi_\mu(t)=e^{-|t|}$.

43. La fonction caractéristique de la mesure de Dirac δ_0 concentrée en 0 est $\Phi_{\delta_0}(t)\equiv 1$.

(iii) : si le moment $\boldsymbol{m}_n(\mu)$ existent pour $n \geq 1$, alors les moments $\boldsymbol{m}_k(\mu)$ existent pour $k = 0, \ldots, n$; de plus, $\Phi_\mu(t)$ possède un développement de Taylor-Young d'ordre n en $t = 0$ soit :

$$\Phi_\mu(t) = 1 + \sum_{k=1}^{n} \frac{(\boldsymbol{i}t)^k}{k!} \boldsymbol{m}_k(\mu) + \circ(t^n) \tag{35}$$

En particulier, $\Phi_\mu(t)$ est n fois dérivable en $t = 0$ et pour tout $k = 0, \cdots, n$:

$$\frac{d^k \Phi_\mu}{dt^k}(0) = \boldsymbol{m}_k(\mu)$$

Preuve. (i)-(ii) : La partie (i) est immédiate. Pour (ii), remarquons qu'avec $t, h \in \mathbb{R}$

$$|\Phi_\mu(t+h) - \Phi_\mu(t)| = \left| \int e^{\boldsymbol{i}tx} \big(e^{\boldsymbol{i}hx} - 1\big) \mu(dx) \right| \leq \int \left| e^{\boldsymbol{i}hx} - 1 \right| \mu(dx) =: \eta(h)$$

Le *« théorème de la convergence dominée »* assure que $\eta(h) \to 0$ quand $h \to 0$.

(iii) : Remarquons d'abord que si k est un entier positif alors, le fait que x^k soit μ-intégrable entraîne que x^k est aussi μ-intégrable (si $|x| \geq 1$ alors $|x|^{k-1} \leq |x|^k$). L'existence du moment $\boldsymbol{m}_n(\mu)$ entraîne donc l'existence des moments $\boldsymbol{m}_k(\mu)$ pour $0 \leq k \leq n$. Maintenant, fixons $t \in \mathbb{R}$; par la *« formule de Taylor-Lagrange »* (d'ordre n) appliquée séparément à la partie réelle de la fonction $x \mapsto e^{\boldsymbol{i}tx}$ ainsi qu'à sa partie imaginaire, nous obtenons l'existence de deux réels $\alpha = \alpha(x)$ et $\beta = \beta(x)$ dans $[0\,;1]$ tels que

$$e^{\boldsymbol{i}tx} = \left(\sum_{k=0}^{n-1} \frac{(\boldsymbol{i}tx)^k}{k!} \right) + \frac{(\boldsymbol{i}tx)^n}{n!} \Big(\cos(\alpha tx) + \boldsymbol{i} \sin(\beta tx) \Big)$$

soit encore

$$e^{\boldsymbol{i}tx} = \left(\sum_{k=0}^{n} \frac{(\boldsymbol{i}tx)^k}{k!} \right) + \frac{(\boldsymbol{i}tx)^n}{n!} \Big(\cos(\alpha tx) + \boldsymbol{i} \sin(\beta tx) - 1 \Big) \tag{36}$$

En intégrant (36) par rapport à $\mu(dx)$, nous obtenons que

$$\Phi_\mu(t) = \sum_{k=0}^{n} \frac{(\boldsymbol{i}t)^k}{k!} \boldsymbol{m}_k(\mu) + t^n \varepsilon(t) \tag{37}$$

où nous avons posé

$$\varepsilon(t) := \int \frac{(\boldsymbol{i}x)^n}{n!} \Big(\cos(\alpha tx) + \boldsymbol{i} \sin(\beta tx) - 1 \Big) \mu(dx)$$

Or d'une part pour tout x

$$\lim_{t \to 0} \frac{(\boldsymbol{i}x)^n}{n!} \Big(\cos(\alpha tx) + \boldsymbol{i} \sin(\beta tx) - 1 \Big) = 0$$

et d'autre part pour tout $t \in \mathbb{R}$ (et pour $n \geq 3$),

$$\left| \frac{(\boldsymbol{i}x)^n}{n!} \Big(\cos(\alpha tx) + \boldsymbol{i} \sin(\beta tx) - 1 \Big) \right| \leq |x^n|$$

Par suite, le fait que $x \mapsto x^n$ soit par hypothèse intégrable (existence du moment $\boldsymbol{m}_n(\mu)$), permet de déduire – par une nouvelle application du *« théorème de la convergence dominée »* – que $\varepsilon(t) \to 0$ quand $t \to 0$: l'identité (37) démontre ainsi la validité de (iii).

□

8.7. Dans la suite nous noterons

$$\mathscr{P}_{\mathrm{d}}(\mathbb{R}) := \mathscr{M}_{\mathrm{d}}(\mathbb{R}) \cap \mathscr{P}(\mathbb{R}) \qquad \mathscr{P}_{\mathrm{c}}(\mathbb{R}) := \mathscr{M}_{\mathrm{c}}(\mathbb{R}) \cap \mathscr{P}(\mathbb{R})$$

$$\mathscr{P}_{\mathrm{ac}}(\mathbb{R}) := \mathscr{M}_{\mathrm{ac}}(\mathbb{R}) \cap \mathscr{P}(\mathbb{R}) \qquad \mathscr{P}_{\mathrm{ps}}(\mathbb{R}) := \mathscr{M}_{\mathrm{ps}}(\mathbb{R}) \cap \mathscr{P}(\mathbb{R})$$

Il est immédiat que ces ensembles sont tous des sous-ensembles convexes de $\mathscr{M}(\mathbb{R})$; le résultat suivant découle directement du Théorème 8.1.

Théorème 8.4. *Pour toute mesure $\mu \in \mathscr{P}(\mathbb{R})$ il existe trois réels $\alpha, \beta, \gamma \geq 0$ avec $\alpha + \beta + \gamma = 1$ ainsi qu'un triplet de probabilités $(\mu_{\mathrm{d}}, \mu_{\mathrm{ac}}, \mu_{\mathrm{ps}}) \in \mathscr{P}_{\mathrm{d}}(\mathbb{R}) \times \mathscr{P}_{\mathrm{ac}}(\mathbb{R}) \times \mathscr{P}_{\mathrm{ps}}(\mathbb{R})$ de sorte que*

$$\mu = \alpha\mu_{\mathrm{d}} + \beta\mu_{\mathrm{ac}} + \gamma\mu_{\mathrm{ps}}$$

et où $\beta\mu_{\mathrm{ac}} + \gamma\mu_{\mathrm{ps}}$ est la partie continue de μ ; de plus cette décomposition est unique.

8.8. Une fonction $F : \mathbb{R} \to [0\,;1]$ croissante, continue à droite et t.q. $(F(-\infty), F(+\infty)) = (0,1)$ est appelée *« une fonction de répartition »*. Toute probabilité μ sur $\mathbb{R}$ est associée à la fonction de répartition $F_\mu : \mathbb{R} \to [0\,;1]$ t.q. $F_\mu(x) = \mu(]-\infty\,;x])$: l'application $\mu \mapsto F_\mu$ ainsi définie réalise une bijection de $\mathscr{P}(R)$ sur l'ensemble convexe des fonctions de répartition. Supposons maintenant que $\mu \in \mathscr{P}_{\mathrm{ac}}(\mathbb{R})$ et notons $f_\mu \in L^1(\mathbb{R})$ la densité de μ ; alors par définition de la fonction de répartition de μ, nous savons que pour tout $-\infty < a \leq b < +\infty$

$$F_\mu(b) - F_\mu(a) = \mu([a\,;b]) = \int_a^b f_\mu(x)dx$$

Par la définition d'une fonction absolument continue (c.f. Définition 8.2) il vient :

Proposition 8.5. *$\mu \in \mathscr{P}(\mathbb{R})$ est absolument continue ssi F_μ est absolument continue.*

8.9. Nous avons déjà noté que la fonction caractéristique d'une probabilité μ absolument continue s'exprime grâce à la transformée de Fourrier de sa densité f_μ, en ce sens que $\Phi_\mu = \sqrt{2\pi}\,\mathfrak{F}_+[f_\mu]$, où $\mathfrak{F}_+[\cdot]$ est la transformée de Fourier L^1. Par suite, si nous supposons que Φ_μ est aussi dans $L^1(\mathbb{R})$, alors nous pouvons appliquer la formule d'inversion, ce qui donne pour presque tout $x \in \mathbb{R}$:

$$f_\mu(x) = \frac{1}{\sqrt{2\pi}}\mathfrak{F}_-[\Phi_\mu](x) = \frac{1}{2\pi}\int \Phi_\mu(t)e^{-itx}dt \tag{38}$$

Proposition 8.6. *Si la fonction caractéristique Φ_μ associée à $\mu \in \mathscr{P}(\mathbb{R})$ est Lebesgue intégrable alors μ est absolument continue et de densité (bornée) $f_\mu(x) = 1/\sqrt{2\pi}\mathfrak{F}_-[\Phi_\mu](x)$.*

8.10. Le cas purement discret est – d'une certaine manière – analogue au cas absolument continue. Pour voir cela considérons $\mu \in \mathscr{P}_{\mathrm{d}}(\mathbb{R})$ et commençons par supposer que le support μ est inclus dans l'ensemble $\mathbb{Z}$ des nombres entiers relatifs. Si pour tout $n \in \mathbb{Z}$ nous posons $p_n := \mu\{n\}$ alors la fonction caractéristique de μ est $\Phi_\mu(t) = \sum_n p_n e^{int}$. Il est remarquable que $\Phi_\mu : \mathbb{R} \to \mathbb{C}$ soit une fonction 2π-périodique dont les coefficients de Fourier $p = (\ldots, p_{-1}, p_0, p_1, \ldots)$ forment un vecteur probabilité sur $\mathbb{Z}$ en ce sens que $\sum_n p_n = 1$: puisque $p_n^2 \leq p_n$, il est alors immédiat que $\sum_n p_n^2 < +\infty$, ce qui signifie que $p \in \ell^2(\mathbb{Z})$. En particulier, nous pouvons dire que $\Phi_\mu \in L^2[0\,;2\pi]$ (espace de Hilbert des fonctions boréliennes 2π-périodiques et de carré Lebesgue-intégrable sur $[0\,;2\pi]$). Par la définition des coefficients de Fourier, nous obtenons ainsi que

$$p_n = \frac{1}{2\pi}\int_0^{2\pi} \Phi_\mu(t)e^{-int}dt \tag{39}$$

Dans le cas général, la probabilité discrète μ n'a aucunes raisons d'avoir son support inclus dans un réseau comme $\mathbb{Z}$: si nous supposons que le support de μ est $\{x_n\}_{n=1}^{\infty}$, alors, en notant $p_n := \mu\{x_n\}$, la fonction caractéristique Φ_μ est en général *« quasi-périodique »* et la formule des coefficients de Fourier (39) est remplacée par

$$p_n = \lim_{n} \frac{1}{2R} \int_{-R}^{R} \Phi_\mu(t) e^{-ix_n t} dt \tag{40}$$

Le résultat peut s'énoncer de manière générale (i.e. pour μ continue, discrète où mixte) de la façon suivante :

Théorème 8.7. *Soit Φ_μ la fonction caractéristique associée à une probabilité $\mu \in \mathscr{P}(\mathbb{R})$: alors*

$$(\forall x \in \mathbb{R}) \qquad \mu\{x\} = \lim_{R \to +\infty} \frac{1}{2R} \int_{-R}^{R} \Phi_\mu(t) e^{-itx} dt$$

Preuve. En utilisant le *« théorème de Fubini »*, nous obtenons que

$$\begin{aligned}
\frac{1}{2R} \int_{-R}^{R} \Phi_\mu(t) e^{-itx} dt &= \frac{1}{2R} \int_{-R}^{R} \Big(\int e^{ity} \mu(dy) \Big) e^{-ixt} dt \\
&= \frac{1}{2R} \int \Big(\int_{-R}^{R} e^{it(y-x)} dt \Big) \mu(dy) \\
&= \frac{1}{2R} \int \Big[e^{it(y-x)} \Big]_{t=-R}^{t=R} \mu(dy) \\
&= \int \left(\frac{e^{iR(y-x)} - e^{iR(y-x)}}{2iR(y-x)} \right) \mu(dy) = \int S\big(R(y-x)\big) \mu(dy)
\end{aligned}$$

où nous avons posé $S(u) = \sin(u)/u$. Or dans la limite où $R \to +\infty$, nous avons d'une part $S\big(R(y-x)\big) \to \mathbf{1}_{\{x\}}(y)$; d'autre part, la fonction $y \mapsto S\big(R(y-x)\big)$ est dominée par $\mathbf{1}_{\mathbb{R}}(y)$ ($\equiv 1$) qui est une fonction μ-intégrable (car μ est une probabilité) : la conclusion vient avec le *« théorème de la convergence dominée »*.

□

Le Théorème 8.7, démontre que la fonction caractéristique d'une probabilité $\mu \in \mathscr{P}_{\mathrm{d}}(\mathbb{R})$ est *caractéristique* de μ ; le théorème suivante généralise cela à tous les types probabilités.

Théorème 8.8 (P. Levy). *Soit μ une probabilité dans $\mathscr{P}(\mathbb{R})$: alors, pour tout $-\infty < a \leq b < +\infty$*

$$\lim_{R \to +\infty} \frac{1}{2\pi} \int_{-R}^{R} \frac{e^{-iat} - e^{-ibt}}{it} \Phi_\mu(t) dt = \mu(]a\,;b[) + \frac{1}{2}\Big(\mu\{a\} + \mu\{b\}\Big) \tag{41}$$

Preuve. Nous commençons en remarquant que

$$\left| \frac{e^{-ita} - e^{-itb}}{it} \right| = \left| \int_a^b e^{itx} dx \right| \leq |b-a|$$

Par suite en utilisant le *« théorème de Fubini »*, nous pouvons écrire que

$$\begin{aligned}
\frac{1}{2\pi} \int_{-R}^{R} \frac{e^{-ita} - e^{-itb}}{it} \Phi_\mu(t) dt &= \frac{1}{2\pi} \int_{-R}^{R} \frac{e^{-ita} - e^{-itb}}{it} \Big(\int e^{itx} \mu(dx) \Big) dt \\
&= \int \left(\frac{1}{2\pi} \int_{-R}^{R} \frac{e^{it(x-a)} - e^{it(x-b)}}{it} dt \right) \mu(dx) \\
&= \int \left(\frac{1}{2\pi} \int_{-R}^{R} \frac{\sin\big(t(x-a)\big) - \sin\big(t(x-b)\big)}{t} dt \right) \mu(dx)
\end{aligned}$$

(la fonction $t \mapsto \big(\cos t(x-a) - \cos t(x-b)\big)/t$ est impaire et donc son intégrale est nulle sur l'intervalle $[-R\,;R]$). Si nous notons $U_R(x) := \frac{1}{2\pi}\int_{-R}^{R}\frac{\sin(tx)}{t}dt$, alors

$$\int_{-R}^{R}\frac{e^{-ita}-e^{-itb}}{it}\Phi_\mu(t)dt = \int \Big(U_R(x-a)-U_R(x-b)\Big)\mu(dx)$$

Maintenant, nous pouvons écrire en posant $y = tx$

$$U_R(x) := \frac{1}{2\pi}\int_{-R}^{R}\frac{\sin(tx)}{t}dt = \frac{1}{\pi}\int_{0}^{R}\frac{\sin(tx)}{t}d = \frac{1}{\pi}\int_{0}^{xR}\frac{\sin(y)}{y}dy = \frac{\mathrm{sg}(x)}{\pi}\int_{0}^{|x|R}\frac{\sin(y)}{y}dy$$

où $\mathrm{sg}(x)$ est le signe de x (avec $\mathrm{sg}(0) = 0$). Or, avec la valeur de *« l'intégrale de Dirichlet »*

$$\lim_{R\to+\infty} U_R(x) = \frac{\mathrm{sg}(x)}{\pi}\int_{0}^{\infty}\frac{\sin(y)}{y}dy = \frac{\mathrm{sg}(x)}{2}$$

Par suite, il existe $M \geq 0$ t.q. $\|U_R\|_\infty \leq M$, pour tout $R > 0$ et donc, par une application licite du théorème de la convergence dominée, il vient finalement que

$$\lim_{R\to+\infty}\int \Big(U_R(x-a)-U_R(x-b)\Big)\mu(dx) = \mu(]a\,;b[) + \frac{1}{2}\Big(\mu\{a\}+\mu\{b\}\Big)$$

□

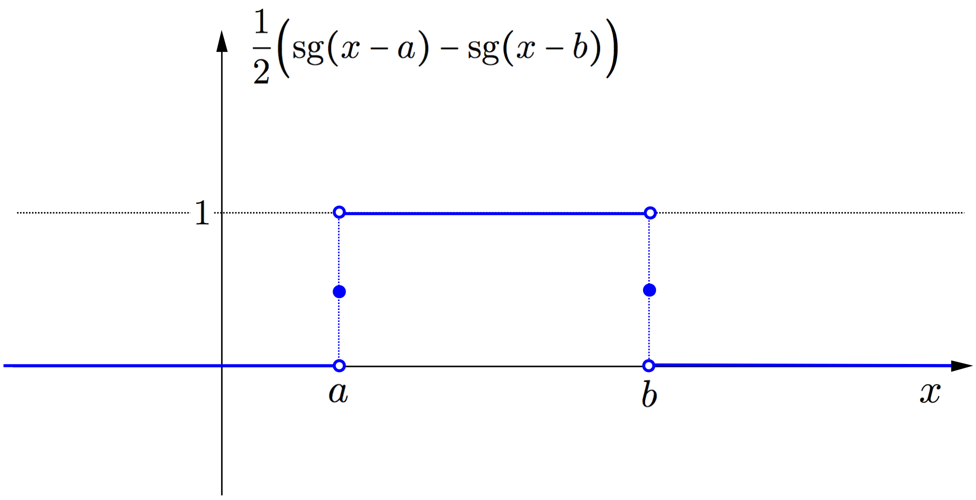


FIGURE 2. *Par définition $U(x) = \mathrm{sg}(x)/2$, où $\mathrm{sg}(x)$ est le signe de x, avec $\mathrm{sg}(0) = 0$. Nous illustrons ici la fonction $U(x-a) - U(x-b)$ afin d'éclairer la présence du terme $\big(\mu\{a\}+\mu\{b\}\big)/2$ dans (41).*

8.11. Il y a plusieurs manières de munir $\mathscr{P}(\mathbb{R})$ d'une topologie, une d'entre-elle étant de considérer les probabilités comme des distributions (voir aussi la Remarque 8.11). Cependant, dans un grand nombre d'applications, il suffit d'avoir une notion adéquate pour la convergence des suites de probabilités : la définition suivante remplit ce rôle.

Définition 8.9. *Une suite $\mu_1, \mu_2, \dots$ dans $\mathscr{P}(\mathbb{R})$ « convergent faiblement vers μ » dans $\mathscr{P}(\mathbb{R})$ ssi pour toute fonction f dans l'espace $\mathscr{C}_b(\mathbb{R})$ des fonction s réelles continues et bornées sur $\mathbb{R}$*

$$\lim_{n\to+\infty}\int f(x)\mu_n(dx) = \int f(x)\mu(dx)$$

Si la suite de probabilités $\mu_1, \mu_2, \ldots$ tend faiblement vers μ alors les fonctions caractéristiques $\Phi_{\mu_1}, \Phi_{\mu_2}, \ldots$ tendent simplement vers Φ_μ. Le théorème suivant (*« théorème de continuité de Paul Lévy »*) précise les conditions de la réciproque de cette propriété.

Théorème 8.10 (P. Lévy)**.** *Soit $\mu_1, \mu_2, \ldots$ une suite de probabilités sur $\mathbb{R}$; si les fonctions caractéristiques $\Phi_{\mu_1}, \Phi_{\mu_2}, \ldots$ convergent ponctuellement vers $\phi : \mathbb{R} \to \mathbb{C}$ continue en 0, alors les μ_n tendent faiblement vers une probabilité μ telle que $\Phi_\mu(x) = \phi(x)$, pour tout $x \in \mathbb{R}$.*

Preuve. Voir par exemple [Bas71, § V-1-4, p. 332], [Ber92, p. 51] ou [KF94, Chap. VI]. □

Remarque 8.11. *Nous avons défini directement la notion de convergence faible pour les suites de $\mathscr{P}(\mathbb{R})$, sans référence à une quelconque topologie sur $\mathscr{P}(\mathbb{R})$. Pour définir une telle topologie, rappelons que si $(E, \|\cdot\|)$ est un espace de Banach, alors son dual topologique E' peut-être muni de « la topologie $*$-faible » (c.f.* [Bre87]*) : une suite $\varphi_1, \varphi_2, \ldots$ de formes linéaires dans E' convergent alors $*$-faiblement vers φ ssi pour tout $x \in E$, la suite des $\varphi_n(x)$ tends vers $\varphi(x)$, quand $n \to +\infty$. Si nous considérons comme espace de Banach, l'espace $\mathscr{C}_b(\mathbb{R})$ des fonctions réelles continues et bornées sur $\mathbb{R}$, muni de la norme infinie $\|\cdot\|_\infty$, alors $\mathscr{P}(\mathbb{R})$ est une partie convexe du dual topologique de $\mathscr{C}_b(\mathbb{R})$: il est alors immédiat que la convergence faible d'une suite de probabilités $\mu_1, \mu_2, \ldots$ vers une probabilité μ coïncide avec la convergence $*$-faible dans le dual topologique de $\mathscr{C}_b(\mathbb{R})$ – notons au passage que $\mathscr{P}(\mathbb{R})$ est une partie convexe et $*$-faiblement fermée de $\mathscr{C}_b(\mathbb{R})$.*

8.12. Un aspect central de la théorie des probabilités (telle que formalisée en particulier par Kolmogorov [44]) porte sur la notion de *« variable aléatoire »*. Nous nous limitons ici aux propriétés élémentaires des variables aléatoires réelles. Etant donné $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, une variable aléatoire (réelle) est spécifiée par une application borélienne $X : \Omega \to \mathbb{R}$; la probabilité $\mu := \mathbb{P} \circ X^{-1}$ (définie sur les boréliens de $\mathbb{R}$) est par définition *« la loi de X »*. Alors l'espérance de X (si elle est définie) est par définition

$$\mathbb{E}(X) := \int X(\omega)\mathbb{P}(d\omega) = \int x\mu(dx)$$

Pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ borélienne, l'application $Y = f(X) = \Omega \to \mathbb{R}$ est encore une variable aléatoire dont l'espérance (si toutefois elle existe) est

$$\mathbb{E}(Y) = \int Y(\omega)\mathbb{P}(d\omega) = \int f(X(\omega))\mathbb{P}(d\omega) = \int f(x)\mu(dx)$$

Notons alors, que pour tout $k = 0, 1, \ldots$ l'espérance $\mathbb{E}(X^k)$ (si elle existe) coïncide avec le k-ème moment $\boldsymbol{m}_1(\mu)$ de la loi de X. Enfin, si l'espérance de X est bien définie, alors la variance de X est l'espérance de $(X - \mathbb{E}(X))^2$ c'est-à-dire le nombre (de $[0 ; +\infty]$)

$$\mathrm{Var}(X) = \mathbb{E}\big((X - \mathbb{E}(X))^2\big) = \int \big(X(\omega) - \mathbb{E}(X)\big)^2\big)\mathbb{P}(d\omega) = \int (x - \mathbb{E}(X))^2\mu(dx)$$

En d'autres termes, la variance $\mathrm{Var}(X)$ coïncide avec la variance $\mathrm{Var}(\mu)$ de la loi de X. De manière analogue au cas de la variance d'une probabilité nous pouvons utiliser les

44. En 1929 Kolmogorov étend de manière significative *« la loi faible des grands nombres »* sous la forme de *« la lois forte des grands nombres »* : l'énoncé même de cette lois nécessité la notion de convergence presque sûre, qui vient directement de la théorie de la mesure telle que développée à partir de la construction de l'intégrale de Lebesgue (1902).

propriétés de linéarité de l'espérance pour écrire que

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}\big(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2\big)\\ &= \mathbb{E}\big(X^2\big) - 2\mathbb{E}(X\mathbb{E}(X)) + \mathbb{E}(X)^2\\ &= \mathbb{E}\big(X^2\big) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 = \mathbb{E}\big(X^2\big) - \mathbb{E}(X)^2\end{aligned}$$

8.13. Nous nous sommes limités jusqu'ici à des variables aléatoires réelles ; par définition, une variable aléatoire complexe (i.e. à valeurs complexes) s'écrit $Z = X + \boldsymbol{i}Y$, où X et Y sont des variables aléatoire réelles ; l'espérance de Z est alors définie en posant

$$\mathbb{E}(Z) = \mathbb{E}(X) + \boldsymbol{i}\mathbb{E}(Y)$$

En particulier, si X est une variable aléatoire réelle, alors pour tout $z = a + \boldsymbol{i}b \in \mathbb{C}$:

$$e^{zX} = e^{(a+\boldsymbol{i}b)X} = e^{aX}\big(\cos(bX) + \boldsymbol{i}\sin(bX)\big)$$

est une variable aléatoire complexe. Si μ est la loi de X alors $\Psi_X(z) := \Psi_\mu(z)$ (pour z dans le domaine du plan complexe adéquat) est *« la fonction génératrice des moments de X »* ; en utilisant l'espérance des variables aléatoires complexes il vient alors

$$\Psi_X(z) = \mathbb{E}(e^{zX})$$

De même $\Phi_X(t) = \Psi_X(\boldsymbol{i}t) = \Phi_\mu(t)$ est la *« fonction caractéristique de X »* de sorte que

$$\Phi_X(t) = \mathbb{E}(e^{\boldsymbol{i}tX})$$

8.14. Il existe un certain nombre d'inégalités probabilistes qui donnent des indications importantes sur la *répartition* des probabilités. La proposition suivante nous donne deux exemples classiques de telles inégalités.

Proposition 8.12. *Soit X une variable aléatoire réelle possédant une espérance ; (i) : si X est à valeurs positives ou nulle, alors « l'inégalité de Markov » affirme que*

$$(\forall a \geq 0) \qquad \mathbb{P}\big\{Z \geq a\big\} \leq \frac{\mathbb{E}(Z)}{a} \tag{42}$$

(ii) : si X est simplement à valeurs réelles, alors « l'inégalité de Tchebychev » [45] *affirme que*

$$(\forall a \geq 0) \qquad \mathbb{P}\big\{|X - \mathbb{E}(X)| \geq a\big\} \leq \frac{\mathrm{Var}(X)}{a^2} \tag{43}$$

Preuve. (i) : La variable aléatoire X étant positive, nous avons $X(\omega) \geq a\mathbf{1}_{\{X\geq a\}}(\omega)$, pour tout $\omega \in \Omega$. L'existence de l'espérance de X nous permet alors d'écrire que

$$\mathbb{E}(X) \geq \mathbb{E}\big(a\mathbf{1}_{\{X\geq a\}}\big) = a\mathbb{P}\{X \geq a\}$$

Nous en déduisons directement *« l'inégalité de Markov »* telle que donnée en (42).

(ii) : Supposons que X est une variable aléatoire possédant une espérance $\mathbb{E}(X)$. Si la variance de X est infinie alors l'inégalité (43) est évidente. Dans le cas où la variance $\mathrm{Var}(X)$ est finie, $Z := (X - \mathbb{E}(X))^2$ est une variable aléatoire positive dont l'espérance est $\mathbb{E}(Z) = \mathrm{Var}(X)$. En appliquant l'inégalité de Markov à Z avec $a > 0$ il vient $\mathbb{P}\{Z \geq a^2\} \leq \mathbb{E}(Z)/a^2$: du fait que $Z(\omega) \geq a^2$ ssi $|X(\omega) - \mathbb{E}(X)| \geq a$, nous obtenons ainsi *« l'inégalité de Tchebychev »* telle que donnée en (42).

45. L'inégalité de Tchebychev est un résultat simple et très puissant de la théorie des probabilités. Son énoncé est dû à Bienaymée (un ami de Tchebychev) ; elle a été démontrée par Tchebychev. On la présente habituellement comme une conséquence de l'inégalité de Markov (élève de Tchebychev).

□

8.15. Deux variables (réelles)[46] X et Y sur $(\Omega, \mathcal{F}, \mathbb{P})$ sont dites indépendantes – et on note $X \perp Y$ – ssi pour tout couple (A, B) de boréliens de $\mathbb{R}$

$$\mathbb{P}\Big(\{X \in A\} \cap \{Y \in B\}\Big) = \mathbb{P}\{X \in A\}\,\mathbb{P}\{Y \in B\}$$

Proposition 8.13. *Soient X et Y deux variables aléatoires ayant chacune un moment d'ordre 2 ; alors (i) : la variable aléatoire XY possède aussi un moment d'ordre 1 et*

$$\mathbb{E}^2(XY) \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$$

(ii) : Si de plus X et Y sont indépendantes, alors

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

Preuve (esquisse). (i) : Soit $(\Omega, \mathcal{F}, \mathbb{P})$, l'espace probabilisé commun à X et Y. Le fait que X et Y possèdent chacune un moment d'ordre 2 signifie que ce sont toutes deux des fonctions de carré $\mathbb{P}$-intégrable : par suite *« d'après l'inégalité de Cauchy-Schwarz »*, XY est une fonction $\mathbb{P}$-intégrable telle que $\mathbb{E}^2(XY) \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$.

(ii) : Pour $N \geq 1$ un entier donné, soient $I_0 :=]-1/2^N\,; 1/2^N[$ et $I_k := [k/2^N\,; (k+1)/2^N[$, pour tout $k = 1, 2, \ldots$; si $I_{-k} := -I_k$, alors nous avons les *approximations* (c.f. Fig. 3)

$$X(\omega) \approx \sum_{k=-\infty}^{+\infty} \frac{k}{2^N} \mathbf{1}_{X^{-1}(I_k)}(\omega) =: \tilde{X}(\omega) \quad \text{et} \quad Y(\omega) \approx \sum_{k=-\infty}^{+\infty} \frac{k}{2^N} \mathbf{1}_{Y^{-1}(I_k)}(\omega) =: \tilde{Y}(\omega)$$

Alors, par un calcul direct (qu'il faut justifier)

$$\begin{aligned}
\mathbb{E}(\tilde{X}\tilde{Y}) &= \int \tilde{X}(\omega)\tilde{Y}(\omega)\mathbb{P}(d\omega) \\
&= \int \left(\sum_{p,q} \frac{p}{2^N} \frac{q}{2^N} \mathbf{1}_{X^{-1}(I_p) \cap Y^{-1}(I_q)}(\omega) \right) \mathbb{P}(d\omega) \\
&= \sum_{p,q} \frac{p}{2^N} \frac{q}{2^N} \mathbb{P}\Big(\{X^{-1}(I_p)\} \cap \{Y^{-1}(I_q)\} \Big) \\
&= \sum_{p,q} \frac{p}{2^N} \frac{q}{2^N} \mathbb{P}\{X^{-1}(I_p)\}\mathbb{P}\{Y^{-1}(I_q)\} = \left(\sum_p \mathbb{P}\{X^{-1}(I_p)\} \right) \left(\sum_q \frac{q}{2^N} \mathbb{P}\{Y^{-1}(I_q)\} \right)
\end{aligned}$$

soit encore

$$\mathbb{E}(\tilde{X}\tilde{Y}) = \left(\int \tilde{X}(\omega)\mathbb{P}(d\omega) \right) \left(\int \tilde{Y}(\omega)\mathbb{P}(d\omega) \right) = \mathbb{E}(\tilde{X})\mathbb{E}(\tilde{Y})$$

Nous en déduisons $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ en faisant tendre $N \to +\infty$.

□

Corollaire 8.14. *Si X et Y sont deux variables aléatoires (réelles) indépendantes, alors*

$$\Phi_{X+Y} = \Phi_X \Phi_Y$$

46. Lorsque nous considérons une famille finie, ou une suite de variables aléatoires, nous supposerons toujours qu'elles sont toutes définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

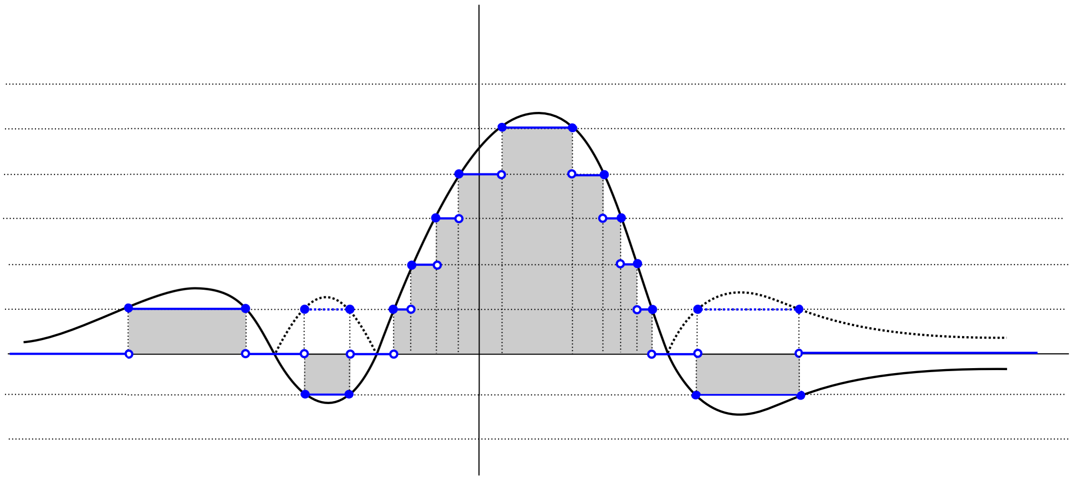


FIGURE 3. *Etant donnée $f \in L^1(\mathbb{R})$ et $N \geq 1$ un entier, nous définissons*

$$f_N := \sum_{k=-\infty}^{\infty} k/2^N \mathbf{1}_{I_k}$$

où $I_0 :=]-1/2^N\,;1/2^N[$, $I_k := [k/2^N\,;(k+1)/2^N[$, pour tout $k = 1, 2, \ldots$ et $I_{-k} := -I_k$; alors pour tout $x \in \mathbb{R}$

$$|f_N(x)| = \sum_{k=-\infty}^{\infty} |k|/2^N \mathbf{1}_{I_k}(x) \leq |f(x)|$$

et d'autre part $f_N(x) \to f(x)$ p.p., quand $N \to +\infty$.

Preuve (esquisse). Soient f et g deux fonctions réelles, boréliennes et définies sur $\mathbb{R}$; si $X \perp Y$ alors $f(X) \perp f(Y)$; le résultat s'obtient alors par une application de la Proposition 8.13 en écrivant

$$e^{\boldsymbol{i}t(X+Y)} = \cos(tX)\cos(tY) - \sin(tX)\sin(tY) + \boldsymbol{i}\Big(\cos(tX)\sin(tY) + \sin(tX)\cos(tY)\Big)$$

de sorte que

$$\begin{aligned}
\Phi_{X+Y}(t) &= \mathbb{E}\big(\cos(tX)\cos(tY)\big) - \mathbb{E}\big(\sin(tX)\sin(tY)\big) \\
&\qquad + \boldsymbol{i}\,\mathbb{E}\big(\cos(tX)\sin(tY)\big) + \boldsymbol{i}\,\mathbb{E}\big(\sin(tX)\cos(tY)\big) \\
&= \mathbb{E}\big(\cos(tX)\big)\mathbb{E}\big(\cos(tY)\big) - \mathbb{E}\big(\sin(tX)\big)\mathbb{E}\big(\sin(tY)\big) \\
&\qquad + \boldsymbol{i}\,\mathbb{E}\big(\cos(tX)\big)\mathbb{E}\big(\sin(tY)\big) + \boldsymbol{i}\,\mathbb{E}\big(\sin(tX)\big)\mathbb{E}\big(\cos(tY)\big) \\
&= \mathbb{E}\big(\cos(tX) + \boldsymbol{i}(\sin(tX)\big)\mathbb{E}\big(\cos(tY) + \boldsymbol{i}\sin(tY)\big)
\end{aligned}$$

soit encore $\Phi_{X+Y}(t) = \Phi_X(t)\Phi_Y(t)$.

□

8.16. Soit $X_1, X_2, \ldots$ une suite de variables aléatoires *« indépendante et identiquement distribuée (i.i.d.) »* : cela signifie que les X_i ont même loi et que $X_i \perp X_j$ dès que $i \neq j$; en particulier (sous réserve) $\mathbb{E}(X_1)$ et $\mathrm{Var}(X_1)$ représente respectivement l'espérance et la variance communes des X_k. Si nous notons $S_n = X_1 + \cdots + X_n$, alors il est immédiat que $\mathbb{E}(S_n) = n\mathbb{E}(X_1)$. Pour calculer le moment d'ordre 2 de S_n, nous utilisons le fait que les variables aléatoires X_k sont deux à deux indépendantes : nous obtenons

$$\begin{aligned}\mathbb{E}(S_n^2) &= \mathbb{E}\left(\sum_{p=1}^{n} X_p^2 + \sum_{1\leq p\neq q\leq n} X_p X_q\right)\\ &= \sum_{p=1}^{n}\mathbb{E}(X_p^2) + \sum_{1\leq p\neq q\leq n}\mathbb{E}(X_p)\mathbb{E}(X_q) = n\mathbb{E}(X_1^2) + (n^2-n)\mathbb{E}^2(X_1)\end{aligned}$$

Nous pouvons alors calculer la variance de S_n, en écrivant

$$\mathrm{Var}(S_n) = n\mathbb{E}(X_1^2) + (n^2-n)\mathbb{E}^2(X_1) - \Big(n\mathbb{E}(X_1)\Big)^2 = n\Big(\mathbb{E}(X_1^2) - \mathbb{E}^2(X_1)\Big) = n\mathrm{Var}(X_1)$$

Proposition 8.15. *Soit $S_n = X_1+\cdots+X_n$, où $X_1, X_2, \ldots$ forment une suite de variables i.i.d. dont l'espérance commune $\mathbb{E}(X_1)$ est bien définie ; alors $\mathbb{E}(S_n) = n\mathbb{E}(X_1)$ et $\mathrm{Var}(S_n) = n\mathrm{Var}(X_1)$.*

Remarque 8.16. *En vue de « l'argument probabiliste de Bernstein » permettant d'établir « le théorème d'approximation de Weierstrass » (c.f. § 4.2), rappelons que la variable aléatoire X est une variable de Bernoulli de paramètre α (avec $0 \leq \alpha \leq 1$) si sa loi est $\mu = (1-\alpha)\delta_0 + \alpha\delta_1$ (où nous notons δ_a la mesure de Dirac au point a). Alors X possède des moments de tous ordres, i.e. $\mathbb{E}(X^0) = 1$ et pour $k = 1, 2, \ldots$*

$$\mathbb{E}(X^k) = \int x^k \mu(dx) = (1-\alpha)\int x^k \delta_0(dx) + \alpha\int x^k \delta_1(dx) = \alpha$$

En particulier la moyenne et la variance de X valent respectivement $\mathbb{E}(X) = \alpha$ et $\mathrm{Var}(X) = \alpha - \alpha^2 = \alpha(1-\alpha)$. Notons aussi au passage que la fonction génératrice des moments Ψ_X est holomorphe dans tout le plan complexe, avec

$$\Psi_X(z) = \int e^{zx}\mu(dx) = (1-\alpha) + \alpha e^z = 1 + \sum_{k=1}^{\infty}\frac{\alpha}{k!}z^k$$

(où nous retrouvons bien que $\Psi_X(z) = \sum_{k=0}^{\infty}\mathbb{E}(X^k)z^k/k!$). Maintenant considérons que $S_n = X_1 + \cdots + X_n$, où $X_1, X_2, \ldots$ est une suite de variables i.i.d., la loi commune des X_k étant une loi de Bernoulli de paramètre α. Alors S_n est une variable aléatoire discrète de support inclus dans $\{0, \ldots, n\}$ d'espérance $\mathbb{E}(S_n) = n\alpha$ et de variance $\mathrm{Var}(S_n) = n\alpha(1-\alpha)$. Par définition S_n est une « loi binomiale de paramètre (n, α) » : on vérifie que

$$0 \leq \forall k \leq n, \qquad \mathbb{P}\{S_n = k\} = \mathrm{C}_n^k \alpha^k (1-\alpha)^{n-k}$$

8.17. Soit $S_n = X_1 + \cdots + X_n$, où $X_1, X_2, \ldots$ est une suite de variables i.i.d. centrées (i.e. d'espérances nulles). Dans le cas (particulier) où les X_k ont un variance $\mathrm{Var}(X_1)$ finie, une application de l'inégalité de Tchebychev à la variable aléatoire S_n donne

$$\mathbb{P}\left\{\left|\frac{S_n(\omega) - n\mathbb{E}(X_1)}{n}\right| < \varepsilon\right\} = 1 - \mathbb{P}\Big\{\big|S_n(\omega) - n\mathbb{E}(X_1)\big| \geq n\varepsilon\Big\} \leq 1 - \frac{\mathrm{Var}(S_n)}{(n\varepsilon)^2} = 1 - \frac{\mathrm{Var}(X_1)}{n\varepsilon^2}$$

En faisant tendre n vers l'infini, il vient :

$$\lim_{n\to+\infty}\mathbb{P}\left\{\left|\frac{S_n(\omega) - n\mathbb{E}(X_1)}{n}\right| < \varepsilon\right\} = 1$$

Nous venons de démontrer *« la loi faible des grands nombres »* dans un cas particulier ; l'hypothèse portant sur la moyenne des X_k (variables centrées) peut être levée très facilement. Par contre la loi faible des grands nombres reste valable lorsque la variance des X_k est infinie (voir [Ber92, p. 45-48] pour la démonstration générale). La *« loi forte des* *grands nombres »* (démontrée en 1929 par Kolmogorov) est d'un autre ordre de difficulté : c'est un théorème très proche du *« théorème ergodique de Birkhoff »* (démontré en 1931) et appartenant pleinement à la théorie de la mesure.

Théorème 8.17. *Soit $S_n := X_1 + \cdots + X_n$, où les X_k forment une suite i.i.d. de variables aléatoires réelles définies sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et supposées $\mathbb{P}$-intégrables ; alors :*

(i) : la loi faible des grands nombres affirme que pour tout $\varepsilon > 0$

$$\lim_{n\to+\infty} \mathbb{P}\left\{\left|\frac{S_n(\omega) - n\mathbb{E}(X_1)}{n}\right| < \varepsilon\right\} = 1$$

(ii) : la loi forte des grands nombres affirme que pour $\mathbb{P}$-presque tout $\omega \in \Omega$

$$\lim_{n\to+\infty} \frac{S_n(\omega) - n\mathbb{E}(X_1)}{n} = 0$$

8.18. A la convergence faible d'une suite de probabilités dans $\mathscr{P}(\mathbb{R})$ correspond une notion de convergence pour les suites de variables aléatoires appelée la *« convergence en loi »*.

Définition 8.18. *Soit $X_1, X_2, \ldots$ une suite de variables aléatoires et soit μ_n la lois de X_n, pour tout $n = 1, 2, \ldots$; on dit alors que la suite des X_n converge en loi vers une variable aléatoire X ssi la suite des probabilités $\mu_1, \mu_2, \ldots$ « converge faiblement » vers la lois μ de X ; on notera alors $X_n \xrightarrow{\mathscr{L}} X$, quand $n \to +\infty$.*

Le résultat suivant est un corollaire du Théorème 8.10 : il est aussi appelé *« le théorème de continuité de Levy »*.

Théorème 8.19. *$X_n \xrightarrow{\mathscr{L}} X$ ssi $\Phi_{X_n}(t) \to \Phi_X(t)$, pour tout $t \in \mathbb{R}$.*

Le *« théorème de la limite centrale »* est essentiellement un corollaire du *« théorème de continuité de Levy »*.

Théorème 8.20. *Soit $S_n := X_1 + \cdots + X_n$, où les X_k forment une suite i.i.d. de variables aléatoires réelles définies sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ supposées $\mathbb{P}$-intégrables et de variance $\sigma^2 = \mathrm{Var}(X_1)$; le théorème de la limite centrale affirme que si $\sigma < +\infty$, alors (pour $n \to +\infty$)*

$$\frac{S_n(\omega) - n\mathbb{E}(X_1)}{\sigma\sqrt{n}} \xrightarrow{\mathscr{L}} \mathrm{N}(0,1)$$

où $\mathrm{N}(0,1)$ est une variable aléatoire « normale centrée réduite » ; autrement dit :

$$\forall f \in \mathscr{C}_b(\mathbb{R}), \qquad \lim_{n\to+\infty} \int f\left(\frac{S_n(\omega) - n\mathbb{E}(X_1)}{\sigma\sqrt{n}}\right)\mathbb{P}(d\omega) = \frac{1}{\sqrt{2\pi}} \int f(t) e^{-t^2/2} dt$$

Preuve (esquisse). Il est possible de se ramener au cas particulier où les variables aléatoires X_k sont centrées et de variance $\sigma^2 = 1$. Dans ce cas, si nous notons $Y_n := S_n/\sqrt{n}$, alors le théorème de la limite centrale affirme que $Y_n \xrightarrow{\mathscr{L}} X$ où X est un variable aléatoire normale centrée et réduite ; en d'autres termes, la loi de X est la mesure gaussienne

$$\mu(dx) = \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx$$

Nous partons du fait (c.f. Corollaire 6.5) que

$$\Phi_X(t) = \sqrt{2\pi}\, \mathfrak{F}_{+}[\mu](t) = e^{-t^2/2}$$

de sorte que par le *« théorème de continuité de Levy »* il s'agit de démontrer que $\Phi_{Y_n}(t)$ converge simplement vers $e^{-t^2/2}$. La fonction caractéristique de Y_n, se calcule grâce à l'indépendance des X_k : ainsi, en utilisant le Corollaire 8.14 nous pouvons écrire que

$$\begin{aligned}\Phi_{Y_n}(t) &= \mathbb{E}\big(e^{it(X_1+\cdots+X_n)/\sqrt{n}}\big)\\ &= \mathbb{E}\big(e^{itX_1/\sqrt{n}}\big)\cdots\mathbb{E}\big(e^{itX_n/\sqrt{n}}\big) = \Big(\mathbb{E}\big(e^{itX_1/\sqrt{n}}\big)\Big)^n = \Big(\Phi_{X_1}(t/\sqrt{n})\Big)^n\end{aligned}$$

Mais les deux premiers moments de X_1 étant $\mathbb{E}(X_1) = 0$ et $\mathbb{E}(X_1^2) = 1$, il vient (c.f. équation (35) in Proposition 8.3) :

$$\Phi_{X_1}(t) = 1 + \mathbb{E}(X_1)\frac{it}{1!} + \mathbb{E}(X_1^2)\frac{(it)^2}{2!} + \circ(t^2) = 1 - \frac{t^2}{2} + \circ(t^2)$$

Finalement nous avons (dans la limite où $n \to +\infty$)

$$\Phi_{Y_n}(t) = \Phi_{X_1}^n(t/\sqrt{n}) \approx \left(1 - \frac{(t/\sqrt{n})^2}{2}\right)^n = \left(1 - \frac{t^2/2}{n}\right)^n \to e^{-t^2/2} = \Phi_X(t)$$

□

8.19. Nous avons défini une variable aléatoire (réelle) comme une application borélienne X définie sur *« un espace probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ ad-hoc »*, à valeurs dans $\mathbb{R}$; nous en avons alors déduit la loi de X comme la probabilité $\mu = \mathbb{P} \circ X^{-1}$, c'est-à-dire un élément de $\mathscr{P}(\mathbb{R})$. Cependant lorsqu'on modélise un phénomène aléatoire physique dont le résultat (obtenu par une mesure) est un nombre réel, le premier objet qu'il est possible de considérer (à partir d'un grand nombre d'expériences successives) est *« une version empirique »* de la fonction de répartition F_μ associée à μ. La question qui se pose alors est de savoir comment il est possible de *« reconstruire une version de l'espace $(\Omega, \mathcal{F}, \mathbb{P})$ »* à partir de la fonction de répartition F_μ. Pour répondre à cette question, le point de départ est de considérer *« la fonction quantile »* $Q_\mu : [0;1] \to \mathbb{R}$ t.q.

$$Q_\mu(x) = \inf\left\{x \in \mathbb{R}\ ;\ F_\mu(x) > y\right\}$$

Théorème 8.21 (du quantile)**.** *Notons Q_μ la fonction quantile associée à la fonction de repartition F_μ d'une mesure $\mu \in \mathscr{P}(\mathbb{R})$; si $\lambda(dx) = dx$ est la mesure de Lebesgue sur l'intervalle $[0;1]$, alors μ est l'image réciproque de λ par Q_μ, en ce sens que $\mu = \lambda \circ Q_\mu^{-1}$.*

Si $X : \Omega \to \mathbb{R}$ est une variable aléatoire réelle définie sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et de loi $\mu := \mathbb{P} \circ X^{-1}$ alors, d'après le théorème du quantile nous pouvons considérer $(\Omega, \mathcal{F}, \mathbb{P}) = ([0;1], \mathcal{B}, \lambda)$ (avec $\mathcal{B}$ la sigma-algèbre des boréliens de $[0;1]$) et identifier X avec $Q_\mu : [0;1] \to \mathbb{R}$. En d'autres termes, la fonction quantile Q_μ *– vue comme variable aléatoire sur $([0;1], \mathcal{B}, \lambda)$ –* peut être considérée comme une forme canonique de la variable aléatoire X. En pratique, avec un générateur de nombres aléatoires (uniforme) dans $[0;1]$ et une version empirique de F_μ, nous pouvons utiliser la fonction quantile pour obtenir un générateur aléatoire suivant la loi de X.

9. Dérivées faibles – Espaces de Sobolev

9.1. Il est difficile de dater l'émergence de *« la notion de fonction »* ; mentionnons la polémique entre d'Alembert, Euler et Daniel Bernoulli autour de la résolution de *« l'équation* *de la corde vibrante »* et dont un aspect porte sur la définition même de la notion de fonction. Pour fixer les idées, disons que lorsqu'en 1872 Weierstrass analyse les propriétés fines d'une fonction continue sur l'intervalle unité et dérivable en aucun point (*« la fonction de Weierstrass »*), il base son travail sur des définitions rigoureuses que nous utilisons toujours de nos jours. Mais à la fin du 19-ème, il apparaît de nouveaux objets mathématiques qui ressemblent à des fonctions mais ne répondent pas clairement aux définitions courantes. En particulier l'ingénieur-physicien-mathématicien Heaviside utilise des techniques de dérivation de fonctions dans des cas où les notions classiques ne s'appliquent pas. D'autre part, en mathématique, la *« théorie de la mesure »* et de *« l'intégrale de Lebesgue »* élargit encore un peu plus les contours de la notion de fonction : ainsi, une fonction peut-être définie – ou même dérivable – presque partout. L'ensemble de ces développements qui répondent à des questions aussi bien pratiques que théoriques, vont se rejoindre dans une théorie des fonctions généralisées : *« la théorie des distributions »* [47].

9.2. *« L'équation de Schrödinger »* (1926) est une équation aux dérivées partielles (EDP) qui vient se rajouter aux deux EDP classiques : *« l'équation des ondes »* et *« l'équation de la chaleur »* (voir e.g. le paragraphe « *Les trois EDP classiques* » dans [KLR98, p. 231]). A la fin des années 20, le travail de Dirac en mécanique quantique montre que les techniques utilisées par Heaviside ne sont pas de simples artifices de calcul (ce qui avait été reproché à Heaviside). C'est au milieu des années 30 que Sobolev introduit une notion de *« fonctionnelles généralisées »* [48] lui permettant d'exprimer un nouveau type de solutions des EDP [49] : *« les solutions faibles »*. Il y a en germe dans ce travail *« la notion de dérivation au sens faible »*. Nous abordons cette question par une formule d'intégration par parties très générale.

Proposition 9.1. *Soient f et g deux fonctions (complexes) définies sur $\mathbb{R}$; si f et g sont absolument continues (c.f. Définition 8.2) de dérivées respectives f' et g' (fonctions définies presque partout et localement intégrables), alors pour tout $-\infty < a \le b < +\infty$*

$$\int_a^b f(x)g'(x)dx = \Big[f(x)g(x)\Big]_a^b - \int_a^b f'(x)g(x)dx$$

Preuve. (D'après [GW90, Proposition 14.5.6].) Comme $f(x) = f(a) + \int_a^x f'(u)du$,

$$\begin{aligned}
\int_a^b f(x)g'(x)dx &= \int_a^b \Big(f(a) + \int_a^x f'(y)dy\Big) g'(x)dx \\
&= f(a)\Big[g(b) - g(a)\Big] + \int_a^b \Big(\int_a^x f'(y)dy\Big) g'(x)dx \\
\text{(Théorème de Fubini)} \quad &= f(a)\Big[g(b) - g(a)\Big] + \int_a^b f'(y)\Big(\int_b^y g'(x)dx\Big) dy \\
&= f(a)\Big[g(b) - g(a)\Big] + \int_a^b f'(y)\Big[g(b) - g(y)\Big]dy \\
&= f(a)\Big[g(b) - g(a)\Big] + \Big[f(b) - f(a)\Big]g(b) - \int_a^b f'(y)g(y)dy
\end{aligned}$$

47. Nous n'abordons pas ici la théorie des distributions
48. Comme les appelle Laurent Schwartz c.f. [Kan04].
49. En particulier l'équation des ondes.

soit encore, après simplification :

$$\int_a^b f(x)g'(x)dx = -f(a)g(a) + f(b)g(b) - \int_a^b f'(y)g(y)dy$$

□

Soit f une fonction absolument continue sur $\mathbb{R}$ et φ une fonction continument dérivable sur $\mathbb{R}$ et dont le support est inclus dans un intervalle compact $[a\,;b]$: alors φ est aussi une fonction absolument continue et par une application directe de la Proposition 9.1 (sur l'intervalle $[a\,;b]$), nous obtenons que

$$\int f(x)\varphi'(x)dx = -\int f'(x)\varphi(x)dx$$

Cette remarque nous amène à la définition suivante.

Définition 9.2. *Soit $\mathscr{D}(\mathbb{R})$ l'espace des fonctions réelles de $\mathscr{C}^\infty(\mathbb{R})$ qui sont à support compact (i.e. identiquement nulle en dehors d'un intervalle compact) : une fonction f dans $L^1_{\text{loc}}(\mathbb{R})$ est dite « dérivable au sens faible », s'il existe une fonction $g \in L^1_{\text{loc}}(\mathbb{R})$ t.q.*

$$\forall \varphi \in \mathscr{D}(\mathbb{R}), \qquad \int f(x)\varphi'(x)dx = -\int g(x)\varphi(x)dx \tag{44}$$

La fonction g est alors appelée une dérivée de f au sens faible.

Proposition 9.3. *$f \in L^1_{\text{loc}}(\mathbb{R})$ est dérivable au sens faible ssi elle est absolument continue ; dans ce cas, si g ($\in L^1_{\text{loc}}(\mathbb{R})$) est la dérivée faible de f, alors $f'(x) = g(x)$ pour presque tout $x \in \mathbb{R}$.*

La preuve de la Proposition 5.2 est basée sur la formule de l'intégration par parties généralisée donnée dans la Proposition 9.1 combinée aux deux propriétés de l'espace $\mathscr{D}(\mathbb{R})$ énoncée dans la proposition suivante.

Proposition 9.4. *Si f est une fonction de $L^1_{\text{loc}}(\mathbb{R})$, alors :*

$$(i) \ : \ \forall \varphi \in \mathscr{D}(\mathbb{R}),\ \int f(x)\varphi(x)dx = 0 \implies f(x) = 0 \quad p.p.$$

$$(ii) \ : \ \forall \varphi \in \mathscr{D}(\mathbb{R}),\ \int f(x)\varphi'(x)dx = 0 \implies \exists C \in \mathbb{R},\ f(x) = C \quad p.p.$$

Preuve. (i) : Pour tout $-\infty < a < b < +\infty$ et tout $1 \le p \le +\infty$, nous considérons l'espace $L^p[a\,;b]$ comme un sous-espace de $L^1_{\text{loc}}(\mathbb{R})$. Pour tout $f \in L^1[a\,;b]$, l'application A qui a $\varphi \in L^\infty[a\,;b]$ associe $A\varphi := \int f(x)\varphi(x)dx$, est une forme linéaire continue sur $L^\infty[a\,;b]$: si nous supposons que A s'annule sur $\mathscr{D}(\mathbb{R})$, alors il découle de la densité de $\mathscr{D}(\mathbb{R})$ dans $L^\infty[a\,;b]$ que A est identiquement nulle sur $L^\infty[a\,;b]$: cela entraîne (théorème de représentation de Riesz [50]) que $f(x) = 0$ pour presque tout $x \in [a\,;b]$.

(ii) (d'après [Bre87, p. 122-123]) : Soit $U \in \mathscr{D}(\mathbb{R})$ t.q. $\int U(x)dx = 1$; alors, pour tout $\psi \in \mathscr{D}(\mathbb{R})$, la fonction $\psi - (\int \psi(y)dy)\,U$ est dans $\mathscr{D}(\mathbb{R})$ et d'intégrale nulle ; sa primitive

$$\varphi(x) := \int_{-\infty}^x \left(\psi(t) - \left(\int \psi(y)dy\right)U(t)\right)dt \tag{45}$$

50. Nous avons déjà énoncé le théorème de Riesz-Markov (c.f. § 8.1) ; rappelons ici que le *« le Théorème de représentation de Riesz L^p pour $1 \le p < +\infty$ »* affirme que si $(X, \mathfrak{A}, \mu)$ est un espace mesuré sigma-fini et q le conjugué de p (i.e. $1 < q \le +\infty$ t.q. $1/p + 1/q = 1$) : alors, toute forme linéaire A bornée sur $L^p(X, \mathfrak{A}, \mu)$ est associée à une unique fonction f de $L^q(X, \mathfrak{A}, \mu)$ t.q. $\|f\|_q = \|A\|_p$ (norme associée) et pour laquelle $Ag = \int f(x)g(x)\mu(dx)$, pour tout $g \in L^p(X, \mathfrak{A}, \mu)$.

appartient donc elle aussi à $\mathscr{D}(\mathbb{R})$. Si f est une fonction de $L^1_{\text{loc}}(\mathbb{R})$ t.q. $\int f(x)\varphi'(x)dx = 0$, pour tout $\varphi(x) \in \mathscr{D}(\mathbb{R})$, alors dans le cas particulier de la fonction φ en (45) il vient :

$$\begin{aligned}
0 &= \int f(x)\varphi'(x)dx \\
&= \int f(x)\Big(\psi(x) - \Big(\int \psi(y)dy\Big)U(x)\Big)dx \\
&= \int f(x)\psi(x)dx - \iint \psi(y)f(x)U(x)dxdy \\
&= \int f(x)\psi(x)dx - \iint \psi(x)f(y)U(y)dxdy = \int \Big(f(x) - \int f(y)U(y)dy\Big)\psi(x)dx
\end{aligned}$$

La fonction ψ étant arbitraire dans $\mathscr{D}(\mathbb{R})$, nous pouvons déduire de la partie (i) que

$$f(x) = \int f(y)U(y)dy \quad \text{p.p.}$$

ce qui prouve (ii).

□

Preuve de la Proposition 5.2. Comme $\mathscr{D}(\mathbb{R})$ est un sous-espace de l'espace des fonctions absolument continues sur $\mathbb{R}$, il découle immédiatement de la Proposition 9.1 que si f est absolument continue de dérivée f' (définie pour presque tout x et dans $L^1_{\text{loc}}(\mathbb{R})$), alors f' est nécessairement une dérivée faible de f. Réciproquement, supposons que $g \in L^1_{\text{loc}}(\mathbb{R})$ soit une dérivée faible de f, en ce sens que pour tout $\varphi \in \mathscr{D}(\mathbb{R})$:

$$\int f(x)\varphi'(x)dx = -\int g(x)\varphi(x)dx \tag{46}$$

La primitive de g, soit $x \mapsto G(x) = f(0) + \int_0^x g(u)du$ est absolument continue et sa dérivée coincide avec g presque partout. Fixons $\varphi \in \mathscr{D}(\mathbb{R})$ arbitrairement donnée : comme G et φ sont des fonctions localement intégrables et absolument continues, nous pouvons appliquer de nouveau la formule de l'intégration par parties généralisée sur un intervalle $[a\,;b]$ contenant le support de φ, de sorte que par (46) il vient

$$\int G(x)\varphi'(x)dx = -\int g(x)\varphi(x)dx = \int f(x)\varphi'(x)dx$$

soit encore :

$$\int \big(f(x) - G(x)\big)\varphi'(x)dx = 0 \tag{47}$$

La fonction $\varphi \in \mathscr{D}(\mathbb{R})$ étant arbitraire, nous déduisons de la partie (ii) de la Proposition 9.4 combinée à (47) que $f(x) - G(x) = f(0) - G(0) = 0$ pour presque tout $x \in \mathbb{R}$.

□

9.3. Notons (par anticipation) $\mathscr{H}^1(\mathbb{R})$ l'espace des fonctions (complexes) absolument continues $f \in L^2(\mathbb{R})$ et dont la dérivée faible f' est aussi dans $L^2(\mathbb{R})$.

Proposition 9.5. *Si f et g sont deux fonctions de $\mathscr{H}^1(\mathbb{R})$, alors fg' et $f'g$ sont dans $L^1(\mathbb{R})$ et*

$$\int f(x)g'(x)dx = -\int f'(x)g(x)dx \tag{48}$$

Preuve. Si f et g sont dans $\mathscr{H}^1(\mathbb{R})$ alors fg est une fonction de $L^1(\mathbb{R})$ qui possède une version dans $\mathscr{C}(\mathbb{R})$. En identifiant fg à sa version continue, il est licite de considérer deux suites de réels $a_1, a_2, \ldots$ et $b_1, b_2, \ldots$ avec $a_n \to -\infty$ et $b_n \to +\infty$, t.q. $f(a_n)g(a_n)$ et

$f(b_n)g(b_n)$ tendent vers 0 quand $n \to +\infty$: d'après la formule d'intégration par parties (c.f. Proposition 9.1) nous avons pour tout $n \geq 1$

$$\int_{a_n}^{b_n} f(x)g'(x)dx = \Big[f(x)g(x)\Big]_{a_n}^{b_n} - \int_{a_n}^{b_n} f'(x)g(x)dx \tag{49}$$

Mais d'autre part, nous savons aussi que fg' et $f'g$ sont dans $L^1(\mathbb{R})$: les suites de fonctions $\mathbf{1}_{[a_n;b_n]}fg'$ et $\mathbf{1}_{[a_n;b_n]}f'g$ sont donc respectivement dominées par les fonctions intégrables $|fg'|$ et $|f'g|$. Partant de (49), nous pouvons appliquer (deux fois) le théorème de la convergence dominée, de sorte que par définition des suites $a_1, a_2, \ldots$ et $b_1, b_2, \ldots$, nous obtenons (48) en faisant tendre $n \to +\infty$.

□

D'après la Proposition 9.5, si $f, g \in \mathscr{H}^1(\mathbb{R})$, alors $\langle f|g'\rangle + \langle f'|g\rangle = 0$ et donc :

$$\big\langle f + f'|g + g'\big\rangle = \langle f|g\rangle + \big\langle f'|g'\big\rangle := \langle\!\langle f|g\rangle\!\rangle_{\mathscr{H}^1}$$

Il est facile de vérifier que $\langle\!\langle\cdot|\cdot\rangle\!\rangle_{1,2}$ définit un produit scalaire (hermiten) sur $\mathscr{H}^1(\mathbb{R})$: dans la suite, nous noterons $\|\cdot\|_{1,2}$ la norme associée à ce produit produit scalaire.

Proposition 9.6. *Le $\mathbb{C}$-espace vectoriel $\mathscr{H}^1(\mathbb{R})$ formé des fonctions complexes absolument continues sur $\mathbb{R}$ qui sont de carré sommable et dont la dérivée faible f' est aussi de carré sommable, est un espace de Hilbert pour le produit scalaire $\langle\!\langle\cdot|\cdot\rangle\!\rangle_{\mathscr{H}^1}$.*

Preuve. Muni du produit scalaire $\langle\!\langle\cdot|\cdot\rangle\!\rangle_{\mathscr{H}^1}$, l'espace $\mathscr{H}^1(\mathbb{R})$ est un espace pré-hilbertien (complexe) : nous devons montrer que c'est aussi un espace de Banach. Soit $f_1, f_2, \ldots$ une suite de fonctions de $\mathscr{H}^1(\mathbb{R})$ qui est de Cauchy pour la norme $\|\cdot\|_{\mathscr{H}^1}$. Si nous notons f'_n la dérivée faible de f_n, alors $\|f_n\|^2_{\mathscr{H}^1} = \|f_n\|_2^2 + \|f'_n\|_2^2$; il est donc immédiat que les fonctions f_n et f'_n forment des suites de fonctions de $L^2(\mathbb{R})$ qui sont de Cauchy pour la norme $\|\cdot\|_2$: nous notons f et g les fonctions de $L^2(\mathbb{R})$ qui sont les limites (dans $L^2(\mathbb{R})$) de ces deux suites de fonctions. Nous aurons prouvé que $\mathscr{H}^1(\mathbb{R})$ est un espace de Hilbert si nous pouvons vérifier que g est la dérivée faible de f. Comme f_n et f'_n convergent dans $L^2(\mathbb{R})$ vers f et g respectivement, d'après *« le théorème de Riesz-Fisher »* (et plus précisément la Proposition 2.10), il existe une suite croissante d'indices $n_1, n_2, \ldots$ t.q. $f_{n_k}(x)$ et $f'_{n_k}(x)$ convergent respectivement vers $f(x)$ et $g(x)$, pour presque tout $x \in \mathbb{R}$: par suite, pour toute fonction φ dans $\mathscr{D}(\mathbb{R})$ et pour tout $k \geq 1$

$$\int f_{n_k}(x)\varphi'(x)dx = -\int f'_{n_k}(x)\varphi(x)dx$$

de sorte que (*« théorème de la convergence dominée »*) dans la limite où $k \to +\infty$:

$$\int f(x)\varphi'(x)dx = -\int g(x)\varphi(x)dx$$

□

L'espace $\mathscr{H}^1(\mathbb{R})$ appartient à une famille d'espaces de Hilbert introduits par Sobolev.

Définition 9.7. *Pour $p = 1, 2, \ldots$, l'espace de Sobolev $\mathscr{H}^p(\mathbb{R})$ est le sous-espace de $L^2(\mathbb{R}) =: \mathscr{H}^0(\mathbb{R})$ formé des fonctions f qui sont p fois dérivables au sens faible et t.q. les dérivées $\partial^k f$ (dérivées k-ème) soient dans $L^2(\mathbb{R})$, pout tout $k = 1, \ldots, p$.*

Le théorème suivant s'obtient par induction :

Proposition 9.8. *La suite des espaces de Sobolev $\mathscr{H}^p(\mathbb{R})$ ($p \geq 0$ entier) est décroissante, avec*

$$L^2(\mathbb{R}) = \mathscr{H}^0(\mathbb{R}) \supset \mathscr{H}^1(\mathbb{R}) \supset \mathscr{H}^2(\mathbb{R}) \supset \cdots$$

De plus, pour $p \geq 0$ nous avons $\mathscr{H}^{p+1}(\mathbb{R}) \subset L^2(\mathbb{R}) \cap \mathscr{C}^p(\mathbb{R})$ et plus précisément :

$$f \in \mathscr{H}^{p+1}(\mathbb{R}) \implies \Big(0 \leq \forall k < p,\ \partial^k f \in L^2(\mathbb{R}) \cap \mathscr{C}^{p-k}(\mathbb{R}) \quad \textit{et} \quad \partial^p f \in \mathscr{H}^1(\mathbb{R})\Big)$$

La généralisation de la Proposition 9.6 s'énonce comme suit :

Proposition 9.9. *Pour tout entier $p \geq 1$, le $\mathbb{C}$-espace vectoriel $\mathscr{H}^p(\mathbb{R})$ est un espace de Hilbert pour le produit scalaire*

$$(f, g) \mapsto \langle\langle f|g \rangle\rangle_{\mathscr{H}^p} := \sum_{k=0}^{p} \left\langle \partial^k f | \partial^k g \right\rangle$$

Si $\|\cdot\|_{1,p}$ dénote la norme associée ce produit scalaire, il vient pour tout $f \in \mathscr{H}^p(\mathbb{R})$

$$\|f\|^2_{\mathscr{H}^p} = \sum_{k=0}^{p} \|\partial^k f\|_2^2$$

Le théorème suivant donne une caractérisation des espaces de Sobolev $\mathscr{H}^p(\mathbb{R})$ grâce à la transformée de Fourier L^2 (ce résultat nous sera très utile dans la démonstration du principe d'incertitude de Heisenberg-Gabor sur $\mathscr{H}^2(\mathbb{R})$: c.f. § 10).

Théorème 9.10. *Soit $p \geq 1$ un entier ; alors les trois propositions suivantes sont équivalentes*
- *(i) : f appartient à $\mathscr{H}^p(\mathbb{R})$;*
- *(ii) : pour tout $0 \leq k \leq p$ la dérivée k-ème $\partial^k f$ est dans $L^2(\mathbb{R})$ et $\mathfrak{F}[\partial^k f] = (\boldsymbol{i}Q)^k \mathfrak{F}[f]$;*
- *(iii) : $Q^k \mathfrak{F}[f]$ est une fonction de $L^2(\mathbb{R})$, pour tout $0 \leq k \leq p$.*

Preuve du Théorème 9.10. $(i) \implies (ii)$: Par définition, si f est dans $\mathscr{H}^p(\mathbb{R})$, alors $f \in \mathscr{C}^k(\mathbb{R})$ pour tout $0 \leq k < p$ avec $\partial^k f \in L^2(\mathbb{R})$; de plus la fonction continue $\partial^{p-1} f$ est dans $\mathscr{H}^1(\mathbb{R})$ et sa dérivée au sens faible $\partial(\partial^{p-1} f) = \partial^p f$ est aussi dans $L^2(\mathbb{R})$, ce qui démontre la première partie de (ii). Pour la deuxième partie de (ii), nous considérons pour tout entier $p \geq 1$ la proposition :

$$(\mathcal{R}_p) \ : \quad f \in \mathscr{H}^p(\mathbb{R}) \implies \Big(0 \leq \forall k \leq p, \quad Q^k \mathfrak{F}[f] \in L^2(\mathbb{R}) \quad \text{et} \quad \mathfrak{F}[\partial^k f] = (\boldsymbol{i}Q)^k \mathfrak{F}[f]\Big)$$

Noue allons démontrer par récurrence sur p que $(\mathcal{R}_p)$ est vraie pour tout $p \geq 1$. Nous commençons par démontrer le lemme suivant que est l'amorce de cette récurrence.

Lemme 9.11. *Si f est une fonction de $\mathscr{H}^1(\mathbb{R})$, alors $\partial f \in L^2(\mathbb{R})$ et $\mathfrak{F}[\partial f] = \boldsymbol{i}Q\mathfrak{F}[f]$.*

Preuve. Pour $f \in \mathscr{H}^1(\mathbb{R})$ (identifiée à sa version continue sur $\mathbb{R}$). Nous utilisons l'existence de deux suites $a_1, a_2, \ldots$ et $b_1, b_2, \ldots$ tendant respectivement vers $-\infty$ et $+\infty$ et telles que $f(a_n)$ et $f(b_n)$ tendent vers 0 quand $n \to +\infty$; en appliquant l'intégration par parties généralisée (c.f. Proposition 9.1), il vient pour tout $n \geq 1$ et $y \in \mathbb{R} \setminus \{0\}$:

$$\frac{1}{\sqrt{2\pi}} \int_{a_n}^{b_n} f(x) e^{-\boldsymbol{i}xy} dx = -\frac{1}{\boldsymbol{i}y\sqrt{2\pi}} \Big[f(x)e^{-\boldsymbol{i}xy}\Big]_{x=a_n}^{b_n} + \frac{1}{\boldsymbol{i}y} \left(\frac{1}{\sqrt{2\pi}} \int_{a_n}^{b_n} \partial f(x) e^{-\boldsymbol{i}xy} dx \right)$$

soit encore, en notant $f_n := \mathbf{1}_{[a_n ; b_n]} f$ et $\partial f_n := \mathbf{1}_{[a_n ; b_n]} \partial f$

$$\mathfrak{F}[\partial f_n](y) = \boldsymbol{i}y\mathfrak{F}[f_n](y) + \frac{1}{\sqrt{2\pi}} \Big[f(x)e^{-\boldsymbol{i}xy}\Big]_{x=a_n}^{b_n}$$

(identité qu'il est licite de considérer valable dans $L^2(\mathbb{R})$, pour tout $n = 1, 2, \ldots$). Mais dans la limite où $n \to +\infty$, nous avons [51] $f_n \to f$ et $\partial f_n \to \partial f$ dans $L^2(\mathbb{R})$: la transformée de Fourier L^2 étant un endomorphisme continu de $L^2(\mathbb{R})$, nous pouvons conclure – en utilisant l'opérateur de position – que $\mathfrak{F}_-[\partial f] = \boldsymbol{i}Q\mathfrak{F}_-[f]$: ceci conclut la preuve du lemme. □

Pour compléter la récurrence portant sur les propositions $(\mathcal{R}_p)$, le Lemme 9.11 nous autorise à considérer $p \geq 1$ t.q. $(\mathcal{R}_p)$ soit vraie. Alors d'après l'inclusion $\mathscr{H}^p(\mathbb{R}) \supset \mathscr{H}^{p+1}(\mathbb{R})$, la proposition $(\mathcal{R}_p)$ sera récurrente si nous pouvons vérifier que $\mathfrak{F}_-[\partial^{p+1} f] = (\boldsymbol{i}Q)^{p+1}\mathfrak{F}_-[f]$ dès que $f \in \mathscr{H}^{p+1}(\mathbb{R})$: or, pour une telle fonction f, nous savons (par définition) que $\partial^p f$ appartient à $\mathscr{H}^1(\mathbb{R})$ et que la dérivée faible $\partial(\partial^p f)$ de $\partial^p f$ coïncide avec $\partial^{p+1} f$; grâce au Lemme 9.11 appliqué à $\partial^p f$ et par définition de p (i.e. « *l'hypothèse de récurrence* »), nous pouvons alors écrire que

$$\mathfrak{F}_-[\partial^{p+1} f] = \mathfrak{F}_-[\partial(\partial^p f)] = \boldsymbol{i}Q\mathfrak{F}_-[\partial^p f] = \boldsymbol{i}Q(\boldsymbol{i}Q)^p\mathfrak{F}_-[f] = (\boldsymbol{i}Q)^{p+1}\mathfrak{F}_-[f]$$

$(ii) \implies (iii)$: Comme la transformée de Fourier L^2 est un endomorphisme de $L^2(\mathbb{R})$, le fait que $\partial^k f$ soit une fonction de $L^2(\mathbb{R})$ combiné à l'identité $\mathfrak{F}_-[\partial^k f] = (\boldsymbol{i}Q)^k\mathfrak{F}_-[f]$ assure que $Q^k\mathfrak{F}_-[f]$ est aussi une fonction de $L^2(\mathbb{R})$.

$(i) \impliedby (iii)$: Nous allons effectuer une récurrence sur l'entier $p \geq 1$: pour cela nous considérons la proposition :

$$(\mathcal{R}'_p) \ : \quad \Big(0 \leq \forall k \leq p, \quad Q^k\mathfrak{F}_-[f] \in L^2(\mathbb{R})\Big) \implies f \in \mathscr{H}^p(\mathbb{R})$$

Le lemme suivant que est l'amorce de la récurrence.

Lemme 9.12. *Si f est une fonction de $L^2(\mathbb{R})$ telle que $Q\mathfrak{F}_-[f]$ – i.e. $y\mathfrak{F}_-[f](y)$ – soit aussi dans $L^2(\mathbb{R})$, alors f appartient à $\mathscr{H}^1(\mathbb{R})$ (en d'autres termes la proposition $(\mathcal{R}'_1)$ est vraie).*

Preuve. Soit $f \in L^2(\mathbb{R})$ et notons $\hat{f} := \mathfrak{F}_-[f]$: nous supposons que $Q\hat{f}$ est une fonction de $L^2(\mathbb{R})$. Par définition, la fonction [52]

$$g := \boldsymbol{i}\mathfrak{F}_+\left[Q\hat{f}\right] \tag{50}$$

appartient $L^2(\mathbb{R})$: elle est localement intégrable et en posant

$$G(x) := \int_0^x g(y)dy \tag{51}$$

nous définissons une fonction absolument continue dont la dérivée faible est g. Attention, G est à priori localement intégrable mais pas nécessairement dans $\mathscr{H}^1(\mathbb{R})$; cependant, comme $f \in L^2(\mathbb{R})$, nous aurons prouvé que $f \in \mathscr{H}^1(\mathbb{R})$ si nous parvenons à démontrer

$$\exists C \in \mathbb{R}, \quad f(x) = G(x) + C \quad \text{(pour presque tout } x \in \mathbb{R}) \tag{52}$$

Pour $\varphi \in \mathscr{D}(\mathbb{R})$ arbitrairement donné, le théorème de Plancherel pour la transformée de Fourier L^2, nous donne $\langle f|\varphi'\rangle = \big\langle \mathfrak{F}_-[f]|\mathfrak{F}_-[\varphi']\big\rangle = \big\langle \hat{f}|\mathfrak{F}_-[\varphi']\big\rangle$: mais comme $\varphi' \in \mathscr{D}(\mathbb{R})$,

51. Pour $g \in L^2(\mathbb{R})$ et $g_n := \mathbf{1}_{[a_n;b_n]}g$ les $|g_n - g|^2$ forment une suite de fonctions intégrables dominée par la fonction intégrable $|g|^2$: comme $|g_n(x) - g(x)|^2 \to 0$ pour presque tout x le théorème de la convergence dominée assure que $\|g_n - g\|_2^2 \to 0$ quand $n \to 0$.

52. La définition de g proposée dans (50) vient par analogie avec la transformée de Fourier L^1 : en effet, nous savons d'après la Proposition 6.3, que si h est une fonction de $L^1(\mathbb{R}) \cap \mathscr{C}^1(\mathbb{R})$ avec $h' \in L^1(\mathbb{R})$ et de transformée de Fourier $\hat{h}$ alors $\partial\mathfrak{F}_-[h'] = \boldsymbol{i}Q\hat{h}$; si de plus $Q\hat{h}$ est aussi dans $L^1(\mathbb{R})$, alors nous pouvons appliquer la formule d'inversion, ce qui nous donne $h' = \boldsymbol{i}\mathfrak{F}_+[Q\hat{h}]$.

nous savons (c.f. Proposition 6.3[53]) que $\mathfrak{F}_-[\varphi'] = \boldsymbol{i}Q\mathfrak{F}_-[\varphi]$ et donc $\langle f|\varphi'\rangle = \left\langle \hat{f}|\boldsymbol{i}Q\mathfrak{F}_-[\varphi]\right\rangle$. Or par hypothèse, $Q\hat{f}$ est dans $L^2(\mathbb{R})$, ce qui nous permet d'écrire que $\left\langle \hat{f}|\boldsymbol{i}Q\mathfrak{F}_-[\varphi]\right\rangle = -\left\langle \boldsymbol{i}Q\hat{f}|\mathfrak{F}_-[\varphi]\right\rangle$; nous pouvons maintenant transposer la transformée de Fourier et utiliser la définition de g donnée en (50) pour obtenir finalement que

$$\langle f|\varphi'\rangle = -\left\langle \boldsymbol{i}\mathfrak{F}_+[Q\hat{f}]\,|\varphi\right\rangle = -\langle g|\varphi\rangle \tag{53}$$

Les fonctions G et φ étant toutes deux localement intégrables et absolument continues (de dérivées respectives g et φ'), nous pouvons appliquer la formule d'intégration par parties généralisée (c.f. Proposition 9.1), de sorte que pour $a < b$ tels que l'intervalle $[a\,;b]$ contienne le support de φ :

$$\int G(x)\varphi'(x)dx = \int_a^b G(x)\varphi'(x)dx = \Big[G(x)\varphi(x)\Big]_a^b - \int_a^b g(x)\varphi(x)dx = -\int g(x)\varphi(x)dx$$

Du fait que $\int g(x)\varphi(x)dx = \langle g|\varphi\rangle$, nous pouvons combiner avec (53), ce qui donne :

$$\int G(x)\varphi'(x)dx = -\langle g|\varphi\rangle = \langle f|\varphi'\rangle = \int f(x)\varphi'(x)dx$$

soit encore

$$\int \Big(f(x) - G(x)\Big)\varphi'(x)dx = 0$$

Les fonctions f et G sont dans $L^1_{\mathrm{loc}}(\mathbb{R})$ et φ est une fonction arbitraire de $\mathscr{D}(\mathbb{R})$: la partie (ii) de la Proposition 9.4 entraîne donc la validité de (52) : le lemme est démontré.

□

Pour compléter la récurrence portant sur les propositions $(\mathcal{R}'_p)$, le Lemme 9.12 nous autorise à considérer $p \geq 1$ t.q. $(\mathcal{R}'_p)$ soit vraie. Soit alors f une fonction – nécessairement dans $L^2(\mathbb{R})$ – telle que $Q^k\mathfrak{F}_-[f]$ appartienne à $L^2(\mathbb{R})$, pour tout $0 \leq k \leq p+1$: alors, d'après $(\mathcal{R}'_p)$ il est immédiat que f est dans $\mathscr{H}^p(\mathbb{R})$. Afin de montrer que f appartient à $\mathscr{H}^{p+1}(\mathbb{R})$, il nous suffit donc de montrer que la fonction continue $\partial^p f$ est dans $\mathscr{H}^1(\mathbb{R})$, soit encore d'après le Lemme 9.12 – i.e. d'après $(\mathcal{R}'_1)$ – que $Q\mathfrak{F}_-[\partial^p f]$ est dans $L^2(\mathbb{R})$. Or la fonction f étant dans $\mathscr{H}^p(\mathbb{R})$, l'implication $(i) \implies (ii)$ assure que $\mathfrak{F}_-[\partial^p f] = (\boldsymbol{i}Q)^p\mathfrak{F}_-[f]$, de sorte que par l'hypothèse faite sur f, nous avons que

$$L^2(\mathbb{R}) \ni (\boldsymbol{i}Q^{p+1})\mathfrak{F}_-[f] = \boldsymbol{i}Q\Big((\boldsymbol{i}Q)^p\mathfrak{F}_-[f]\Big) = \boldsymbol{i}Q\mathfrak{F}_-[\partial^p f]$$

Ceci achève la preuve de l'implication $(i) \impliedby (iii)$.

□

Remarque 9.13. *Les coefficients binômiaux C_p^k étant tous compris entre 1 et 2^p, pour tout entier $p \geq 1$ et tout $y \in \mathbb{R}$, nous avons la double inégalité*

$$0 \leq \forall k \leq p, \qquad y^{2k} \leq (1+y^2)^p \leq 2^p(1+y^2+\cdots+y^{2p}) \tag{54}$$

D'après le Théorème 9.10, $f \in \mathscr{H}^p(\mathbb{R})$ ssi pour tout $0 \leq k \leq p$ la fonction $y^k\hat{f}(y)$ est dans $L^2(\mathbb{R})$. D'après la double inégalité dans (54), ceci équivaut au fait que $\sqrt{(1+y^2)^p}\,\hat{f}(y)$ soit dans $L^2(\mathbb{R})$.

53. La transformée de Fourier L^1 et L^2 coïncident sur $\mathscr{S}(\mathbb{R})$; ici, φ étant supposé dans $\mathscr{D}(\mathbb{R})$, c'est aussi un élément de $\mathscr{S}(\mathbb{R})$. En fait on peut montrer que pour $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ les transformées de Fourier L^1 et L^2 sont identifiables à des fonctions égales presque partout (c.f. [GW90, Proposition 22.1.6]).

10. **Le principe d'incertitude de Heisenberg-Gabor**

10.1. L'espace de Sobolev $\mathscr{H}^2(\mathbb{R})$ joue un rôle important dans la mécanique quantique unidimensionnelle (et par extension dans d'autres cas) ; cela peut se comprendre, en particulier à cause de la présence du laplacien dans l'équation de Schrödinger. Dans ce paragraphe nous nous plaçons dans $\mathscr{H}^2(\mathbb{R})$ pour démontrer une version du *« principe d'incertitude de la théorie du signal sur $L^2(\mathbb{R})$ »* aussi appelé *« principe d'incertitude de Heisenberg-Gabor »*. Considérons pour cela que f soit une fonction de la sphère unité de $L^2(\mathbb{R})$, de sorte que (c.f. théorème de Plancherel) sa transformée de Fourier $\hat{f} = \mathfrak{F}[f]$ est aussi une fonction de la sphère unité de $L^2(\mathbb{R})$; en d'autres termes, les mesures positives finies $\mu_f(dx) = |f(x)|^2dx$ et $\mu_{\hat{f}}(dy) = |\hat{f}(y)|^2dy$ sont toutes les deux des probabilités. Rappelons alors (c.f. § 8.4) que les *« k-ème moments »* de μ_f et $\mu_{\hat{f}}$ (s'ils existent) sont respectivement

$$\boldsymbol{m}_k(\mu_f) = \int x^k|f(x)|^2dx \quad \text{et} \quad \boldsymbol{m}_k(\mu_{\hat{f}}) = \int y^k|\hat{f}(y)|^2dy$$

Par définition les moments d'ordre 0 de μ_f et $\mu_{\hat{f}}$ valent tous deux 1 ; les moments $\boldsymbol{m}_1(\mu_f)$ et $\boldsymbol{m}_1(\mu_{\hat{f}})$ (lorsqu'ils existent) représentent les *« positions moyennes »* respectives de μ_f et $\mu_{\hat{f}}$: cela explique en particulier la dénomination donnée à l'opérateur position Q défini au § 6.2, du fait que (sous réserve que f soit dans le bon sous-espace de $L^2(\mathbb{R})$) :

$$\boldsymbol{m}_1(\mu_f) = \int x|f(x)|^2dx = \int xf(x)\overline{f(x)}dx = \langle Qf|f\rangle = \langle f|Qf\rangle =: \langle f|Q|f\rangle$$

(et de même pour $\boldsymbol{m}_1(\mu_{\hat{f}})$). D'autre part, lorsque les moments d'ordre 1 existent, les *« écartd types »* de μ_f et $\mu_{\hat{f}}$ sont les quantités (finies ou infinies) suivantes :

$$\sigma(\mu_f) := \sqrt{\boldsymbol{m}_2(\mu_f) - \boldsymbol{m}_1(\mu_f)^2} \quad \text{et} \quad \sigma(\mu_{\hat{f}}) := \sqrt{\boldsymbol{m}_2(\mu_{\hat{f}}) - \boldsymbol{m}_1(\mu_{\hat{f}})^2}$$

Théorème 10.1 (Heisenberg-Gabor). *Soit f une fonction de l'espace de Sobolev $\mathscr{H}^2(\mathbb{R})$ t.q. $\int |f(x)|^2dx = 1$ et dont la transformée de Fourier $\hat{f}$ est aussi dans $\mathscr{H}^2(\mathbb{R})$; si $\mu_f(dx) = |f(x)|^2dx$ et $\mu_{\hat{f}}(dx) = |\hat{f}(x)|^2dx$, alors les écarts types $\sigma(\mu)$ et $\sigma(\mu_{\hat{f}})$ sont finis et vérifient :*

$$\sigma(\mu_f)\sigma(\mu_{\hat{f}}) \geq 1/2$$

Nous commençons par démontrer deux lemmes intermédiaires.

Lemme 10.2. *Soit f une fonction de $L^2(\mathbb{R})$ t.q. $\int |f(x)|^2dx = 1$ et dont la transformée de Fourier $\hat{f} = \mathfrak{F}[f]$ appartienne à l'espace de Sobolev $\mathscr{H}^2(\mathbb{R})$; alors les deux premiers moments $\boldsymbol{m}_1(\mu_f)$ et $\boldsymbol{m}_2(\mu_f)$ de la mesure de probabilité $\mu_f(dx) = |f(x)|^2dx$ sont finis.*

Preuve. Le fait que $\hat{f}$ soit dans $\mathscr{H}^2(\mathbb{R})$, assure que $xf(x)$ et $x\overline{f(x)}$ sont des fonctions de $L^2(\mathbb{R})$ (c.f. Théorème 9.10) : cela entraîne que $x|f(x)|^2 = (x\overline{f(x)})f(x)$ et $x^2|f(x)|^2 = (x\overline{f(x)})(xf(x))$ sont toutes deux des fonctions de $L^1(\mathbb{R})$.

□

Lemme 10.3 (Recentrage). *Soient $\mu_f(dx) = |f(x)|^2dx$ et $\mu_{\hat{f}}(dy) = |\hat{f}(y)|^2dy$, où f une fonction de l'espace de Sobolev $\mathscr{H}^2(\mathbb{R})$ t.q. $\int |f(x)|^2dx = 1$ et dont la transformée de Fourier $\hat{f}$ est aussi dans $\mathscr{H}^2(\mathbb{R})$. Si nous notons $\mu_g(dx) = |g(x)|^2dx$ et $\mu_{\hat{g}}(dy) = |\hat{g}(y)|^2dy$ où*

$$g(x) = f\big(x + \boldsymbol{m}_1(\mu_g)\big)e^{-i\boldsymbol{m}_1(\mu_{\hat{g}})x}$$

(avec $\hat{g}(y) = \mathfrak{F}[g](y)$) alors les deux probabilités μ_g et $\mu_{\hat{g}}$ sont centrées[54] *et de plus :*

$$\sigma(\mu_g) = \sqrt{\boldsymbol{m}_2(\mu_g)} = \sigma(\mu_f) \quad \textit{et} \quad \sigma(\mu_{\hat{g}}) = \sqrt{\boldsymbol{m}_2(\mu_{\hat{g}})} = \sigma(\mu_{\hat{f}})$$

Preuve. Le Lemme 10.2 assure que les deux premiers moments des probabilités μ_f et μ_g existent : par le calcul direct il vient pour $k = 1, 2$ (en posant $u = x + \boldsymbol{m}_1(\mu_f)$) :

$$\boldsymbol{m}_k(\mu_g) = \int x^k \left| f(x + \boldsymbol{m}_1(\mu_f)) e^{-i\boldsymbol{m}_1(\mu_{\hat{f}})x} \right|^2 dx = \int \left(u - \boldsymbol{m}_1(\mu_f)\right)^k |f(u)|^2 \, du$$

Par suite, $\boldsymbol{m}_1(\mu_g) = 0$ pour $k = 1$ et pour $k = 2$ il vient $\boldsymbol{m}_2(\mu_g) = \sigma(\mu_f)^2$. Par la Proposition 6.2 nous obtenons de même que $\boldsymbol{m}_1(\mu_{\hat{g}}) = 0$ avec $\boldsymbol{m}_2(\mu_{\hat{g}}) = \sigma(\mu_{\hat{f}})^2$ du fait que

$$\hat{g}(y) = \mathfrak{F}\left[f\left(x + \boldsymbol{m}_1(\mu_f)\right)e^{-i\boldsymbol{m}_1(\mu_{\hat{f}})x}\right](y) = \hat{f}(y + \boldsymbol{m}_1(\mu_{\hat{f}}))e^{-i\boldsymbol{m}_1(\mu_f)p}$$

□

Preuve de Théorème 10.1. Le fait que f et $\hat{f} = \mathfrak{F}[f]$ soient des fonctions de l'espace de Sobolev $\mathscr{H}^2(\mathbb{R})$, assure d'une part (c.f. Lemme 10.2) que les deux premiers moments des probabilités μ_f et $\mu_{\hat{f}}$ existent : en particulier les écarts type $\sigma(\mu_f)$ et $\sigma(\mu_{\hat{f}})$ sont finis. D'autre part, d'après le Théorème 9.10, l'hypothèse $f \in \mathscr{H}^2(\mathbb{R})$ assure aussi que (dans $L^2(\mathbb{R})$) nous avons $\mathfrak{F}[f'](y) = \boldsymbol{i}y\hat{f}(y)$, de sorte que par l'identité de Plancherel

$$\int |f'(x)|^2 dx = \int |\mathfrak{F}[f'](y)|^2 dy = \int y^2 |\hat{f}(y)|^2 dy = \boldsymbol{m}_2(\mu_{\hat{f}}) \tag{55}$$

Enfin (nouvelle application du Théorème 9.10) $xf(x)$ est une fonction de $L^2(\mathbb{R}) \cap \mathscr{C}^1(\mathbb{R})$: mais la fonction $f(x)$ elle même étant aussi dans $L^2(\mathbb{R}) \cap \mathscr{C}^1(\mathbb{R})$, la fonction $x|f(x)|^2$ est dans $L^1(\mathbb{R}) \cap \mathscr{C}^1(\mathbb{R})$: il existe donc deux suites de réels $a_1 a_2, \ldots$ et $b_1, b_2, \ldots$ avec $a_n \to -\infty$ et $b_n \to +\infty$, t.q. $a_n|f(a_n)|^2$ et $b_n|f(b_n)|^2$ tendent vers 0 quand $n \to +\infty$. Comme $f(x)$ est dans $\mathscr{C}^1(\mathbb{R})$, la fonction $x \mapsto |f(x)|^2 = f(x)\overline{f(x)}$ est elle aussi dans $\mathscr{C}^1(\mathbb{R})$ et de plus $(|f(x)|^2)' = 2\mathfrak{re}(f(x)\overline{f(x)'})$. Grâce à une intégration par parties (classique) sur l'intervalle $[a_n ; b_n]$, nous obtenons pour tout $n \geq 1$:

$$\int_{a_n}^{b_n} x\left(|f(x)|^2\right)' dx = \left[x|f(x)|^2\right]_{a_n}^{b_n} - \int_{a_n}^{b_n} |f(x)|^2 dx$$

Dans la limite où $n \to +\infty$, la convergence dominée entraîne que

$$\int x\left(|f(x)|^2\right)' dx = -\int |f(x)|^2 dx = -1$$

Nous pouvons alors écrire :

$$1 = \left| \int x\left(f(x)\overline{f(x)}\right)' dx \right| \leq \int \left| 2x\,\mathfrak{re}\left(f(x)\overline{f'(x)}\right) \right| dx \leq 2 \int \left| xf(x)\overline{f'(x)} \right| dx$$

Or $xf(x)$ et $f'(x)$ étant dans $L^2(\mathbb{R})$, par *« l'inégalité de Cauchy-Schwarz »* il vient :

$$1 \leq 2\left(\int x^2 |f(x)|^2 \, dx\right)^{1/2} \left(\int |f'(x)|^2 \, dx\right)^{1/2}$$

c'est-à-dire en utilisant (55) :

$$1 \leq 2\sqrt{\boldsymbol{m}_2(\mu_f)}\sqrt{\boldsymbol{m}_2(\mu_{\hat{f}})} \tag{56}$$

La conclusion découle directement du lemme recentrage (c.f. Lemme 10.3) à partir de l'équation (56) en considérant la fonction $g(x) = f(x + \boldsymbol{m}_1(\mu))e^{-i\boldsymbol{m}_1(\mu_{\hat{f}})x}$.

54. En ce sens que leurs moyennes respectives $\boldsymbol{m}_1(\mu_g)$ et $\boldsymbol{m}_1(\mu_{\hat{g}})$ sont toutes deux nulles.

□

10.2. Le cas d'égalité de l'inégalité de Heisenberg-Gabor – dans $\mathscr{S}(\mathbb{R})$ – peut se déduire simplement de l'argument précédent, du cas d'égalité de l'inégalité de Cauchy-Schwarz.

Corollaire 10.4. *Les fonctions gaussiennes $\varphi(x) = \varphi(0)\, e^{-ax^2/2}$ ($a > 0$) sont les seuls signaux (normalisés et centrés) de $\mathscr{S}(\mathbb{R})$ vérifiant le cas d'égalité de l'inégalité de Heisenberg-Gabor.*

Preuve. Si $\mu_\varphi = |\varphi(x)|^2 dx$ et $\mu_{\hat{\varphi}} = |\hat{\varphi}(y)|^2 dy$ sont les probabilités associées à un signal $\varphi(x)$ de $\mathscr{S}(\mathbb{R})$ supposé normalisé et centré, alors par une intégration par partie :

$$\int x\big(2\varphi(x)\varphi'(x)\big)dx = -\int \varphi^2(x)dx = -1$$

de sorte que (c.f. argument du Théorème 10.1)

$$\frac{1}{2} = \left|\int x\varphi(x)\varphi'(x)dx\right| \leq \left(\int x^2\varphi^2(x)dx\right)^{1/2}\left(\int \big(\varphi'(x)\big)^2 dx\right)^{1/2} = \sigma(\mu_\varphi)\,\sigma(\mu_{\hat{\varphi}}) \tag{57}$$

Nous déduisons alors de la suite des inégalités dans (57) et du cas d'égalité l'inégalité de Cauchy-Schwarz (réel), que le cas d'égalité dans l'inégalité de Heisenberg-Gabor (i.e. si $\sigma(\mu_\varphi)\,\sigma(\mu_{\hat{\varphi}}) = 1/2$) est équivalent à l'existence d'un nombre α réel t.q. $\varphi'(x) = -\alpha x\varphi(x)$ (l'introduction du signe « – » étant arbitraire) ; mais l'intégration de cette équation différentielle est simples : elle est équivalente à la quadrature $d/dx\big(\varphi(x)e^{\alpha x^2/2}\big) = 0$, ce qui donne finalement $\varphi(x) = \varphi(0)e^{-\alpha x^2/2}$. La condition $\varphi \in \mathscr{S}(\mathbb{R})$ entraîne que $\alpha > 0$.

□

10.3. Nous allons retrouver le cas d'égalité de l'inégalité de Heisenberg-Gabor de manière plus laborieuse, en étudiant directement les *« fonctions gaussiennes à écart type complexe »* (c.f. § 6.3). Pour voir cela, soient a et b sont deux réels quelconques et notons

$$G_{a+ib}(x) := e^{-(a+ib)x^2} \tag{58}$$

Pour vérifier le cas d'égalité de l'inégalité de Heisenberg-Gabor (dans le cas centré), nous considérons le signal gaussien normalisé $\varphi(x) = \varphi(0)\, G_{a+ib}(x)$ avec $a > 0$ et où $\varphi(0) = \varphi(0) > 0$ est la constant de normalisation (i.e. assurant que $\|\varphi\| = 1$), de sorte que

$$\mu_\varphi(dx) = |\varphi(x)|^2 dx = |\varphi(0)|^2\, e^{-2ax^2}\, dx$$

est une probabilité gaussienne centrée (i.e. $\boldsymbol{m}_1(\mu_\varphi) = 0$) et d'écart type

$$\sigma(\mu_\varphi) = \frac{1}{2\sqrt{a}}$$

D'autre part, nous savons (c.f. Lemme 6.4) que

$$\hat{\varphi}(y) = \mathfrak{F}[\varphi](y) = \frac{\varphi(0)}{\sqrt{2\pi}}\int e^{-(a+ib)x^2}e^{-iyx}dx = \frac{\varphi(0)}{\sqrt{2(a+ib)}}\, e^{-\frac{y^2}{4(a+ib)}}$$

Par suite la probabilité

$$\mu_{\hat{\varphi}}(dy) = |\hat{\varphi}(y)|^2 dy = \frac{|\varphi(0)|^2}{2\sqrt{a^2+b^2}}\, e^{-\frac{ay^2}{2(a^2+b^2)}}\, dy$$

est aussi gaussienne centrée et d'écart type

$$\sigma(\mu_{\hat{\varphi}}) = \sqrt{\frac{a^2+b^2}{a}}$$

Finalement nous obtenons la valeur exact du produit des écarts types de μ_φ et $\mu_{\hat{\varphi}}$, soit :

$$\sigma(\mu_\varphi)\sigma(\mu_{\hat{\varphi}}) = \frac{1}{2}\sqrt{1+\frac{b^2}{a^2}} \tag{59}$$

Ainsi, pour le signal normalisé $\varphi = \varphi(0)\, e^{-(a+\boldsymbol{i}b)x^2}$ $(a > 0)$, nous avons vérifié directement l'inégalité $\sigma(\mu_\varphi)\sigma(\mu_{\hat{\varphi}}) \geq 1/2$, le cas d'égalité ayant lieu ssi $b = 0$.

RÉFÉRENCES

[Bas71] J. Bass. *Cours de mathématiques : Topologie. Intégration. Distributions. Equations intégrales. Analyse Harmonique.* Cours de mathématiques. Masson et Cie, 1971.

[Ber92] M.A. Berger. *An Introduction to Probability and Stochastic Processes*. Springer Texts in Statistics. Springer New York, 1992.

[Bre87] H. Brezis. *Analyse fonctionnelle : théorie et applications*. Éditions Masson - Paris, 1987.

[D'A47a] Jean Le Rond D'Alembert. *Recherches sur la courbe que forme une corde tendue mise en vibration*. Histoire de l'Académie royale des sciences et des belles lettres de Berlin, T. 1, 1747.

[D'A47b] Jean Le Rond D'Alembert. *Réflexions sur la cause générale des vents*. Paris : David l'aîné, 1747.

[Fou22] J.-B. Fourier. *Théorie analytique de la chaleur*. Firmin Didot, Père et Fils - Paris, 1822.

[Gam01] T.W. Gamelin. *Complex Analysis*. Undergraduate Texts in Mathematics. Springer New York, 2001.

[GW90] C. Gasquet and P. Witomski. *Analyse de Fourier et applications : filtrage, calcul numérique, ondelettes*. Masson, 1990.

[JvN35] P. Jordan and J. von Neumann. On inner products in linear metric spaces. *Ann. of Math.*, 36 no 3 :719–723, 1935.

[Kan04] J.-M. Kantor. Mathématiques d'Est en Ouest – Théorie et pratique : l'exemple des distributions. *Gazette de la SMF*, 100 :33–43, 2004.

[KF94] A.N. Kolmogorov and S. Fomine. *Éléments de la théorie des fonctions et de l'analyse fonctionnelle*. ELLIPSE (Traduit de Éd. Mir en russe), (1994).

[KLR98] J. P. Kahane and P. G. Lemarié-Rieusset. *Séries de Fourier et Ondelettes*. Cassini - Paris, 1998.

[Loc94] G. Lochak. *La Géométrisation de la physique*. Flammarion, Paris, 1994.

[LS89] L. Lusternik and V. Sobolev. *Précis d'analyse fonctionnelle*. Mathématiques (Mir). Mir, 1989.

[Oli16a] E. Olivier. Mécanique quantique II : opérateurs non bornés. *Bull. d'Inf. App. & App.*, 104 :115–153, 2016.

[Oli16b] E. Olivier. Mécanique quantique III : de de Broglie à Schrödinger. *Bull. d'Inf. App. & App.*, 105 :171–209, 2016.

[Oli17] E. Olivier. Mécanique quantique IV : Du principe de Fermat à l'intégrale de chemin de Feynman. *Bull. d'Inf. App. & App.*, 106 :–, 2017.

[RSN55] F. Riesz and B. SZ.-Nagy. *Leçons d'analyse fonctionnelle*. Gauthiers-Villars, 1955.

[Rud64] W. Rudin. *Principles of mathematical analysis*. International series in pure and applied mathematics. McGraw-Hill, 1964.

[ST89] M. Samuelides and L. Touzillier. *Analyse fonctionnelle*. Éditions Cépaduès, 1989.

[ST90] M. Samuelides and L. Touzillier. *Analyse harmonique*. Collection La Chevêche. Cépadues-Éditions, 1990.

[VN55] J. Von Neumann. *Mathematical Foundations of Quantum Mechanics*. Investigations in physics. Princeton University Press, 1955.

[Wie04] N Wiener. *Fourier Transform and Certain of it Applications*. Cambridge Mathematical Library. Dover Publications Inc. New York (First published by The Cambridge Univ. Pr. in 1933), 2004.

(2016) B.I.A.A. no **103** BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
ISSN 0291-5413 (Université d'Aix-Marseille)

Les fondements de la théorie de la relativité générale [extrait]

Albert EINSTEIN

Résumé. – Traduction [1] d'un extrait de « *Die Grundlage der allgemeinen Relativitätstheorie* » par A. Einstein, publié dans « *Annalen der Physik. IV. Folge. Band 49.* » (Traduction par Jean-Michel KNIPPEL).

La théorie décrite dans ce qui suit, représente probablement la généralisation la plus large de la théorie, désignée actuellement comme « la théorie de la relativité » ; cette dernière, je la nomme par la suite « générale », pour la distinguer de la première, « la théorie de la relativité restreinte », que je supposerai connue. La généralisation de la théorie de la relativité a été très facilitée par la forme qui a été donnée à la théorie de la relativité restreinte par Minkowski, un mathématicien qui a révélé clairement l'équivalence formelle des coordonnées spatiales et des coordonnées temporelles, et qui a été exploitée dans la construction de la théorie. Les outils mathématiques nécessaires à la théorie de la relativité générale se trouvent dans le « calcul différentiel et intégral » qui repose sur les recherches de Gauss, Riemann et Christoffel sur des variétés non-euclidiennes et celles apportées par Ricci et Levi-Civita dans un système et déjà appliquées à des problèmes de la physique théorique. J'ai mis, dans la section B du présent essai, tout ce qui nous est nécessaire ; le physicien, qui n'est pas au fait des outils mathématiques supposés connus, trouvera une aide développée de manière simple et transparente, si bien que l'étude de la littérature mathématique pour la compréhension du présent traité ne sera pas nécessaire. Enfin, je suis reconnaissant à mon ami, le mathématicien Grossmann, qui par son aide m'a épargné, non seulement, l'étude de la littérature mathématique spécifique, mais aussi m'a assisté lors de la recherche des équations du champ de gravitation.

1. On trouve la traduction de M. Solovine des trois articles historiques d'Einstein sur les fondements de la relativité générale aux éditions Gabay (réimpression de 2009) ; notons que l'introduction ici traduite, ne figure pas dans cette édition.

1916. № 7.

ANNALEN DER PHYSIK.

VIERTE FOLGE. BAND 49.

1. *Die Grundlage der allgemeinen Relativitätstheorie;*

von A. Einstein.

Die im nachfolgenden dargelegte Theorie bildet die denkbar weitgehendste Verallgemeinerung der heute allgemein als „Relativitätstheorie" bezeichneten Theorie; die letztere nenne ich im folgenden zur Unterscheidung von der ersteren „spezielle Relativitätstheorie" und setze sie als bekannt voraus. Die Verallgemeinerung der Relativitätstheorie wurde sehr erleichtert durch die Gestalt, welche der speziellen Relativitätstheorie durch Minkowski gegeben wurde, welcher Mathematiker zuerst die formale Gleichwertigkeit der räumlichen Koordinaten und der Zeitkoordinate klar erkannte und für den Aufbau der Theorie nutzbar machte. Die für die allgemeine Relativitätstheorie nötigen mathematischen Hilfsmittel lagen fertig bereit in dem „absoluten Differentialkalkül", welcher auf den Forschungen von Gauss, Riemann und Christoffel über nichteuklidische Mannigfaltigkeiten ruht und von Ricci und Levi-Civita in ein System gebracht und bereits auf Probleme der theoretischen Physik angewendet wurde. Ich habe im Abschnitt B der vorliegenden Abhandlung alle für uns nötigen, bei dem Physiker nicht als bekannt vorauszusetzenden mathematischen Hilfsmittel in möglichst einfacher und durchsichtiger Weise entwickelt, so daß ein Studium mathematischer Literatur für das Verständnis der vorliegenden Abhandlung nicht erforderlich ist. Endlich sei an dieser Stelle dankbar meines Freundes, des Mathematikers Grossmann, gedacht, der mir durch seine Hilfe nicht nur das Studium der einschlägigen mathematischen Literatur ersparte, sondern mich auch beim Suchen nach den Feldgleichungen der Gravitation unterstützte.

BIAA

Bulletin d'Informatique Approfondie & Applications
Computation - Information
Volumes 2016

103 - 104 - 105

Aix Marseille université

Publication trimestrielle, gratuite, de l'Université d'Aix Marseille
Dépôt légal : juin 2016 - http://www.biaa.eu - ISSN 0291 - 5413

Couverture : dessin de Michel Avezard (Zevar) interprété par la MMIAGe (1985-1987)

(2016) B.I.A.A. no **104** BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
ISSN 0291-5413 (Université d'Aix-Marseille)

ÉDITORIAL : Jacques Arsac, astronome et informaticien

Maurice NIVAT [1]

Résumé. – (N.D.L.R.) : Jacques Arsac est décédé en début 2014. Maurice Nivat nous autorise à diffuser le texte qu'il a écrit pour ses obsèques et qui est paru le 21 janvier 2014 dans "Un blog d'informaticien(ne)s de la Société Informatique de France" (http ://binaire.blog.lemonde.fr).

Jacques Arsac était avant tout un homme généreux. Astronome, ayant découvert l'informatique et ce qu'elle pouvait apporter à l'astronomie et aux astronomes il n'a eu de cesse d'en répandre l'usage d'abord auprès de ses collègues astronomes puis auprès de tous ceux, nombreux, qui avaient besoin de calculs.

C'est René de Possel qui a créé l'institut de Programmation mais, au vu et au su de tout le monde, c'est Jacques Arsac qui en était l'âme, le moteur, l'homme orchestre qui faisait tout car il y avait tout à faire pour savoir un peu ce qu'était la programmation avant d'en diffuser l'enseignement, pour délimiter un peu le champ de la nouvelle discipline qui en France a pris le nom d'Informatique alors qu'aux Etats-Unis, son pays natal, elle est restée science des calculateurs (computers).

Comme il n'y avait rien, Jacques Arsac se battait toute la journée et je pense une partie de la nuit, pour écrire des compilateurs, faciliter le travail de tous ceux qui voulaient utiliser ce que nous appelions une machine à calculer, pour jeter les bases d'un enseignement de la programmation, pour recruter et former des gens qui puissent l'enseigner, pour attirer des étudiants vers cette nouvelle discipline, pour que les autres disciplines se poussent un peu pour faire de la place à cette nouvelle venue. Il ne cherchait rien pour lui-même, il négligeait de publier beaucoup des travaux qu'il faisait, il voulait seulement que ça marche et que l'informatique se développe, il faisait tout pour et il a réussi.

Je ne peux m'empêcher de penser quand je pense à Jacques Arsac aux moines défricheurs du onzième ou du douzième siècle. A vrai dire il en avait la foi chrétienne, ardente et exigeante, qui a accompagné toute sa vie d'homme et de scientifique.

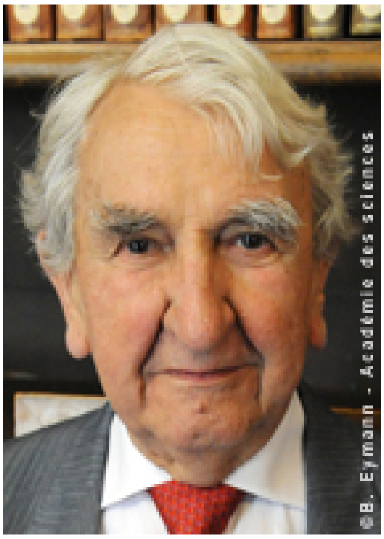
Jacques Arsac croyait tellement à ce qu'il faisait, à l'avenir de sa discipline qu'il s'est investi pendant dix ans pour en transporter l'enseignement au lycée, payant là aussi beaucoup de sa personne pour adapter les enseignements, les matériels et les logiciels et pour recruter et former des professeurs capables d'enseigner l'informatique quelque soit leur discipline d'origine (c'était un point sur lequel il insistait beaucoup, que tous devaient être capables d'apprendre et d'enseigner l'informatique). Jacques Arsac a mis dans cette expérience lycéenne tout son énergie, sa force de travail, sa foi en l'avenir et il a perdu : en 1995, je crois un décret est venu mettre fin à dix ans d'expérience de l'enseignement de l'informatique au lycée, à vrai dire on n'a jamais su pourquoi. Jacques Arsac en a été profondément meurtri.

Les dernières années de sa vie ont été essentiellement occupées à une réflexion sur la science et sa foi chrétienne qui l'ont amené à écrire plusieurs livres et à fonder

1. Membre du conseil scientifique de la SIF (Société Informatique de France).

une association des scientifiques chrétiens aujourd'hui bien vivante.

Jacques Arsac a été avant tout un homme d'action, surtout soucieux de faire partager son savoir, dont l'action a permis à l'informatique de se développer rapidement à l'université de Paris, et un homme de foi, foi en la science, foi dans le progrès que pouvait apporter ces insupportables machines à calculer qui jusqu'en 1981 et l'avènement des PCs étaient vraiment délicates à manier pour des performances ridicules à côté de la moindre tablette d'aujourd'hui. Un des tous premiers en France, il a cru en ces machines qu'il manipulait en virtuose, si imparfaites fussent-elles, et il a annoncé la place qu'elles allaient prendre dans nos vies. Tous les informaticiens et utilisateurs de machines actuelles lui sont redevables de sa vision prophétique.



Jacques Arsac

(2016) B.I.A.A. no **104** BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
ISSN 0291-5413 (Université d'Aix-Marseille)

Le nombre d'or, π et la pyramide de Khéops : une analyse arithmétique

Christian FAIVRE [1 2]

Résumé. – Le nombre d'or et π sont présents partout dans la grande pyramide de Khéops ainsi que dans la coudée (l'unité de mesure des anciens Egyptiens). En effet, de nombreuses proportions (issues de mesures de la grande pyramide) donnent des valeurs très proches de π ou bien du nombre d'or. Le but de cet article sera d'expliquer ces diverses relations.

1. Deux approximations de π

Nous allons débuter cette étude par deux approximations de π (dont l'une au moins est très connue) qui permettront de mettre en évidence la méthodologie générale qui sera explicitée par la suite.

Tout le monde connait la célèbre approximation $\frac{355}{113}$ de π d'Adrien Métius. C'est une fraction simple mais remarquablement proche de π. En fait les six premières décimales de $\frac{355}{113}$ sont exactement les mêmes que celles de π. En effet

$$\frac{355}{113} = 3.14159292\cdots$$

alors que

$$\pi = 3.14159265\cdots$$

Donnons un autre exemple (toujours une approximation de π) due au génial mathématicien Indien S. Ramanujan mort prématurément à l'âge de 33 ans. L'approximation de Ramanujan de π est

$$\sqrt[4]{\frac{2143}{22}}$$

Cette approximation est encore meilleure que celle d'Adrien Métius car cette fois les huit premières décimales sont les mêmes que celles de π :

$$\sqrt[4]{\frac{2143}{22}} = 3.1415926525\cdots$$

alors que

$$\pi = 3.1415926535\cdots \tag{1}$$

Il est naturel (à la lumière de ces deux exemples) de se demander comment on peut trouver de telles approximations. Est-ce simplement en cherchant plus ou moins au hasard ou bien y-a-t-il un moyen plus systématique d'y parvenir ? Cela pourrait être envisageable pour l'approximation d'Adrien Métius mais semble difficilement crédible pour celle de Ramanujan (comment penser à une approximation faisant intervenir la racine quatrième d'une fraction). C'est ce que nous allons voir dans la section suivante.

1. Université d'Aix-Marseille
2. christian.faivre@univ-amu.fr

2. Le développement en fraction continue

En fait il existe un outil pour trouver de telles approximations : c'est le développement en fraction continue. Celui-ci est l'outil de base pour trouver les meilleures fractions qui approchent un nombre donné. Le début du développement en fraction continue de π est (voir l'annexe A pour une introduction aux fractions continues) :

$$\pi = [3, 7, 15, 1, 292, \dots]$$

On constate que le quatrième quotient partiel est un nombre assez élevé (292), en fait un nombre inhabituellement grand pour un quatrième quotient partiel. Donc la fraction $[3, 7, 15, 1]$ va être une bonne approximation de π. Calculons donc cette fraction. On a

$$[3, 7, 15, 1] = 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1}}}$$

Donc

$$\begin{aligned} [3, 7, 15, 1] &= 3 + \cfrac{1}{7 + \cfrac{1}{16}} \\ &= 3 + \frac{16}{7 \times 16 + 1} \\ &= 3 + \frac{16}{113} \\ &= \frac{339 + 16}{113} \\ &= \frac{355}{113} \end{aligned}$$

c'est-à-dire précisément l'approximation d'Adrien Métius. Il est intéressant de regarder les autres fractions que l'on déduit du début du développement en fraction continue de π, c'est-à-dire

$$[3, 7], \qquad [3, 7, 15], \qquad [3, 7, 15, 1, 292]$$

On a

$$[3, 7] = 3 + \frac{1}{7}, \qquad [3, 7, 15] = 3 + \cfrac{1}{7 + \cfrac{1}{15}}$$

et

$$[3, 7, 15, 1, 292] = 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1 + \cfrac{1}{292}}}}$$

Après calcul, on trouve

$$[3,7]=\frac{22}{7},\qquad [3,7,15]=\frac{333}{106},\qquad [3,7,15,1,292]=\frac{103993}{33102}$$

(voir l'annexe A4 pour une méthode de calcul plus rapide de ces fractions). La fraction $\frac{22}{7}$ est une fraction classique très simple proche de π (cela peut se deviner à cause du quotient partiel 15 modérément grand) puisque

$$\frac{22}{7}=3.1428\cdots$$

Quant à $\frac{333}{106}$, c'est une fraction du même ordre de complexité que $\frac{355}{113}$ mais moins bonne que celle d'Adrien Métius puisque

$$\frac{333}{106}=3.141509\cdots$$

(quatre décimales exactes au lieu de six avec A. Métius). Quant à la fraction $\frac{103993}{33102}$, c'est une approximation meilleure que celle d'Adrien Métius puisque

$$\frac{103993}{33102}=3.1415926530\cdots$$

(neuf décimales exactes avec π, voir (1)) mais c'est une fraction nettement plus complexe. Ainsi, on peut dire que l'approximation d'Adrien Métius réalise le meilleur compromis approximation-complexité.

Après avoir constaté la présence d'un grand quotient partiel dans le début du développement de π, il est assez naturel de se demander si le même phénomène se produit dans les puissances de π. En ce qui concerne l'approximation de Ramanujan, il faut regarder le développement en fraction continue de π^4. On a

$$\pi^4=[97,2,2,3,1,16539,1,\ldots]$$

Là encore, on remarque un très grand quotient partiel (16539) parmi les tous premiers quotients partiels (c'est le cinquième quotient partiel) et donc la fraction $[97,2,2,3,1]$ sera une excellente approximation de π^4 i.e.

$$\pi^4\simeq[97,2,2,3,1]$$

Mais

$$[97,2,2,3,1]=\frac{2143}{22}$$

après un calcul élémentaire comme plus haut, d'où l'approximation

$$\pi^4\simeq\frac{2143}{22}$$

et donc

$$\pi\simeq\sqrt[4]{\frac{2143}{22}}$$

Il est à noter qu'un phénomène aussi net de grand quotient partiel (parmi les tous premiers quotients partiels) ne donne un résultat vraiment intéressant que pour π et π^4 (et dans une moindre mesure pour π^3) au moins si l'on examine les dix premières puissances de π comme le montrent les développements en fraction continue :

$$\begin{aligned}
\pi &= [3, 7, 15, 1, 292, \ldots] \\
\pi^2 &= [9, 1, 6, 1, 2, 47, 1, \ldots] \\
\pi^3 &= [31, 159, 3, 7, \ldots] \\
\pi^4 &= [97, 2, 2, 3, 1, 16539, 1, \ldots] \\
\pi^5 &= [306, 50, 1, 4, 60, 1, \ldots] \\
\pi^6 &= [961, 2, 1, 1, 3, \ldots] \\
\pi^7 &= [3020, 3, 2, 2, 3, 2, \ldots] \\
\pi^8 &= [9488, 1, 1, 7, 1, 1, \ldots] \\
\pi^9 &= [29809, 10, 14, 1, 9, \ldots] \\
\pi^{10} &= [93648, 21, 15, 1, 4, 2, \ldots]
\end{aligned}$$

Le développement de π^3 montre donc que 31 est une bonne approximation de π^3. En fait

$$\pi^3 = 31,0063\cdots$$

Il est très raisonnable d'imaginer que le grand quotient partiel 292 dans le développement en fraction continue de π a dû inciter Ramanujan à regarder si un phénomène analogue avait lieu dans les puissances de π. La mise en évidence claire de ce phénomène dans π^4 l'a donc conduit à proposer l'approximation de π, soit

$$\sqrt[4]{\frac{2143}{22}}$$

3. A la recherche de relations simples entre π et le nombre d'or

Rappelons que le nombre d'or (noté ici G) est égal à :

$$G = \frac{1+\sqrt{5}}{2}$$

On va essayer ici de mettre en évidence des relations entre π et G en utilisant la même méthodologie, c'est-à-dire des développements en fraction continue et la présence de grands quotients partiels dans le début du développement.

3.1. **Relation entre π et G^2.** La première relation sera une relation entre π et G^2. Si on regarde le développement en fraction continue de π/G^2, on a

$$\frac{\pi}{G^2} = [1, 5, 2175, 2, 8, 60, \ldots]$$

On voit un grand quotient partiel (2175), donc la fraction $[1, 5]$ sera une bonne approximation de π/G^2 :

$$\frac{\pi}{G^2} \simeq [1, 5]$$

Mais

$$[1, 5] = 1 + \frac{1}{5} = \frac{6}{5}$$

donc on en déduit

$$\pi \simeq \frac{6}{5} G^2$$

A titre de vérification, comparons les développements décimaux de π et $\frac{6}{5}G^2$:

$$\pi = 3.14159\cdots \simeq 3.1416$$

alors que

$$\frac{6}{5}G^2 = 3.14164\cdots$$

3.2. **Relation entre $\sqrt{G}$ et π.** Essayons maintenant de trouver une relation entre π et $\sqrt{G}$. Si on regarde le développement en fraction continue de $\pi\sqrt{G}$ on a

$$\pi\sqrt{G} = [3, 1, 259, 1, 13, \ldots]$$

La présence du grand quotient partiel 259 implique donc que

$$\pi\sqrt{G} \simeq [3, 1]$$

Comme

$$[3, 1] = 3 + \frac{1}{1} = 4$$

on a donc

$$\pi\sqrt{G} \simeq 4$$

d'où

$$\sqrt{G} \simeq \frac{4}{\pi}$$

A titre de vérification, comparons ici aussi les développements décimaux de $\sqrt{G}$ et $\frac{4}{\pi}$.

$$\sqrt{G} = 1.2720\cdots \simeq 1.272$$

alors que

$$\frac{4}{\pi} = 1.2732\cdots \simeq 1.273$$

3.3. **Autres relations en entre π et G.** Nous avons également cherché s'il existait d'autres relations du type

$$\pi \simeq \frac{p}{q}G^\alpha$$

où $\frac{p}{q}$ est une fraction simple. Remarquons que les deux approximations obtenues précédemment, c'est-à-dire

$$\pi \simeq \frac{6}{5}G^2 \qquad \text{et} \qquad \sqrt{G} \simeq \frac{4}{\pi}$$

correspondent respectivement à $\alpha = 2$ et $\alpha = -\frac{1}{2}$ puisque

$$\sqrt{G} \simeq \frac{4}{\pi}$$

est équivalente à

$$\pi \simeq \frac{4}{\sqrt{G}} = 4\,G^{-1/2}$$

Nous avons étudié cela pour toutes les valeurs α suivantes

$$-3 \quad -2 \quad -1 \quad -\frac{1}{2} \quad -\frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{2} \quad 1 \quad 2 \quad 3$$

Comme plus haut nous avons utilisé le développement en fraction continue de π/G^α.

En conclusion cela n'a rien donné, sauf pour les valeurs déjà citées plus haut, c'est-à-dire $\alpha = 2$ et $\alpha = -\frac{1}{2}$. Par exemple pour $\alpha = 3$ on obtient

$$\frac{\pi}{G^\alpha} = [0, 1, 2, 1, 6, 1, 2, 1, 1, 8, \ldots]$$

4. Quelques applications

4.1. π, G et la valeur de la coudée. L'unité de mesure des anciens Egyptiens est la coudée qui vaut approximativement 52 cm. Il existe d'ailleurs une règle en bois d'une longueur d'une coudée exposée au musée du Louvre à Paris (figure 1). Cette règle date du pharaon Toutankhamon (qui a régné entre 1336–1327 av. JC, 18^{e} dynastie) et a appartenu à Mâya son ministre des finances. Cette règle a une longueur indiquée de 52.3 cm. Il est à

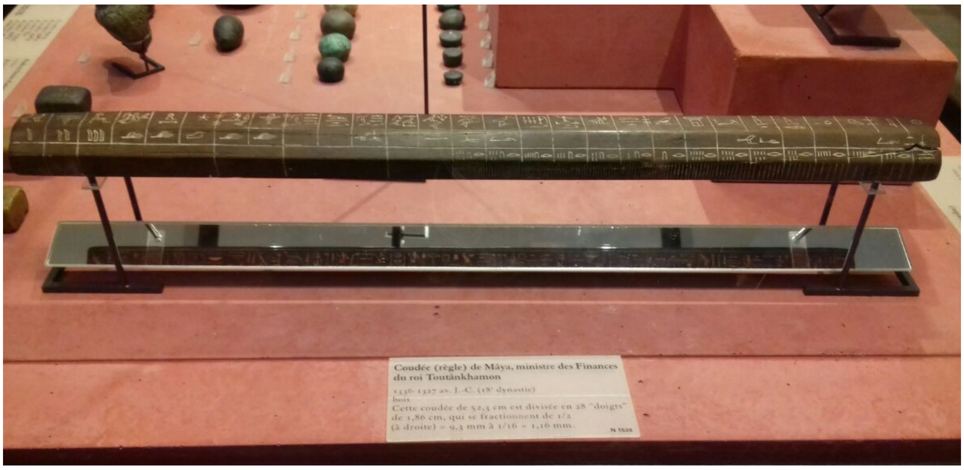


FIGURE 1. Une coudée égyptienne (Musée du Louvre).

noter que la valeur de la coudée a légèrement varié au cours de l'histoire. Essayons de déterminer la valeur qu'elle avait lors de la construction de la pyramide de Khéops.

Tous les auteurs s'accordent sur les dimensions idéales de la pyramide : 280 coudées pour la hauteur et 440 coudées pour le côté de sa base. Les mesures réelles des quatre cotés de sa base faites par le Survey Department of Egypt donne les valeurs suivantes (en mètres)

$$230.391 \quad 230.357 \quad 230.253 \quad 230.454$$

La meilleure estimation de la valeur théorique du coté est alors de prendre la moyenne de ces quatre valeurs, ce qui donne 230.3637. Si on divise cette valeur par 440, on obtient une approximation précise de la valeur de la coudée égyptienne et cela donne 0.5236 en fait plus précisément

$$0.5235539\cdots$$

Ainsi 52.36 cm semble être une estimation très précise de la coudée de la pyramide de Khéops. Certains auteurs (comme par exemple J.P Lauer) optent pour la valeur 52.36 cm pour la coudée alors que d'autres (comme par exemple J. Dorner) ont proposé 52.35 cm. C'est d'ailleurs la valeur utilisée par l'architecte G. Dormion ([Dor04] p.48) dans son livre sur la pyramide de Khéops.

La coudée n'est pas une unité propre aux anciens Egyptiens mais se retrouve aussi chez les bâtisseurs de cathédrales. On emploie aussi l'empan et le pied ainsi que la palme et la paume (voir figure 2). Toutes ces mesures sont censées être dans les proportions du nombre d'or, c'est-à-dire que le pied est égal au produit de G par l'empan, et la coudée est égale au produit de G par le pied i.e.

$$\text{pied} = G \times \text{empan}, \quad \text{coudée} = G \times \text{pied}$$

Ainsi

$$\text{coudée} = G^2 \times \text{empan}$$

Or la valeur standard de l'empan est de 20 cm. Avec cette valeur cela donne pour la coudée (en centimètres)

$$20G^2 = 52.3606\cdots$$

c'est-à-dire 52.36 cm avec une très bonne précision. Il semblerait donc évident que la coudée égyptienne est reliée au nombre d'or. Mais d'une manière étonnante cette coudée est aussi reliée à π car

$$\frac{\pi}{6} \simeq 0.5236$$

En effet

$$\frac{\pi}{6} = 0.5235987\cdots$$

donc une coudée vaut approximativement $\pi/6$ mètres. Mais ceci peut s'expliquer par la relation

$$\pi \simeq \frac{6}{5} G^2$$

que l'on a déjà vue. En effet, on a alors

$$\frac{\pi}{6} \simeq \frac{G^2}{5}$$

d'où

$$\frac{\pi}{6} \simeq \frac{20\,G^2}{100}$$

comme $20\,G^2$ est approximativement la valeur de la coudée (en centimètres), on a donc que $\frac{\pi}{6}$ est approximativement la valeur de la coudée en mètres.

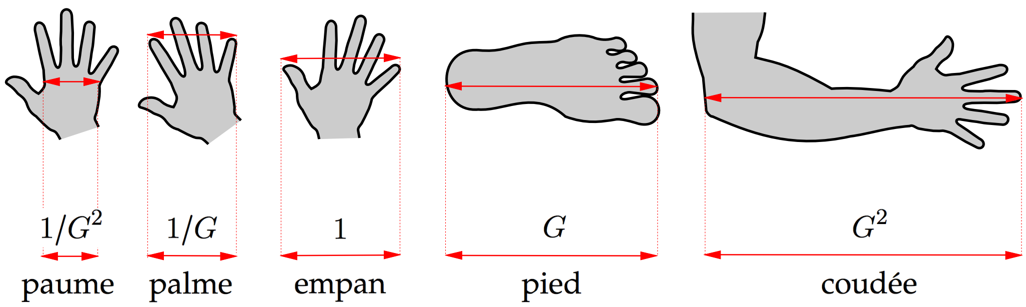


FIGURE 2. Les unités basées sur le corps humain

4.2. **Une autre relation célèbre.** Il existe aussi une autre relation étonnante reliant π, le nombre d'or et la coudée, c'est

$$\pi - G^2 \simeq \text{la coudée en mètres}$$

En effet

$$\pi - G^2 = 0.523558\cdots$$

Mais cette relation peut encore se déduire facilement de

$$\pi \simeq \frac{6}{5}G^2.$$

En effet, on en déduit alors

$$\begin{aligned}\pi - G^2 &\simeq \frac{6}{5}G^2 - G^2\\ &\simeq \frac{1}{5}G^2\end{aligned}$$

Mais comme on l'a déjà vu plus haut $\frac{1}{5}G^2$ est une approximation de la valeur de la coudée en mètres :

$$\frac{1}{5}G^2 = \frac{20\,G^2}{100}$$

Ainsi

$$\pi - G^2 \simeq \text{la coudée en mètres}$$

4.3. **La pente de la pyramide de Khéops.** Comme on l'a déjà dit, la plupart des auteurs s'accordent sur les dimensions initiales de la pyramide : la base est un carré de 440 coudées de côté et la hauteur est égale à 280 coudées. Tout au moins ce sont très vraisemblablement ces dimensions que les anciens Egyptiens avaient en tête lorsqu'ils ont bâti cette pyramide (notons cependant qu'il n'existe aucun document mentionnant les dimensions initiales de la pyramide de Khéops). Ce qui donne une pente (pour les faces) égale à

$$\frac{280}{220}$$

c'est-à-dire après simplification

$$\frac{14}{11}$$

Si la pyramide est construite suivant le nombre d'or, sa pente devrait être égale à $\sqrt{G}$ car ce serait la pente dans le triangle rectangle de côtés 1, $\sqrt{G}$ et d'hypoténuse G (voir figure 3). Nous rappelons que le nombre d'or vérifie la relation

$$G^2 = G + 1$$

et que c'est cette relation qui montre que le triangle de côtés 1, $\sqrt{G}$ et d'hypoténuse G est bien rectangle (la pente est ici le rapport entre le côté opposé et le côté adjacent). Or il se trouve que la fraction $14/11$ est assez proche de $\sqrt{G}$ puisque

$$\frac{14}{11} \simeq 1.2727\cdots, \qquad \sqrt{G} \simeq 1.2720\cdots$$

En fait ce n'est pas une coincidence et le lien entre $14/11$ et $\sqrt{G}$ est encore plus évident si on regarde le développement en fraction continue de $\sqrt{G}$. On a alors

$$\sqrt{G} = [1, 3, 1, 2, 11, 3, 5, 1, \ldots]$$

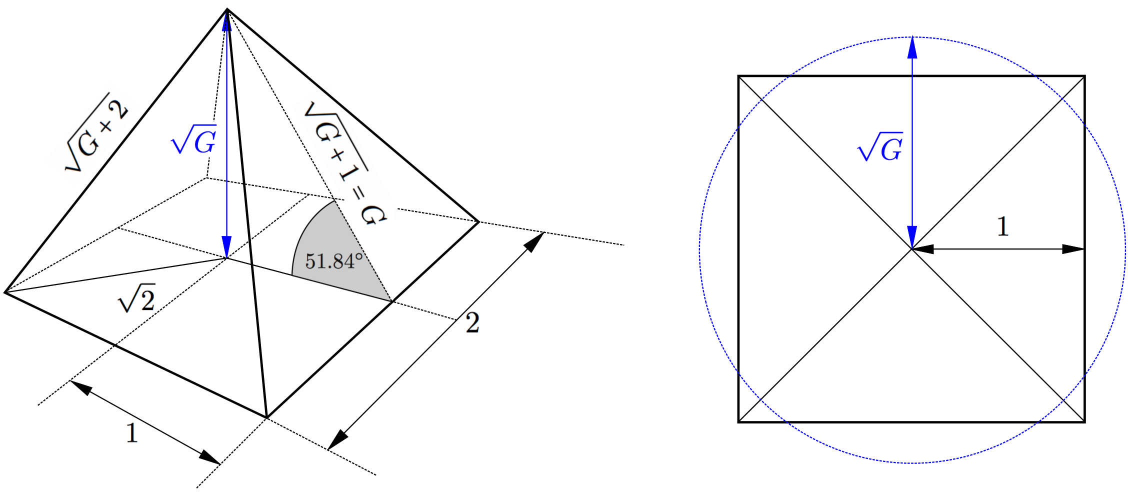


FIGURE 3. Proportions suivant le nombre d'or

Les premières fractions déduites de ce développement sont

$$[1,3] = \frac{4}{3}, \qquad [1,3,1] = \frac{5}{4}, \qquad [1,3,1,2] = \frac{14}{11}$$

On voit donc que $14/11$ est la troisième réduite du développement en fraction continue de $\sqrt{G}$. La réduite suivante approche mieux mais la fraction est moins simple :

$$[1,3,1,2,11] = \frac{159}{125}$$

Notons que $4/3$ est la pente utilisée dans la pyramide de Khéphren et que $5/4$ est celle de la pyramide de Mykérinos.

4.4. **La pyramide du Louvre.** Voici la méthode utilisée par Pei (l'architecte de la pyramide du Louvre) pour déterminer la hauteur de sa pyramide. Si la base de la pyramide est un carré de côté a alors on considère le cercle ayant même périmètre que le carré (voir figure 3). Le rayon de ce cercle est la valeur qu'il faut prendre pour la hauteur de la pyramide. Dans son livre [Goy90, p. 88], l'égyptologue G. Goyon donne cet argument faussement attribué à Hérodote. En fait c'est simplement un argument géométrique pour faire apparaitre une pente égale à $\frac{4}{\pi}$. En effet, essayons de traduire cela mathématiquement. Si R désigne le rayon de ce cercle, alors on a

$$4a = 2\pi R$$

soit

$$R = \frac{2a}{\pi}$$

La valeur de la pente est alors égale à :

$$\frac{R}{a/2} = \frac{2a/\pi}{a/2} = \frac{4}{\pi}$$

Comme on l'a vu $\frac{4}{\pi}$ est proche de $\sqrt{G}$ donc on obtient quasiment la même pente (et donc la même hauteur) par cette méthode qu'avec celle basée sur le triangle d'or.

4.5. **Autres remarques.** Si une pyramide à base carré est construite suivant un triangle d'or, alors on en déduit de nombreuses égalités faisant intervenir π ou G. Cela n'a rien de surprenant et découle du choix du triangle d'or et/ou des relations entre π et G vues plus haut. On se contentera d'en citer deux.

a) L'aire formée par les quatre faces de la pyramide divisée par l'aire de la base est égale à G. En effet, comme on le constate facilement si on prend comme unité la demi-base, alors la hauteur du triangle formée par une face est égale à G (voir figure 3) ; donc l'aire de la face est

$$\frac{2 \times G}{2} = G$$

(rappelons que l'aire d'un triangle est égale à la base multipliée par la hauteur et divisé par 2). Donc pour les quatre faces, on obtient une aire totale égale à $4G$. Par ailleurs l'aire de la base est égale à $2 \times 2 = 4$, d'où l'on obtient

$$\frac{4 \times G}{4} = G$$

b) Le demi-périmètre de la base divisé par la hauteur de la pyramide vaut (approximativement) π. En effet, la hauteur est égale à $\sqrt{G}$ (voir figure 3), le demi-périmètre vaut 4, d'où l'on obtient

$$\frac{4}{\sqrt{G}}$$

Mais comme $\sqrt{G} \simeq \frac{4}{\pi}$ on a

$$\frac{4}{\sqrt{G}} \simeq \frac{4}{\frac{4}{\pi}} = \pi$$

Pour la pyramide de Khéops, les relations décrites dans a) et b) sont vérifiées approximativement : tout cela découle du choix de la pente de $14/11$ (qui est proche de $\sqrt{G}$) et qui fait que cette pyramide est approximativement construite suivant un triangle d'or.

5. Conclusion

Dans ce qui précède, nous avons vu les approximations :

$$\pi \simeq \frac{6}{5}G^2 \qquad \sqrt{G} \simeq \frac{4}{\pi} \tag{2}$$

Ces approximations sont des particularités arithmétiques entre π et le nombre d'or (et l'on a vu comment on pouvait les obtenir naturellement avec le développement en fraction continue) et elles n'ont rien à voir avec la pyramide de Khéops.

En revanche on retrouve G dans ce monument : la valeur de la coudée, ses dimensions, notamment sa pente égale à $14/11$ qui est une approximation de $\sqrt{G}$ (la pente du triangle d'or). Ceci n'est pas vraiment étonnant car on retrouve le nombre d'or dans de nombreuses constructions de l'Antiquité (ainsi que dans les cathédrales) et ces constructions nous paraissent harmonieuses.

Une question récurrente est de savoir si les Egyptiens de l'Ancien Empire connaissaient le nombre d'or. Pour ma part j'en suis convaincu. On pourrait penser que la présence du nombre d'or dans la pyramide de Khéops n'est qu'une coïncidence : les Egyptiens auraient choisi la pente $14/11$ parce ce qu'elle conduisait à un monument harmonieux et qu'il s'agissait d'une fraction simple. Or la fraction $14/11$ n'est pas seulement une approximation de $\sqrt{G}$ mais elle apparaît dans le développement de $\sqrt{G}$ en fraction continue (elle fait partie en quelque sorte de l'ADN de $\sqrt{G}$). Par ailleurs, les Egyptiens de cette époque ne connaissaient peut-être pas le nombre π mais on le retrouve souvent dans la pyramide de Khéops du fait de la présence de G et des approximations (2) citées plus haut. On peut résumer cela simplement en disant

« Quand il y a G, il y a π et réciproquement »

D'autre part, il existe peu de traces sur les connaissances mathématiques des Egyptiens de cette époque et cela nous amène à rester prudent. Cependant quelques problèmes issus du papyrus Rhind peuvent laisser perplexe. Citons en particulier le problème 56 (voir par exemple [Goy90] p. 96) où l'on demande le résultat de la division de 180 par 250. Le résultat est donné (sans démonstration) sous la forme

$$\frac{1}{2} + \frac{1}{5} + \frac{1}{50}$$

C'est ce que l'on appelle maintenant la décomposition égyptienne de la fraction $\frac{180}{250}$:

$$\frac{180}{250} = \frac{1}{2} + \frac{1}{5} + \frac{1}{50}$$

C'est exactement ce que l'on obtient en utilisant l'algorithme classique pour obtenir ce type de décomposition. Il semblerait donc que les Egyptiens connaissaient cet algorithme bien qu'ils n'aient aucun symbole (selon la thèse officielle) pour représenter les fractions générales de la forme p/q, mais seulement les fractions du type $1/q$.

Annexe A : introduction aux fractions continues

Cette annexe donne une introduction (sans démonstration) à la théorie des fractions continues pour permettre au lecteur non familier avec cette théorie de lire ce qui précède. Une fraction continue est une expression de la forme

$$a + \frac{1}{b}$$

ou bien de la forme

$$a + \cfrac{1}{b + \cfrac{1}{c}}$$

ou bien de la forme

$$a + \cfrac{1}{b + \cfrac{1}{c + \cfrac{1}{d}}}$$

et ainsi de suite... Ici a, b, c, d sont tous des nombres entiers ≥ 1 sauf éventuellement le premier a pour lequel on exige $a \in \mathbb{Z}$. Comme ces expressions sont trop encombrantes, on emploie usuellement la notation suivante

$$[a, b] = a + \frac{1}{b}$$

$$[a,b,c] = a + \cfrac{1}{b + \cfrac{1}{c}}$$

$$[a,b,c,d] = a + \cfrac{1}{b + \cfrac{1}{c + \cfrac{1}{d}}}$$

et ainsi de suite... Dans une fraction continue, comme par exemple $[a,b,c,d]$, les nombres a,b,c,d sont appelés les *quotients partiels* pour une raison que l'on expliquera par la suite.

A1 : développement des nombres rationnels. Tout d'abord toute fraction de nombres entiers (i.e. un nombre rationnel) peut s'écrire sous la forme d'une fraction continue, c'est-à-dire que si $\frac{p}{q}$ est une fraction en nombre entiers, alors il existe $m \geq 0$, $a_0 \in \mathbb{Z}$ et $a_1, \ldots, a_m \geq 1$ entiers tels que

$$\frac{p}{q} = [a_0, a_1, \ldots, a_m]$$

Par exemple si l'on considère la fraction $\frac{10}{17}$, alors on a

$$\frac{10}{17} = [0,1,1,2,3]$$

Pour trouver le développement en fraction continue d'une fraction on emploie l'algorithme d'Euclide qui permet de trouver les quotients partiels par une succession de divisions en nombres entiers (encore appelées divisions euclidiennes). Exposons-la sur un exemple (le cas général est identique).

Soit à décomposer par exemple $\frac{39}{14}$ en fraction continue. On effectue tout d'abord la division euclidienne de 39 par 14. Si on pose cette division le quotient est 2 et le reste 11 donc

$$39 = 14 \times 2 + 11$$

Ensuite on divise le diviseur de cette division (c'est-à-dire 14) par le reste (c'est-à-dire 11). On obtient un quotient de 1 et le reste est 3 donc

$$14 = 11 \times 1 + 3$$

Ensuite on recommence avec en faisant à chaque fois la division du diviseur par le reste de la division précédente et on continue jusqu'à obtenir un reste égal à 0. Ici cela continue par la division de 11 par 3, on obtient $11 = 3 \times 3 + 2$, puis la division de 3 par 2 donc $3 = 2 \times 1 + 1$ et enfin on divise 2 par 1 ce qui donne un quotient égal à 2 et un reste nul i.e. $2 = 1 \times 2 + 0$. En résumé, on a

$$\begin{aligned} 39 &= 14 \times 2 + 11 \\ 14 &= 11 \times 1 + 3 \\ 11 &= 3 \times 3 + 2 \\ 3 &= 2 \times 1 + 1 \\ 2 &= 1 \times 2 + 0 \end{aligned}$$

On en déduit comme on va le voir qu'alors

$$\frac{39}{14} = [2, 1, 3, 1, 2]$$

i.e. les quotient partiels sont les quotients des divisions successives, ce qui justifie la dénomination de quotients partiels. En effet, on déduit des égalités plus haut (en divisant à chaque fois par le diviseur) que :

$$\begin{aligned}
\frac{39}{14} &= 2 + \frac{11}{14} \\
\frac{14}{11} &= 1 + \frac{3}{11} \\
\frac{11}{3} &= 3 + \frac{2}{3} \\
\frac{3}{2} &= 1 + \frac{1}{2}
\end{aligned}$$

Alors de proche en proche :

$$\begin{aligned}
\frac{39}{14} &= 2 + \frac{11}{14} \\
&= 2 + \cfrac{1}{1 + \cfrac{3}{11}} \\
&= 2 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{2}{3}}} \\
&= 2 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{1}{1 + \cfrac{1}{2}}}}
\end{aligned}$$

c'est-à-dire

$$\frac{39}{14} = [2, 1, 3, 1, 2]$$

A2 : développement des nombres irrationnels. Si x est un nombre irrationnel, c'est-à-dire un nombre qui ne peut pas s'écrire sous la forme d'une fraction d'entiers (il est connu par exemple que π, e (la base des logarithmes népériens) ou encore le nombre d'or G sont des nombres irrationnels), alors on peut le représenter de manière *unique* sous la forme d'une fraction continue *infinie*, c'est-à-dire qu'il existe une unique suite de nombres $a_0, a_1, a_2, \ldots$ (avec toujours $a_0 \in \mathbb{Z}$ et $a_k \geq 1$ entier pour tout $k \geq 1$) telle que

$$x = [a_0, a_1, a_2, \ldots]$$

où $[a_0, a_1, a_2, \ldots]$ a ici la signification suivante :

$$[a_0, a_1, a_2, \ldots] = \lim_{n\to\infty} [a_0, a_1, a_2, \ldots, a_n]$$

Dans une telle représentation a_0 est forcément la partie entière de x.

Le développement le plus simple est celui du nombre d'or qui est donné par

$$G = [1, 1, 1, \dots]$$

c'est-à-dire que le développement en fraction continue de G est constitué uniquement de 1. On peut le trouver par exemple à l'aide du raisonnement suivant : comme $G^2 = G+1$, on a après division par G :

$$G = 1 + \frac{1}{G}$$

Or écrivons le développement en fraction continue de G i.e. $G = [a_0, a_1, a_2, \dots]$. On a alors en remplaçant

$$[a_0, a_1, a_2, \dots] = 1 + \frac{1}{[a_0, a_1, a_2, \dots]} = [1, a_0, a_1, a_2, \dots]$$

Par unicité du développement en fraction continue, on en déduit

$$a_0 = 1, \quad a_1 = a_0, \quad a_2 = a_1, \quad \dots$$

donc $a_0 = 1$, $a_1 = 1$, $a_2 = 1, \dots$ c'est-à-dire

$$G = [1, 1, 1, \dots]$$

Un autre développement connu (mais plus difficile à démontrer) est celui du nombre e (la base des logarithmes népériens) et qui est donné explicitement par

$$e = [2, 1, 2, 1, 1, 4, 1, 1, 6, 1, \dots]$$

faisant apparaitre les blocs $(1, 2, 1)$, $(1, 4, 1)$, $(1, 6, 1)$ etc... à partir du rang 1. En revanche le développement en fraction continue de π n'est pas connu dans son ensemble (au sens où l'on ne connait pas actuellement de "formule" qui permettrait de décrire l'ensemble des quotients partiels comme c'est le cas pour e ou bien G par exemple). En règle générale on se contente de calculer un certain nombre de quotients partiels et pour cela on passe par le développement décimal de x. Plus précisément à partir d'un certain nombre de décimales de x on va pouvoir obtenir un certain nombre de quotient partiels. La méthode sera expliquée à l'annexe A6.

A3 : grands quotients partiels et approximation. Soit x un nombre irrationnel et soit

$$x = [a_0, a_1, a_2, \dots]$$

son développement en fraction continue. Dans cet article nous avons plusieurs fois utilisé l'argument suivant : si a_n est un grand quotient partiel alors

$$x \simeq [a_0, a_1, \dots, a_{n-1}]$$

i.e. $[a_0, a_1, \dots, a_{n-1}]$ est une fraction qui est une bonne approximation de x. Essayons maintenant de justifier cet argument. Pour simplifier, on va supposer par exemple que a_3 est un grand quotient partiel. On peut écrire

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + A}} \tag{3}$$

avec

$$A = \cfrac{1}{a_3 + \cfrac{1}{a_4 + \ddots}} = \frac{1}{a_3 + B}$$

avec

$$B = \frac{1}{a_4 + \ddots} = [0, a_4, a_5, \ldots]$$

Comme $0 < B < 1$, on a donc

$$\frac{1}{a_3 + 1} < A < \frac{1}{a_3}$$

Si a_3 est grand, alors $\frac{1}{a_3} \simeq 0$ et $\frac{1}{a_3+1} \simeq 0$ donc $A \simeq 0$. D'après l'égalité (3), on aura alors

$$x \simeq a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2}} = [a_0, a_1, a_2]$$

A4 : autres résultats sur l'approximation par des fractions continues. Si x est un nombre irrationnel, écrivons son développement en fraction continue

$$x = [a_0, a_1, a_2, \ldots]$$

Pour tout $n \geq 0$, considérons alors la fraction continue

$$[a_0, a_1, a_2, \ldots, a_n]$$

et écrivons–là sous la forme d'une fraction irréductible $\frac{p_n}{q_n}$ avec $q_n > 0$:

$$\frac{p_n}{q_n} = [a_0, a_1, a_2, \ldots, a_n]$$

D'après la définition du développement en fraction continue de x, on a

$$\lim_{n \to \infty} \frac{p_n}{q_n} = x$$

donc les fractions $\frac{p_n}{q_n}$ s'approchent de plus en plus de x lorsque n devient grand. On peut même donner un estimation de l'erreur. En effet on démontre que

$$\left| x - \frac{p_n}{q_n} \right| < \frac{1}{q_n^2}$$

On démontre aussi que les p_n et q_n peuvent se calculer de proche en proche à partir des quotients partiels par les formules suivantes : on pose $p_{-1} = 1$, $p_0 = a_0$ et $q_{-1} = 0$ $q_0 = 1$. Alors pour tout $n \geq 1$, on a

$$p_n = a_n \, p_{n-1} + p_{n-2} \quad \text{et} \quad q_n = a_n \, q_{n-1} + q_{n-2}$$

Il est plus facile de calculer les fractions $\frac{p_n}{q_n}$ de cette manière que de les déduire de

$$\frac{p_n}{q_n} = [a_0, a_1, a_2, \ldots, a_n]$$

en calculant à chaque fois la fraction continue $[a_0, a_1, a_2, \ldots, a_n]$. A titre d'illustration prenons le cas de π. On a vu dans le texte que $\pi = [3, 7, 15, 1, 292, \ldots]$. On en déduit donc

$$\begin{aligned} p_{-1} &= 1 \\ p_0 &= 3 \\ p_1 &= 7 \times 3 + 1 = 22 \\ p_2 &= 15 \times 22 + 3 = 333 \\ p_3 &= 1 \times 333 + 22 = 355 \\ p_4 &= 292 \times 355 + 333 = 103993 \end{aligned}$$

puis avec $q_{-1} = 0$, $q_0 = 1$

$$\begin{aligned} q_{-1} &= 0 \\ q_0 &= 1 \\ q_1 &= 7 \times 1 + 0 = 7, \\ q_2 &= 15 \times 7 + 1 = 106 \\ q_3 &= 1 \times 106 + 7 = 113 \\ q_4 &= 292 \times 113 + 106 = 33102 \end{aligned}$$

Ainsi

$$\frac{p_1}{q_1} = \frac{22}{7}, \qquad \frac{p_2}{q_2} = \frac{333}{106}, \qquad \frac{p_3}{q_3} = \frac{355}{113}, \qquad \frac{p_4}{q_4} = \frac{103993}{33102}$$

A5 : qualité de l'approximation par des fractions continues. On démontre que les fractions $\frac{p_n}{q_n}$ déduites du développement en fraction continue d'un nombre irrationnel x sont les meilleures fractions qui approchent x au sens suivant : si $\frac{p}{q}$ (avec $q > 0$) est une fraction telle que $q \leq q_n$ et $\frac{p_n}{q_n} \neq \frac{p}{q}$ alors

$$\left| x - \frac{p_n}{q_n} \right| < \left| x - \frac{p}{q} \right|$$

c'est-à-dire que $\frac{p_n}{q_n}$ est plus proche de x que $\frac{p}{q}$. Ainsi par exemple pour π, l'approximation d'Adrien Métius est la meilleure approximation de π parmi toutes les fractions $\frac{p}{q}$ avec $0 < q \leq 355$.

A6 : développer en fraction continue un nombre irrationnel. Cette section expose une méthode pour trouver une partie du développement en fraction continue d'un nombre irrationnel x si l'on connait une partie de son développement décimal. Dans la section sur le développement en fraction continue d'un nombre rationnel nous avons vu que l'on pouvait obtenir son développement en faisant uniquement des divisions euclidiennes en nombres entiers. Ceci peut se programmer facilement sur une machine. Afin d'illustrer la méthode pour un nombre irrationnel supposons par exemple que l'on connaisse jusqu'au trois premières décimales de x. Par exemple supposons que

$$x = 13.203\cdots$$

On considère alors les approximations décimales par défaut 13.203 et par excès 13.204 et l'on écrit ces approximations sous la forme d'une fraction, c'est-à-dire

$$\frac{13203}{1000}, \qquad \frac{13204}{1000}$$

Ensuite on développe ces deux fractions en fraction continue comme on l'a vu dans la section A1. Après calcul, on trouve

$$13.203 = [13, 4, 1, 12, 1, 1, 7], \qquad 13.204 = [13, 4, 1, 9, 5]$$

La plus grande plage de quotients partiels (à partir du début) commune aux deux développements est alors le début du développement en fraction continue de x, c'est-à-dire ici

$$x = [13, 4, 1, \ldots]$$

Prenons un autre exemple, celui de $\pi = 3.14159265\cdots$. On a suivant le même démarche

$$3.141 = [3, 7, 10, 1, 5, 2], \qquad 3.142 = [3, 7, 23, 1, 2]$$

Ainsi $\pi = [3, 7, \ldots]$. Donc on n'obtient qu'un seul quotient partiel en utilisant les trois premières décimales de π. Même avec six décimales on n'obtient toujours qu'un seul quotient partiel car

$$3.141592 = [3, 7, 15, 1, 84, 6, 2], \qquad 3.141593 = [3, 7, 16, 983, 4, 2]$$

En revanche avec 7 décimales, on a

$$\begin{aligned} 3.1415926 &= [3, 7, 15, 1, 243, 1, 1, 9, 1, 1, 4] \\ 3.1415927 &= [3, 7, 15, 1, 354, 2, 6, 1, 4, 1, 2] \end{aligned}$$

donc cette fois on obtient trois quotients partiels et donc

$$\pi = [3, 7, 15, 1, \ldots]$$

Pour récupérer le quatrième quotient partiel (c'est-à-dire 292) il faudrait utiliser les dix premières décimales de π.

A7 : quelques compléments. Donnons quelques compléments récents. Si l'on connait les n premières décimales d'un nombre irrationnel x, la méthode précédente permet donc de déterminer avec certitude un certain nombre de quotients partiels. Notons $k_n(x)$ ce nombre de quotients partiels. Il y a eu récemment plusieurs travaux autour de cette fonction k_n. Premièrement on connait son comportement asymptotique lorsque $n \to \infty$ et cela pour presque tous les nombres irrationnels x (le terme de "presque tout" a un sens précis en mathématique mais on ne cherchera pas à le définir ici car cet exposé se veut élémentaire). C'est un résultat dû à Lochs [Loc64]. Pour presque tout nombre irrationnel x, on a

$$\lim_{n\to\infty} \frac{k_n(x)}{n} = \frac{6 \log 2 \log 10}{\pi^2} \simeq 0.9702$$

Comme la constante $6 \log 2 \log 10/\pi^2$ est proche de 1, on peut presque dire que lorsque n est grand n décimales déterminent n quotients partiels.

Citons aussi un autre résultat plus récent sur k_n qui permet d'avoir un résultat plus concret [Fai98]. Lorsque $n \to \infty$ la distribution des valeurs de

$$\frac{k_n - na}{\sigma\sqrt{n}}$$

se rapproche de la loi normale $N(0, 1)$. Ici a désigne la constante

$$a = \frac{6 \log 2 \log 10}{\pi^2}$$

et $\sigma > 0$ est une autre constante avec $\sigma \simeq 0.7708$ (les quatre décimales sont exactes). En fait la distribution de k_n est très proche (même pour les petites valeurs de n, en fait pour $n \geq 4$) d'une loi normale

$$N(na - 1, \sigma\sqrt{n})$$

où le -1 est une correction de continuité (voir [Fai07] p. 85–91). Ceci permet de donner concrètement une borne précise. Par exemple, pour un niveau de confiance de 95%, si l'on veut k quotients partiels (au moins), il faut prendre un nombre de décimales au moins égal à

$$\frac{\left(-q\sigma + \sqrt{q^2\sigma^2 + 4ak}\right)^2}{4a^2}$$

avec $q = 1.96$ et où les constantes a et σ ont été définies plus haut.

RÉFÉRENCES

[Dor04] Gilles Dormion. *La chambre de Chéops, analyse architecturale*. Fayard, 2004.

[Fai98] C. Faivre. A central limit theorem related to decimal and continued fraction expansions. *Arch. Math. (Basel)*, 70 :455–463, 1998.

[Fai07] C. Faivre. Habilitation à diriger des Recherches. Contributions à l'étude métrique des fractions continues, 2007.

[Goy90] Georges Goyon. *Le secret des batisseurs des grandes pyramides, Khéops*. Pygmalion, 1990.

[Loc64] G. Lochs. Vergleich der Genauigkeit von Dezimalbruch und Kettenbruch. *Abh. Math. Sem. Univ. Hamburg*, 27 :142–144, 1964.

(2016) B.I.A.A. no **104** BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
ISSN 0291-5413 (Université d'Aix-Marseille)

Science et /ou Fiction. Transcription de la conférence enregistrée à Marseille le 5 juin 1974.

Jean-Marie SOURIAU [1]

Résumé. – (N.D.L.R.) : « *Science et/ou Science-Fiction* » fait partie du cahier numéro 9 du séminaire H.S.I.F.S. : « *Séminaire d'Histoire et Sociologie des Idées et des Faits Scientifiques* » de l'Université de Provence et de sa Faculté des Sciences à Marseille. Le texte original des débats de 1974 a été dactylographié à partir d'une bande sonore : il comporte des *coquilles* ainsi qu'un nombre significatif de phrases approximativement retranscrites. Pour ce numéro du *Bulletin*, nous nous sommes limités à une réédition de la partie introductive exposée par Jean-Marie Souriau en préambule au débat. Nous avons essayé de rendre le texte le plus lisible possible, tout en préservant son intégrité ainsi que son caractère original, parfois haut en couleur et typique de sa forme orale originelle. (L'original du texte intégral de la conférence – introduction et débat – est mis à disposition de toute personne intéressée sur simple demande.) Suivant le conseil de l'auteur, nous avons aussi complété et vérifié un certain nombre de références ; pour le confort du lecteur, quelques notes historiques accompagnent ces précisions. Notons aussi que le contenu de ce séminaire de Souriau à été retravaillée par l'auteur lui-même (probablement à partir de ses notes préparatoires), ce qui a donné une version très aboutie dans sa forme littéraire. Ce dernier texte – que vous pouvez lire dans le n° 105 du *Bulletin* – à servi de base à une publication dans la revue « *La Recherche* » d'octobre 1974.

Pour terminer ce préambule, rappelons que le *"pourquoi de ces cahiers ?"* du séminaire H.S.I.F.S. a été développé dans le numéro zéro par l'élucubration collective de son équipe historique, dont notre ami Georges Chappaz en est la mémoire. Ainsi, nous lisons dans ce numéro que « *Les diverses conférences ont régulièrement attiré un public nombreux pour un petit centre comme la Faculté Saint Charles à Marseille (60 à 100 auditeurs) [...]. Volontairement, nous avons conservé l'intégralité de l'ensemble des exposés, et même autant que faire se peut (dans la limite de l'intelligibilité) des "débats" qui ont suivi. Pour nous, c'est l'ensemble, en tant que tel, qui présente un intérêt et qui constitue un témoignage de ce que peuvent être dans un petit centre universitaire les préoccupations des scientifiques, leurs problèmes, leurs façons spécifiques de les aborder* ».

Je dois d'abord quelques excuses préliminaires à l'auditoire [...] : finalement pour des raisons tout à fait accessoires, je n'ai pas eu le temps de préparer assez cette conférence et il y a beaucoup de citations [données] de mémoire et sans doute avec des erreurs. S'il y a des gens qui peuvent me corriger, ça sera utile. Je voudrais dire quelque chose de plus grave [...]. Si j'ai accepté de parler [de science-fiction], c'est tout à fait par étourderie. Je ne me rendais pas compte [...] que ce n'est pas du tout une conférence qu'il fallait faire [...], mais plutôt un cours, ou y consacrer une vie, car le sujet est tellement immense [...].

[Je connais] au moins deux cours [sur le thème de la science fiction] : un qui s'est passé aux Etats-Unis et un qui s'est passé en France. En France c'était dans une faculté de Lettres à Paris ; aux Etats-Unis c'était dans une école militaire, ce qui prouve la différence de points de vue sur ce sujet entre les Européens et les Américains. Je reviendrai sur le cours américain dans un instant.

Donc aujourd'hui je ne peux parler que d'un certain nombre de points [...] qui m'ont paru avoir un certain intérêt, mais bien entendu ça n'a aucune prétention. Il y a beaucoup d'autres choses qu'il faudrait dire ; si en particulier il y a des gens ici qui n'ont jamais lu de science-fiction, il ne faut pas qu'ils croient qu'après cette conférence ils sauront de quoi il s'agit. Bien entendu, il faut prendre en science-fiction,

1. Professeur de l'Université de Provence (1923-2012).

les grands thèmes et c'est là-dessus que je vais organiser cette causerie.

[Parmi] les thèmes de science-fiction, il y en a un qui est tout à fait rebattu, mais je pense qu'il est impossible de le passer sous silence : c'est le thème du voyage dans l'espace [...].

FIGURE 1 *: (Source Wikipédia) Frontispice de « l'Histoire comique » contenant « Les États et empires de la Lune ». de Cyrano de Bergerac. Edité à Amsterdam, en 1709. Le narrateur s'élève dans les cieux grâce à des fioles de rosée.*

Le premier roman de science-fiction auquel je vais faire allusion est paru en 1657. Je vous donne le titre exact : « *Histoire comique des États et Empires de la Lune* » ; bon, je ne vous cacherai pas le nom de son auteur : c'était Cyrano de Bergerac. C'est une œuvre posthume (Cyrano est mort en 1655) [...]. Il y a une allusion aux moyens que Cyrano utilise pour se rendre dans la Lune, des moyens qui sont assez fantaisistes [comme] l'évaporation, la rosée du matin etc. On peut dire voilà de la fantaisie, voilà un auteur qui a choisi ce sujet en l'air et qui a parlé de ça comme il aurait parlé d'autre chose. Cyrano a également écrit [entre autres] une tragédie qui s'appelait « *La Mort d'Agrippine* » [ainsi qu']une comédie qui s'appelait « *Le Pédant Joué* » (par la suite un peu pillée par Molière) : donc il a choisi ce sujet parmi d'autres sujets. Je voudrais faire une étude de l'ambiance dans laquelle ce livre a été écrit qui montre que, contrairement à cet aspect de fantaisie, c'est l'aboutissement d'une intense période d'activité intellectuelle et dans laquelle beaucoup de personnalités se sont connues mutuellement. Je commence par citer une date : 1600, Giordano Bruno a été brûlé pour différentes raisons – bien sûr pour ses opinions. Quel est le rapport entre Giordano Bruno et Cyrano de Bergerac ? Cyrano de Bergerac a cité effectivement Giordano Bruno parmi les gens qui l'ont inspiré : comment ça a pu se produire ? Eh bien, l'une des visions les plus intéressantes de Giordano Bruno [...], c'est la vision anti-aristotélicienne de l'espace infini. Alors que pour Aristote – comme chacun sait – il y avait une sphère, des étoiles et puis derrière un peu de lumière, des anges etc., Bruno semble être l'un des premiers – enfin il y avait Démocrite avant lui – qui ait essayé de rendre réaliste [la vision anti-aristotélicienne] : cette notion de l'infini, de l'espace.

Je cite un autre fait qui semble apparemment sans lien. En 1609, Kepler publie ses deux premières lois de la mécanique céleste. 1610 : Galilée publie le « *Siderius Nuncius* » dans lequel il commence à révéler les découvertes [astronomiques] qu'il a faites [...] grâce à la lunette perfectionnée qui porte son nom (mais qu'il n'avait pas inventée) ; il découvre les tâches du soleil, les montagnes de la Lune : il suffit d'avoir une lunette pour voir ces choses-là. Je dois citer une autre personne, qui était professeur de Philosophie à l'Université d'Aix, qui s'appelait Monsieur Gassend qui a arrangé son nom en Gassendi ; c'était un véritable provençal – enfin il est né à 10 km de Digne [...]. Gassendi a joué un rôle important dans le mouvement des idées à cette époque ; en particulier il a correspondu avec Galilée et il a fait des observations complémentaires à celles de Galilée, quand il habitait Aix : une de ses observations intéressantes était une éclipse de lune.

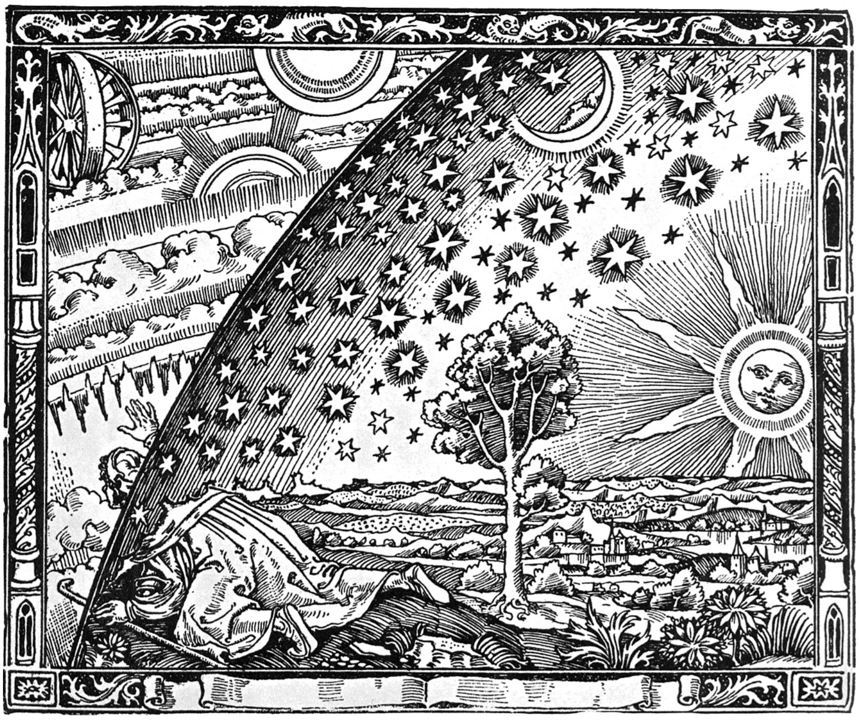


FIGURE 2 *: (D'après Wikipédia-fr) : La « Gravure au pélerin » est une gravure sur bois anonyme, qui réapparait dans le livre de Camille Flammarion publié en 1888, « L'Atmosphère : météorologie populaire », au chapitre « La forme du ciel ».*

Il faut que je revienne à Kepler ; il a publié sa troisième loi en 1618. Ses trois lois permettaient de faire de la mécanique céleste ; il a fait des prédictions, construit des éphémérides et a prédit un passage de Mercure devant le Soleil, phénomène assez rare qui ne se produit qu'une ou deux fois par siècle (la prochaine fois ce sera le 10 novembre prochain). Gassendi a observé [ce passage à partir d'un observatoire aixois]. Il a fait mieux en 1640. Galilée avait énoncé le principe [d'une expérience où] il imaginait un bateau en mouvement de translation rectiligne uniforme sous une brise favorable ; du haut du mât on laisse tomber un projectile [...] : le projectile va décrire une trajectoire parabolique, mais en raison du mouvement du bateau on verra le poids tomber au pied du mât. C'est un exemple qui avait été imaginé et publié par Galilée en 1639 [pour illustrer] la relativité du mouvement. En 1640, Gassendi a fait l'expérience devant Marseille : il a pris un bateau et il a réalisé [l'expérience] : c'était un esprit réaliste. Une autre œuvre de Gassendi [a été de dresser] l'une des premières cartes de la Lune. Au bout d'un certain temps la philosophie ne lui a pas suffit – [Gassendi] est aussi un peu Monseigneur sur les bords : il s'est fait nommer Professeur de Mathématiques au Collège de France en 1645 et ceci pour [la] raison

2. N.D.L.R. : (D'après Wikipédia (fr) dans l'article sur Cyrano de Bergerac) : Dans une page célèbre de sa *« Vie de M. de Molière »*, parue en 1705, Jean-Léonor Le Gallois, sieur de Grimarest, tentera, dans des termes peu flatteurs pour Cyrano, de justifier les emprunts que Molière a fait à son œuvre. Après avoir indiqué que le père du futur comédien s'était résolu à l'envoyer au collège des Jésuites, il écrit :

« Le jeune Pocquelin était né avec de si heureuses dispositions pour les études qu'en cinq années de temps, il fit non seulement ses Humanités, mais encore sa Philosophie. Ce fut au collège qu'il fit connaissance avec deux hommes illustres de notre temps : Mr de Chapelle et Mr Bernier. Chapelle était fils de Mr Luillier, [. . .] [lequel] n'épargna rien pour [lui] donner une belle éducation, jusqu'à lui choisir pour précepteur le célèbre Mr de Gassendi, qui, ayant remarqué dans Molière toute la docilité et toute la pénétration nécessaires pour prendre les connaissances de la philosophie, se fit un plaisir de la lui enseigner en même temps qu'à Messieurs de Chapelle et Bernier. Cyrano de Bergerac, que son père avait envoyé à Paris sur sa propre conduite, pour achever ses études, qu'il avait assez mal commencées en Gascogne, se

[...] qu'il fréquentait un cercle littéraire et philosophique à Paris ; dans ce cercle figuraient des gens comme Molière justement et comme Cyrano de Bergerac [2], si bien que lorsque Cyrano a rédigé dans cette ambiance parisienne son roman de fiction, c'était réellement de la science-fiction : [...] il était imbibé d'un milieu philosophique et scientifique en un certain sens. Je vous signale qu'il a bien été militaire, mais il a abandonné la vie militaire à l'âge de 22 ans. Il est mort [...] assassiné, peut-être : enfin, il est mort [mystérieusement] à 36 ans. En tout cas, il fréquentait ce milieu littéraire [dans lequel] il a créé une œuvre fantaisiste ; donc la fantaisie ici n'est qu'apparente : il y avait toute une ambiance intellectuelle et scientifique qui a donné naissance à cette œuvre dite de science-fiction.

Je vais prendre un deuxième exemple [...] qui est très connu : c'est [celui de Jonathan] Swift qui s'est passé 100 ans plus tard. Swift a décrit dans « *Les voyages de Gulliver* » (1726) un certain nombre de fantaisies et parmi ces fantaisies quelque chose qui s'appelait *Laputa*. Je ne sais pas s'il faut y voir une intention malveillante [en rapport à] l'université [3], [mais] dans cette université de *Laputa*, raconte Swift, les nombreux astronomes – [d'influents personnages portés sur la politique] – avaient découvert à la planète Mars, deux satellites qui vérifiaient les lois de Kepler. Ce n'était pas du tout une coïncidence merveilleuse des horaires et des durées de révolution qu'avait donné Swift : cela signifiait explicitement que c'était conforme aux lois de Kepler. Swift avait donc aussi, bien entendu, l'intention scientifique de se mettre en accord.

Alors évidemment, ceci a beaucoup fait parler, parce qu'effectivement en 1877 [4] on a découvert ces deux satellites très proches de Mars comme Swift l'avait prévu. J'ai lu (sans doute beaucoup d'entre vous ont lu aussi des choses un peu mystérieuses à ce sujet), que Swift était un grand initié. Swift savait : bon, je crois qu'on peut faire la part du hasard bien sûr ; mais il n'y a aucune chance qu'il ait pu avoir connaissance d'une découverte scientifique : ces satellites sont très petits et très difficiles à observer, c'est-à-dire que ce n'est qu'à partir du moment où on a eu un instrument assez puissant, qu'on a pu les voir. Il faut bien se rappeler que si on dessine le Soleil, la Terre a un satellite, voici Mars qui en a deux, Jupiter en a quatre (découverts par Galilée). Il est assez normal de voir une progression géométrique. . . Il se trouve que c'est comme ça.

Il y a eu une suite à l'histoire de ces satellites. En 1965 un astronome soviétique a déclaré qu'en étudiant le mouvement des deux satellites, il y en avait un, le plus proche, qui était freiné. Il était freiné par l'atmosphère martienne ; [mais] l'atmosphère martienne étant très ténue, elle ne pouvait pas expliquer un tel freinage, sauf si le satellite en question était une sphère creuse. De là à penser [qu'il s'agissait] d'un satellite artificiel construit un jour par les Martiens, il n'y avait qu'un pas qui a été rapidement franchi [...]. Depuis, on a envoyé [une sonde] qui a photographié de près les deux satellites de Mars : on voit que ce sont des honnêtes morceaux de montagne qui n'ont absolument pas l'apparence de satellites artificiels [...] : un camouflage bien fait.

---

glissa dans la société des disciples de Gassendi, ayant remarqué l'avantage considérable qu'il en tirerait. Il y fut admis cependant avec répugnance ; l'esprit turbulent de Cyrano ne convenait point avec de jeunes gens qui avaient déjà toute la justesse d'esprit que l'on peut souhaiter dans des personnes toutes formées. Mais le moyen de se débarrasser d'un jeune homme aussi insinuant, aussi vif, aussi gascon que Cyrano ? Il fut donc reçu aux études et aux conversations que Gassendi conduisait avec les personnes que je viens de nommer. Et comme ce même Cyrano était très avide de savoir et qu'il avait une mémoire fort heureuse, il profitait de tout et il se fit un fonds de bonnes choses dont il tira avantage dans la suite. Molière aussi ne s'est-il pas fait un scrupule de placer dans ses ouvrages plusieurs pensées que Cyrano avait employées auparavant dans les siens. "Il m'est permis, disait Molière, de reprendre mon bien où je le trouve. »

Ces lignes sont le seul « document » qui évoque la rencontre dès cette époque de Chapelle et Molière, dont l'amitié n'est avérée, par de nombreux témoignages, qu'à partir de 1659, l'assistance de Molière aux entretiens de Gassendi et la rencontre de Cyrano avec Molière.

3. N.D.L.R. : Extrait du roman de Swift à propos des astronomes de Laputa : « *Plusieurs d'entre eux, principalement ceux qui s'appliquent à l'astronomie, donnent dans l'astrologie judiciaire, quoiqu'ils n'osent l'avouer publiquement ; mais ce que je trouvai de plus surprenant, ce fut l'inclination qu'ils avaient pour la politique et leur curiosité pour les nouvelles ; ils parlaient incessamment d'affaires d'État, et portaient sans façon leur jugement sur tout ce qui se passait dans les cabinets des princes. J'ai souvent remarqué le même caractère dans nos mathématiciens d'Europe, sans avoir jamais pu trouver la moindre analogie entre les mathématiques et la politique ; à moins que l'on ne suppose que, comme le plus petit cercle a autant de degrés que le plus grand, celui qui sait raisonner sur un cercle tracé sur le papier peut également raisonner sur la sphère du monde ; mais n'est-ce pas plutôt le défaut naturel de tous les hommes, qui se plaisent ordinairement à parler et à raisonner sur ce qu'ils entendent le moins ?* »

4. N.D.L.R. : Deimos et Phobos sont découverts par Asaph Hall en 1877.

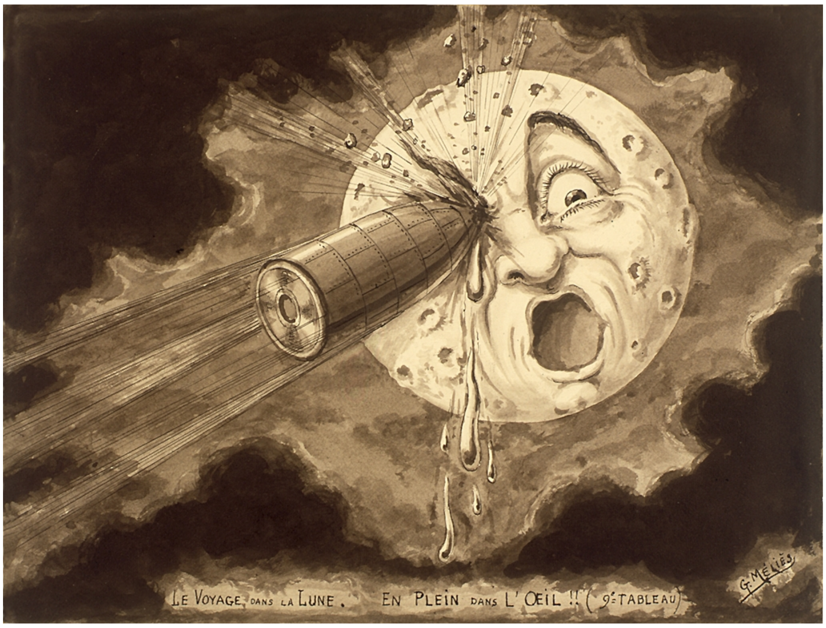


FIGURE 3 *: Mise en scène de « l'obus de Jule verne » par Méliès (1902).*

Bon, je vais encore vous parler d'un sujet bateau, je ne peux pas y échapper : bien entendu, c'est Jules Verne. Jules Verne, [...] c'est quelque chose d'extrêmement différent de Cyrano de Bergerac. Cyrano de Bergerac, ses moyens de s'envoler étaient fantaisistes. Il en citait beaucoup. Pourquoi ? Parce que c'était un atomiste, c'était un disciple de Gassendi et de Démocrite ; il pensait qu'à partir du moment où [...] la Lune était semblable à la Terre, que la Terre était la Lune de la Lune [...], on trouverait un moyen pour y arriver, que la connaissance, la raison, permettrait certainement d'y arriver, que c'était un détail subalterne. Au contraire, Jules Verne a écrit *« De la Terre à la Lune »* au moment d'une technologie triomphante. C'est assez intéressant d'ailleurs de lire ce roman, pour ceux qui ne l'ont pas fait, parce que justement Jules Verne insiste sur le côté réaliste, sur la technologie et même sur la sociologie. Une grande partie du livre consiste à décrire comment on fait une souscription internationale ; il y figure ce que chaque pays a donné avec des commentaires sur l'esprit scientifique et aventureux des différents pays. [...]

Jules Verne qui avait un conseiller scientifique, un polytechnicien, savait qu'il fallait lancer son satellite – enfin son obus – des zones équatoriales et qu'il n'y avait que deux états dans les Etats Unis d'où l'on pouvait faire le lancement : le Texas et la Floride. Il raconte la bagarre entre les délégués du Texas et de la Floride qui voulaient chacun posséder le centre de lancement. Je vous rappelle que le centre de lancement efficace est en Floride [et qu'aujourd'hui] les opérations sont dirigées du Texas, de Houston ; par conséquent là, il y avait une anticipation qui est assez étonnante. Dans l'obus de Jules Verne, il y avait trois astronautes [avec] évidemment quelque chose qui ne marche pas : [il] lançait son satellite avec un canon enterré dans le sol qu'il appelait « *Columbiad* ». [...] On a beaucoup ridiculisé l'idée de lancer un satellite – un obus – avec un canon jusqu'à la Lune et d'y mettre des astronautes. Mais je voudrais signaler une chose, c'est que le véhicule [du premier voyage qui] s'est posé sur la Lune [...] s'appelait *« Columbia »* et que ça n'est pas par hasard ; c'est Amstrong qui avait choisi de l'appeler *« Columbia »* par référence à Jules Verne, si bien qu'il y a un autre aspect réciproque de celui que je vous ai décrit précédemment : l'influence sur les scientifiques, ou ici plutôt des techniciens, de la fiction antérieure.

Les Soviétiques récupèrent leurs cosmonautes sur terre, les Américains récupèrent leur astronautes dans le Pacifique, exactement comme dans le second roman de Jules Verne « *Autour de la Lune* » qui raconte le retour. On ne peut sous-estimer ces références : la fiction a influencé le comportement des scientifiques et des techniciens même quand il s'agissait de choisir entre des programmes qui coûtaient beaucoup d'argent.

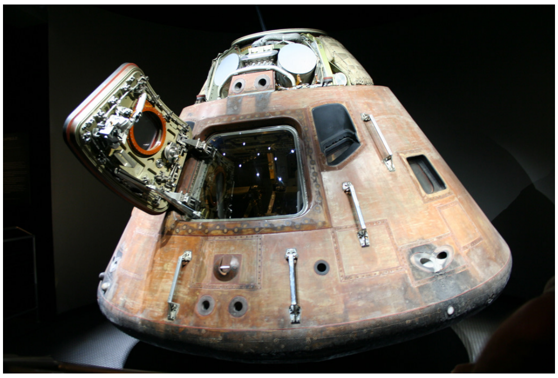


FIGURE 4 : *(Wikimedia Commons) : Capsule Columbia (Appolo 11).*

Un autre fait [bien] connu, est l'influence qu'a eu l'autre roman de Jules Verne « *Vingt milles lieues sous les mers* » paru en 1870. Comme chacun sait, son sous-marin s'appelle le « *Nautilus* » et depuis, il y a eu beaucoup de sous-marins qui se sont appelés « *Nautilus* ». En particulier, il y en a un qui a réédité presque exactement l'un des exploits prêtés par Jules Verne au Capitaine Némo, à savoir [...] percer la banquise et sortir au Pôle. [Cet exploit à été réalisé] par un sous-marin à énergie nucléaire, l'un des premiers [construits] par les Américains. Là [encore], la référence était aussi voulue : Jules Verne avait aussi bien impressionné la Marine Américaine que la Nasa. Ce n'est pas la seule influence de « *Vingt milles lieues sous les mers* » ; quand [le roman] paraît en 1870, Rimbaud a 16 ans et [en 1871] il publie un livre qui s'appelle « *Le bateau ivre* ». Certains pensent que c'est une transcription de certains passages de Jules Verne ; je ne garantis pas : c'est une dispute littéraire dans laquelle je ne suis pas compétent mais enfin, cette influence a donc été dans différents domaines.

Maintenant il faut bien arriver à des faits plus récents. La science-fiction proprement dite [...] remonte selon les bons auteurs à 1926, en Amérique bien sûr, avec la publication du magazine « *Amazing Stories* »[5]. Cette science-fiction américaine est visiblement le fruit d'autres circonstances sociologiques. [1926[6]], la mécanique quantique venait de prendre naissance ; c'était la période de triomphe de la physique et de la technologie : c'est dans cette ambiance que s'est créée la science-fiction américaine, dont on croit trop souvent qu'elle est uniquement technologique [...]. Je ne veux pas vous faire une liste d'auteurs de science-fiction : il y en a actuellement un grand nombre [...]. Il y a dix à vingt ans – ce qu'on appelle maintenant l'âge d'or de la science-fiction – un grand nombre [de ces] auteurs [...] étaient des scientifiques. Je l'ai dit tout à l'heure, [la science-fiction] était un objet d'enseignement universitaire [et] je vous donne une référence : le numéro de « *L'American Journal of Physics* » de février 1973 [dans lequel on trouve l'article « *The Science in Science fiction : A Seminar Course* » de G. P. Calame]. Donc c'est quelque chose de récent. Je vous lis rapidement le résumé : « *Un séminaire basé sur la science utilise des histoires de science-fiction : le cours consiste en lecture d'histoire, discussion de la physique, ce qui sert de fondation scientifique à l'histoire.* » La réaction des étudiants au cours est enthousiaste. Je vous signale quelques auteurs du programme :

– « *Neutron Star* » de Larry Niven (1970) ; [thèmes scientifiques :] forces de marées, étoiles à neutrons *red shift* gravitationnel ;

– « *Tau Zero* » de Poul Anderson (1970) ; thèmes scientifiques : effet Doppler, dilatation du temps, relativité restreinte, cosmologie de l'univers oscillant.

– « *Mission of Gravity* » de Hal Clement (1962) : c'est le récit d'une planète où la gravité est très [forte] ; mais [cette planète] tourne très vite, si bien qu'il y a d'énormes différences de gravitation entre l'équateur et le pôle.

5. N.D.L.R. : (source Wikipedia) : « *Amazing Stories* » est un magazine américain lancé en avril 1926 par Hugo Gernsback. C'était le premier magasine entièrement dédié à la *Science fiction*. Des histoires de science fiction paraissaient déjà régulièrement dans d'autres magasines, certaines publiées par Gernsback, mais « *Amazing Stories* » a réellement permis de définir et de lancer un nouveau genre de *pulp fiction*. Hugo Gernsback (188-1967) est bien un personnage clef de la sciences fiction du 20-ème siècle. C'est lui même un auteur de science-fiction, mais c'est surtout son dévouement pour le développement de ce nouveau genre littéraire qui lui vaut sa célébrité. Il a notamment permis à de nombreux auteurs de faire leurs premiers pas dans les magazines qu'il a créés. Il est considéré comme le créateur du terme *science-fiction* », influencé par Jules Verne et H.G. Wells ; le plus prestigieux prix de la science-fiction porte son (pré)nom : le « *Prix Hugo* ».

6. N.D.L.R. : Publication de la série des mémoires de Schrödinger sur les fondements de la mécanique ondulatoire.

[L'histoire raconte] un voyage de l'équateur au pôle avec, bien entendu, un dispositif antigravitationnel pas encore inventé bien sûr. Je dois signaler que ce livre a été très mal accueilli en France par la critique qui l'accusait d'être un livre purement technologique ; [il] n'a pas séduit les lecteurs européens [bien qu'il soit] considéré comme un classique aux Etats-Unis etc. . .

Ah oui, il y en a un autre qui est intéressant c'est « *Dune* » de Frank Herbert (1965), qui vient d'être publié en France [et dont un des thèmes] scientifiques [est] l'écologie. Il y a plus que de l'écologie dans le récit de « *Dune* » [et en particulier] beaucoup de politique sous-jacente, mais je ne veux pas entrer dans les détails.

Supposons que je veuille parler de la science-fiction américaine à des gens qui n'en ont jamais lu. Je ne pourrai faire qu'une chose aujourd'hui, c'est essayer de faire une classification des thèmes : c'est là-dessus que je vais terminer. Ces thèmes ont un caractère bien particulier ; ils sont interchangeables [et sont formulés dans] un vocabulaire [...] d'initiés ; [cet aspect disparait] au niveau des lecteurs qui en lisent beaucoup, [même si] les sujets et les allusions littéraires sont très fréquentes. [On peut dire que] ces thèmes sont un bien commun. Comme les thèmes musicaux au 18ème siècle, ils sont considérés comme quelque chose sur lequel chaque auteur a le droit d'écrire des variations sur un thème donné, [...] comme les thèmes de Gershwin dans le jazz. Je vais essayer d'en faire un petit tableau qui ne sera [...] pas du tout complet mais enfin, [j'espère en donner] une idée.

Il faut [bien sûr] commencer par le voyage spatial. Je ne reviens pas sur ce thème – on en connait l'importance aux yeux des Américains. [Je vais plutôt parler de] tous les thèmes qui en dérivent et le premier, c'est [évidemment] le contact avec la vie étrangère.

On voit une toute petite allusion à ce contact dans Jules Verne [...] lorsque les cosmonautes passent derrière la Lune ; certes, la face [cachée de la Lune] est obscure, ils n'y voient rien, mais à un moment ils allument des fusées [et] ils aperçoivent des forêts [...]. Jules Verne fait [ici] une allusion extrêmement rapide à la possibilité de vie sur la Lune, non pas sur la face que nous voyons, mais sur l'autre.

La vie étrangère, c'est un thème qui remonte à fort loin et les descriptions [possibles en sont] inépuisables. Je vais citer simplement quelques exemples. Un qui a été très célèbre, c'est les « *Chroniques martiennes* » de Ray Bradbury (1950). C'est un livre qui a produit un grand choc psychologique, un livre très poétique. L'arrivée sur Mars y est à peine décrite : quand les Terriens débarquent sur Mars, les Martiens [meurent] immédiatement de la varicelle et il ne reste plus que des ruines et des allusions poétiques. Les Martiens sont un peuple frêle et composé certainement de littéraires.

Un autre exemple de description de la vie étrangère, c'est un film qui n'a pas eu beaucoup de succès mais [...] quand même intéressant, qui s'appelle « *Danger, planète interdite* ». Il est passé à Marseille il y a un an ou deux. Alors là, c'est une idée très simple : voilà le Soleil, voilà la Terre, les astronomes ayant fait quelques calculs s'aperçoivent qu'il y a une planète qui est juste l'anti-Terre. L'anti-Terre est un thème très ancien : [c'est une planète] qui a la même période [que la Terre, ce qui est] tout à fait permis par la mécanique céleste. On y envoie des cosmonautes [qui] croient arriver sur la planète et ils se retrouvent sur Terre, [ou ce qui semble être la Terre]. L'un des cosmonautes rentre chez lui, il rencontre sa femme et en allant se laver les dents, il aperçoit sur la tablette de la salle de bain des produits d'après rasage. Mais l'étiquette [est] imprimée à l'envers : il s'aperçoit [alors que] sur cet anti-monde – cette anti-Terre complète – tout a été retourné,... [en particulier] sa femme ! Alors bien sûr, ça a une implication : [le cosmonaute réalise] que son [double] anti-cosmonaute, qui est parti [de l'anti-Terre], est arrivé [sur Terre et a rencontré] sa femme...

Un autre contact qui est intéressant, part d'une autre idée ; c'est un roman écrit par [Fred Hoyle] [7] – un célèbre astronome anglais – qui

7. N.D.L.R. : (D'après Wikipedia-en) : Fred Hoyle, astronome connu pour ses travaux sur la nucléosynthèse des éléments au coeur des étoiles, était partisan de la théorie de l'univers statique : c'est lui qui a inventé le terme « *Big Bang* » pour désigner ironiquement la théorie concurrente d'un univers en expansion. Il a aussi impressionné Richard Feynman en prédisant correctement un niveau d'énergie spécial de l'atome de carbone, en raisonnant uniquement sur le fait que sans ce niveau d'énergie, les étoiles ne pourraient pas synthétiser les éléments nécessaires au développement de la vie telle qu'elle existe sur terre.

FIGURE 5 : *(Source Wikipédia) : Couverture de l'édition originale de « The Black Cloud » de Fred Hoyle (1957).*

s'appelle « *Le Nuage Noir* »[8]. Comme le titre l'indique, c'est un nuage qui vient du Cosmos et arrive sur Terre ; [il] risque de détruire le système solaire ; il a l'air d'avoir des facultés d'absorption très grandes et on pense même qu'il est capable d'avaler le soleil, jusqu'au moment où un autre éminent astronome anglais – qui ressemble beaucoup à l'auteur – se dit que ce nuage noir n'est peut-être pas ce que nous croyons. Peut-être [que] c'est une forme de vie ; à l'aide d'un radio-téléscope, il arrive à entrer en contact avec lui. Le nuage noir est très ancien, très intelligent ; il est capable de discuter et au bout d'une assez longue discussion – qui est [en fait] le sujet du livre – le nuage noir dit que ça lui est égal [et qu']il ira manger un autre soleil autour duquel il n'y a pas de vie.

« *A for Andromeda* » de Fred Hoyle et John Elliot[9] est un autre roman, un peu raté vers la fin, mais qui part d'une idée très belle et aussi originale[10]. [Les auteurs imaginent] que le contact avec une vie extérieure peut avoir lieu sans support matériel. Comment ? Eh bien c'est très simple. C'est encore un radio-astronome anglais qui reçoit [directement un message] d'une galaxie – la galaxie d'Andromède ; on s'aperçoit que [l'émission est] un message certainement significatif [formé de points et de traits], on l'enregistre [...]. L'enregistrement [montre que le message] est répété en permanence. Je crois que l'émission dure deux ou trois jours et est réémise sans arrêt. Alors évidemment, les grands esprits se mettent à étudier ce message et à essayer de lui trouver une signification. On trouve assez vite que c'est un plan d'ordinateur suivi d'un programme. On réalise l'ordinateur, [et lorsqu'on implémente] le programme, la pensée des Andromédiens [s'anime]. La première chose que fait l'ordinateur une fois programmé, c'est d'interroger les Terriens pour savoir quelle est la chimie de la planète. Il faut bien penser que ce message est envoyé d'Andromède dans toutes les directions, donc qu'il s'adresse à toutes les formes de vie éventuelles. L'ordinateur donne des réponses et peu à peu, il construit par un procédé biologique une femme qui est dépositaire de la pensée d'Andromède : la vie extra-terrestre est [ainsi] parvenue sur terre par l'intermédiaire d'un message codé. [Mais ensuite] quoi faire de cette femme, évidemment ?

Ça c'est la description de la vie extra-terrestre avec laquelle nous avons des rapports d'observateurs. Mais il y a une autre question, [celle] des rapports plus passionnés, plus passionnels entre la vie extra-terrestre et la vie terrestre. Alors là, c'est un thème qui est tout à fait

8. N.D.L.R. : « *Le Nuage Noir* » de Fred Hoyle ; titre original « *The Black Cloud* » (1957), traduit par Jean Queval (Dunod, 1962).

9. N.D.L.R. : (D'après Wikipedia-en) : Norman James de la BBC, contacta Hoyle pour les droits de la retransmission télé de « *The Black Cloud* », mais Hoyle était intéressé par l'idée d'un nouveau scénario. Lors d'une rencontre entre Norman, John Elliot et Hoyle, ce dernier expliqua les grandes lignes d'une histoire en 8 épisodes qui devait devenir « *A for Andromeda* ». Hoyle s'est inspiré du travail de l'astronome Frank Drake, qui à cette époque avait initié le « *Projet Ozma* », une des premières expériences de recherche d'intelligence extra-terrestre, en d'autres termes l'ancêtre du projet SETI (Search for Extra-Terrestrial Intelligence).

10. N.D.L.R. : On peut remarquer les lignes communes entre le scénario de « *A for Andromeda* » de Hoyle & Elliot et celui du roman de Carl Sagan intitulé « *Contact* » (1985) adapté au cinéma par Robert Zemeckis en 1997. Notons aussi qu'en 1974, Souriau rédige ici tout un paragraphe autour de cette idée de *contact* !

un classique de la science-fiction du premier contact. Il y a d'ailleurs une nouvelle (malheureusement je ne sais plus l'auteur mais peut-être que quelqu'un pourra le dire) qui s'appelle « *Premier contact* »[11] qui raconte simplement comment [...] un vaisseau blanc, terrien, en mission quelque part [...] rencontre un vaisseau jaune qui ne lui dit rien qui vaille. [Les protagonistes] s'aperçoivent qu'ils sont vraiment étrangers [...] : le récit [décrit alors] un petit ballet entre les [deux vaiseaux, où] l'envie d'entrer en contact [se mêle à la] terreur [qu'ils s'inspirent mutuellement]. Le thème a été repris par un Soviétique qui a dit : « *voilà bien les Américains qui ne pensent qu'à la terreur* », et la même histoire a été racontée avec une fraternité immédiate. Les applications politiques sont évidentes.

FIGURE 6 *: Couverture de la première édition (australienne) de « First Contact » de Murray Leinster (1945).*

Mais après ce premier contact [...], il y a les autres contacts des terriens avec les planètes. Bien entendu, la première phase, une des plus anciennes, c'est la colonisation. Les Terriens vont coloniser ; enfin vous savez, quand on a débarqué sur la Lune on a fait la conquête de la Lune ; [la colonisation c'est] conquérir et implanter des colons. C'est une idée d'autant plus à la mode quand on pense à la surpopulation de la Terre. Alors là, il y a évidemment toute une littérature en quelque sorte sur ce sujet qui a des implications politiques. Je ne citerai qu'un exemple qui est un ouvrage de Francis Carsac qui s'appelle « *Ce monde est notre* » : ça veut dire, nous colons terriens, nous sommes chez nous sur ce monde... Cet ouvrage est paru pendant la guerre d'Algérie : on a interviewé Francis Carsac qui a répondu qu'il n'avait pas fait le rapprochement – ce qui est probable d'ailleurs. Je vous signale que Francis Carsac est un Universitaire : son vrai nom c'est Bordes et il est professeur de préhistoire à l'université de Bordeaux.

Je continue les contacts [...] avec le thème de l'adaptation, [disons de] l'adaptation de l'homme à la vie étrangère. Il y a énormément de variations sur ce thème [de] l'homme qui est complètement remodelé : les os sont remplacés par de la glace pour vivre sur les planètes glacées comme Jupiter ou Saturne et il devient incapable de revenir sur Terre, ce qui pose des problèmes psychologiques. Je voudrais simplement insister sur [le fait que] c'est par cette voie des problèmes d'adaptation aux planètes étrangères, qu'est apparu tout un volet de la science-fiction, qui est la science-fiction biologique. Une nouvelle qui doit dater des années 50 s'appelle « *L'Axolotl* »[12] [de Robert Abernathy]. [L'axolotl est un véritable] animal mexicain – je crois – qui n'est pas très beau [...] ; il vit sous sa forme d'axolotl dans presque tous les cas, mais dans des circonstances favorables il évolue et il atteint son stade de développement final [sous la forme d'une] salamandre dorée. C'est donc en fait une larve, dont l'état parfait n'apparait

11. N.D.L.R. : Il s'agit probablement du roman « *First Contact* » de Murray Leinster (1945) qui a remporté le « *Retro Hugo Award for Best Novelette* » en 1996.

12. N.D.L.R. : « *L'Axolotl* » (titre original : « *Axolotl* ». Nouvelle de Robert Abernathy, initialement parue dans « *The Magazine of Fantasy & Science Fiction* » n°32, janvier 1954. Sa version française est publiée en livre de poche dans « *La Grande Anthologie de la Science-Fiction* » dans le volume « *Histoires de Cosmonautes* ». Résumé : un homme décide de partir dans l'espace afin de découvrir ce qui tue les cobayes envoyés jusque-là. On le croit mort, mais il subit une étonnante métamorphose, qui lui ouvre un autre univers.

que dans des conditions particulières [13]. Dans la nouvelle d'Abernathy, l'homme, de même que l'axolotl, est une larve et quand il ira dans l'espace il se transformera, il prendra son état pour lequel il est fait qui est quelque chose de très beau. [L'auteur ne nous] donne évidemment qu'une description approximative.

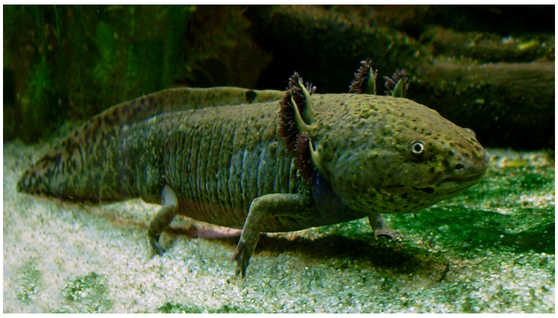


FIGURE 7 *: (Source Wikipédia) : Un axolotl dans son milieu naturel.*

Bien entendu, [cette histoire de métamorphose] se rapproche des histoires des mutants qui sont innombrables et en particulier des mutants télépathes. La télépathie a été utilisée très largement dans cette littérature, pour des raisons sentimentales, parce qu'on aime bien la télépathie. Enfin, c'est beau de voir qu'on peut se comprendre mutuellement, mais [il y a] aussi [...] des raisons pratiques : quand on fait débarquer un Terrien sur une planète étrangère – ça se produit 50 fois par mois dans la littérature de science fiction – [...] s'il faut qu'il commence à apprendre la langue, [on tombe vite dans] un passage à vide du roman, ce qui est très ennuyeux ... Alors, c'est tellement plus simple et naturel [d'avoir des terriens] télépathes : on se comprend sans avoir à parler. Il y a une variante qui est l'appareil, l'ordinateur projecteur de l'homme ; mais je crois qu'on y renonce un peu parce que les essais techniques ne réussissent pas, alors on préfère ne pas en parler : c'est tellement plus simple d'être télépathe pour se comprendre avec la forme de vie étrangère.

Je voudrais aller un peu plus loin. Tout ça se sont des variantes ; la science-fiction biologique est quelque chose d'intéressant [...] : ça conduit souvent à des vues à plus long terme et des vues qui comportent une option philosophique sur le destin de l'humanité ; mais je ne vais pas en parler aujourd'hui. Un ouvrage qui est très beau c'est « *Les enfants d'Icare* » par Arthur Clark et qui fait comprendre progressivement – c'est un roman très long – que l'Homme a vraiment quelque chose de plus que les autres formes de vie. Il y a d'autres formes de vie mais c'est l'enfant chéri de la Galaxie et il est destiné en quelque sorte à en devenir le maître. Il y a un triomphalisme de l'espèce humaine qui est traité avec des applications différentes ; [dans le roman de Clark] c'est [écrit] sur un mode assez poétique et assez saisissant.

Bien entendu, tout ce que je viens de dire comprend sa réciproque. Le sujet réciproque, c'est qu'au lieu que nous soyons des colonisateurs nous sommes des colonisés. Ce thème est très ancien et il a été divulgué de façon très spectaculaire par un film [...] que beaucoup d'entre-vous ont peut-être vu, « *La planète des Singes* » [...]. Je n'en cite qu'un passage. C'est les singes de la classe dirigeante (il y a deux classes), qui font une partie de chasse ; ils chassent les hommes, les hommes sauvages dans la brousse ; ensuite on voit la photographie du chasseur devant son trophée, fait de corps humains accrochés à des piques. Je peux vous le dire pour l'avoir constaté : ça produit un choc psychologique étonnant sur [les ames sensibles] qui vont voir un film pour s'amuser. La terreur qui règne dans la salle devant cette idée que l'homme peut être un gibier pour les singes : on entend, on sent le frisson de terreur, alors qu'on montre sans arrêt des horreurs [...] dans la plupart des films [qui n'ont pas le même effet]. Mais celle-là frappe : pourquoi frappe[-t-elle] plus que des choses bien plus graves, bien plus horribles, comme les tortures ? Il me semble qu'il y a une conscience d'espèce qui joue par rapport à ce que je disais tout à l'heure.

Bon, en allant un tout petit peu plus loin dans ce thème, nous sommes colonisés, oui mais nous ne le savons pas. Alors là le thème est rebattu : nous sommes colonisés par des gens qui nous surveillent qu'on appelle les « *Grands*

13. N.D.L.R. : Le terme *axlotl* est aussi utilisé par Frank Herbert tout au long du cycle de « *Dune* », pour désigner une technologie mystérieuse développée par les *Tleilaxu* : les *cuves axlotls* (*Axlotl tanks* en anglais). Cette technologie est un élément clef de l'univers de « *Dune* » : elle permet la régénéressence de cellules d'un individu mort, afin de récréer une sorte de clone appelé *Ghola* et ayant le potentiel d'accéder – dans des conditions de stress très particulières – à sa (ses) mémoire(s) antérieures(s) – vue(s) comme une sorte de mémoire atavique... Chaque tome du cycle dévoile un peu plus l'horreur des *cuves axlotls*, liée au statut de la femme dans la civilisation *Tleilaxu*...

Galactiques » [14]. Ils nous surveillent comment ? Bien évidemment avec des soucoupes volantes qui sont les témoins de cette surveillance. A ce moment-là, la science-fiction déborde sur une croyance. Il existe une proportion non négligeable des Français et même certains astronomes qui pensent que les soucoupes volantes sont une réalité, ne sont pas une invention d'une folie collective et qu'elles viennent effectivement des extra-terrestres. Un autre argument qui est souvent donné par les tenants de ce point de vue, c'est l'archéologie mystérieuse [...] : une terrasse qui se trouve au Liban [15] était une rampe de lancement des Grands Galactiques la dernière fois qu'ils sont venus nous voir ; la sagesse mystérieuse des pharaons revient, c'est une très vieille histoire. Il y a toute une attitude philosophique qui c'est manifestée [avec le livre de Louis Pauwels et Jacques Bergier] *« Le matin des magiciens »* (1960) [16] que beaucoup d'entre vous connaissent : ce livre, qui jouait cette carte-là, a eu un succès considérable. On va un peu plus loin bien entendu, non seulement nous sommes colonisés, nous sommes surveillés mais nous sommes des extra-terrestres. Un jour les extra-terrestres qui étaient des hommes ont débarqué, ont colonisé la Terre et nous sommes leurs descendants ; nous n'avons rien à voir avec la Terre, nous ne sommes pas des indigènes. C'est difficile à soutenir ne serait-ce qu'au point de vue biologique. La démonstration a été faite puisque le code génétique de nos cellules était rigoureusement le même que celui des plantes ; donc on ne peut pas tenir cette idée. Alors, la seule chose à dire, c'est que les plantes aussi sont extra-terrestres, mais à ce moment-là toute vie sur terre est extra-terrestre et on retrouve une vieille théorie philosophique qu'on appelle la *transpermie* qui veut bien dire ce que ça veut dire, que la vie est envoyée, elle se propage à travers l'Univers, elle se dépose : les graines tombent sur les planètes, elles se déposent et elles fructifient. Je dois dire pour être tout à fait honnête, qu'il y a des découvertes astronomiques récentes qui vont un tout petit peu dans ce sens-là. On a découvert depuis 5 ou 6 ans dans la Galaxie, dans les régions centrales de la Galaxie – enfin entre le centre de la Galaxie et le Soleil – des nuages qu'on a pu analyser par radio-téléscope [...] ; [on y a] découvert [...] de l'eau, du carbone, de l'hydrogène et des acides aminés qui sont le support de la vie sur Terre. Une découverte récente qui a été faite l'année dernière c'est que ces nuages forment effectivement comme je l'ai dessiné ici une sphère centrée sur le centre de la Galaxie ; [elle a un] mouvement d'expansion rapide qui a dû commencer il y a peu de temps à l'échelle galactique, il y a une centaine de millions d'années. De là à dire que le centre de la Galaxie envoie des graines, envoie ses matières organiques qui vont ensemencer les planètes, il n'y a qu'un pas que l'on peut franchir ; c'est une question personnelle.

Bon, j'ai fait cet étalage, cette classification des thèmes mais j'ai oublié le principal et à mon avis le seul, le véritable thème de la science-fiction : c'est le temps. La meilleure preuve, disons que le voyage spatial, la conquête de la Lune était de la science-fiction pour Jules Verne et n'en est plus maintenant. Pourquoi ? Parce que pour Jules Verne c'était le futur et que maintenant c'est le présent : par conséquent la science-fiction, ce n'est pas ce qui se passe ailleurs : c'est ce qui se passe à un autre moment. On parle *d'utopie* pour la science-fiction : il serait plus sage de parler *d'uchronie* – le mot est parfois utilisé – c'est-à-dire quelques chose qui sort du temps tel que nous le concevons.

---

14. N.D.L.R. : Ou « *Grands Anciens* » sont des créatures extraterrestres fictionnelles, originellement issues de l'oeuvre Howard Philipps Lovecraft.

15. N.D.L.R. : Cette terrasse se trouve à Tyr, ville de Phénicie.

16. N.D.L.R. : (D'après Wikipédia-fr) : « *Le Matin des magiciens – introduction au réalisme fantastique* » est un livre de Louis Pauwels et Jacques Bergier publié en octobre 1960 et se présentant comme une « *introduction au réalisme fantastique* ». Cet ouvrage de plus de 500 pages dans son édition originale se présente comme un récit, « *parfois légende et parfois exact* », consacré à « *des domaines de la connaissance à peine explorés [...] aux frontières de la science et de la tradition* ». Son contenu aborde des thèmes aussi divers que l'alchimie, les sociétés secrètes, les civilisations disparues, les récurrences insolites, les religions et les sciences occultes ou l'ésotérisme. Il repose sur des témoignages anciens (comme « *les manuscrits de la mer Morte* »), des recherches et des livres d'auteurs reconnus ou méconnus, des articles de revues spécialisées et des ouvrages de science-fiction ou de littérature fantastique. Ce livre, véritable phénomène éditorial, vendu à un million d'exemplaires, a remis au goût du jour le réalisme fantastique, inspiré la revue « *Planète* » et la collection « *L'Aventure mystérieuse* », où Jacques Bergier publiera plusieurs ouvrages.

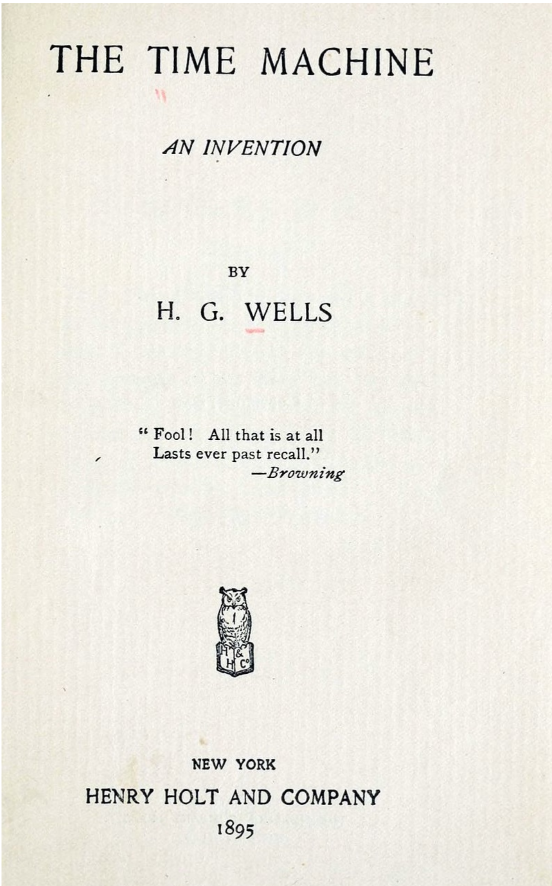

THE TIME MACHINE
AN INVENTION
BY
H. G. WELLS
" Fool! All that is at all
Lasts ever past recall."
—*Browning*
NEW YORK
HENRY HOLT AND COMPANY
1895

FIGURE 8 : *Couverture de l'édition originale de « The Time Machine ».*

Il y a une référence littéraire que je suis obligé de faire : c'est un roman de 1895 [...] écrit par un commis drappier qui avait fait des études et qui est tout d'un coup devenu célèbre à l'âge de 29 ans : il s'agit de Herbert George Wells bien entendu. Ce roman a produit l'effet d'une bombe et a rendu Wells célèbre d'un seul coup : [c'est *« La machine à explorer le temps »*]. Je pense que la plupart d'entre vous l'ont lu. C'est le récit d'un Anglais qui part avec une machine de son invention dans le futur, qui vit quelques temps la vie des gens d'un futur, puis qui vit quelques temps la vie des gens d'un futur très lointain ; ça se passe délibérément à quelques dizaines de millions d'années en avant. [A la fin, le héros] revient chez ses amis et bien entendu, c'est l'occasion pour l'auteur de décrire sa vision de l'évolution ultérieure de la Société.

Il faut dire que ça a paru nouveau ; mais enfin, Platon décrivant l'Atlantide [...] dans un passé lointain, se plaçait dans les mêmes conditions uchroniques qui permettent de décrire la société telle qu'on la conçoit, qu'on la souhaite ou telle qu'on craint qu'elle soit. Je ne parle pas d'innombrables romans qui décrivent ce qui se passe après la destruction de la Terre par la guerre atomique, la renaissance, la reprise de la vie commune par des petits groupes. Il est évident que c'est l'occasion – disons le prétexte même – pour chacun de décrire son modèle de société [sous forme d'une] science-fiction sociologique. Mais ce qui [tient le] plus [de la] science-fiction proprement dite, ce sont les paradoxes temporels. Ces paradoxes on peut quand même les décrire. Je fais un petit dessin... Voici le temps qui s'écoule disons comme ceci. Qu'a fait le héros de Wells ? Eh bien il s'est séparé du temps commun à tout le monde, il est allé dans le futur et il est revenu par le même procédé. Wells donne des explications qui sont disons, non satisfaisantes, mais enfin je ne peux que décrire le schéma de son roman.

Il est bien évident que dès que l'on parle de voyage dans le temps, le voyage dans le futur [tel que décrit par Wells] ne pose pas de trop graves problèmes. Il y a [cependant] des choses qui posent des problèmes beaucoup plus graves : c'est le voyage dans le passé. Bien entendu, il y a une idée qui vient de suite à l'esprit ; alors, c'est le sujet d'un roman de Barjavel – paru en 43 je crois [17] – et qui s'appelle *« Le voyageur imprudent »*. C'est le voyageur qui se rend dans le passé, qui veut sauver Bonaparte au siège de Toulon, qui détourne un coup de pistolet qui menaçait Bonaparte, qui tombe sur un soldat [...] à côté et qui par malheur était l'arrière-grand-père du voyageur, à la suite de quoi tout disparait. Il y a là une paradoxe temporel auquel il est difficile de toutes façons d'échapper. Si on va modifier le passé qu'est-ce qu'il va arriver ? Barjavel s'en tire par une pirouette : l'homme devient flou, il oublie tout, il disparait, tout le monde oublie, enfin ça s'arrange. Un autre exemple qui est une nouvelle qui s'appelle *« L'enfant en proie au temps »* [18] et qui décrit la vie d'une fille-mère [...]. Vous devinez à peu près ce qui se passe : c'est une

17. N.D.L.R. : (d'après Wikipédia-fr) : Le roman connaît une prépublication, sous forme de feuilleton, en 1943, avant d'être édité en volume l'année suivante.

18. N.D.L.R. : (d'après Wikipédia-fr) *« L'Enfant en proie au temps »* de Charles L. Harness : édition française de 1956 trad. Alain Dorémieux. Paru initialement en 1953 dans *The Magazine of Fantasy & Science Fiction no 25 ; in Fiction no 26*, sous le titre original *« Child by Chronos »*.

jeune fille qui vit avec sa mère, sa mère qu'elle déteste, sa mère qu'elle trouve possessive. La mère a un métier très bizarre : elle fait des pronostics ; elle [ne fait faire qu']une seule chose à [...] sa fille, c'est apprendre les titres des journaux ; elle la manipule de façon très bizarre. La fille ne va pas à l'école, puis un jour [arrive] un individu qui est l'amant de sa mère [...] et elle finit par avoir une liaison coupable avec [lui]. A ce moment-là, il y a une petite histoire de fuite dans le temps ; [la fille] revient dans le passé et elle s'aperçoit qu'elle est sa propre mère et [que] l'enfant qu'elle va mettre au monde c'est elle-même.

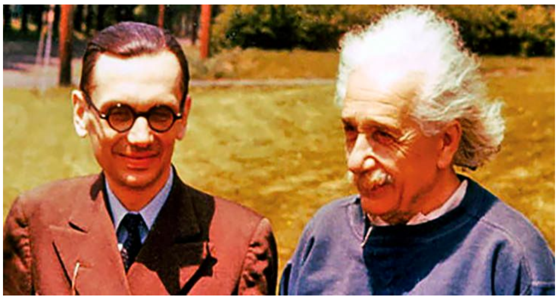


FIGURE 9 : *Kurt Gödel et Albert Einstein à Princeton vers 1948.*

Bon, bien entendu on peut aller un petit peu plus loin encore. Au lieu de faire ces dessins là où il y a un petit épisode dans la ligne de temps, je vais vous expliquer quelque chose que tout le monde va comprendre. C'est l'éternel retour, c'est-à-dire l'idée que l'histoire – l'évolution – va se fermer [...]. Dans la cosmologie, il n'y a rien qui empêche cette possibilité ; en particulier il y a un modèle d'univers qui a été donné par le logicien [Kurt Gödel]. C'est le seul travail [qu'il a fait en] relativité générale : [il propose] un modèle d'univers dans lequel il y a l'éternel retour, c'est-à-dire que les lignes de temps décrites par les galaxies sont fermées sur elles-mêmes. Bien entendu, cette idée-là permet de jouer avec toutes les idées précédentes. Par exemple les « *Grands Galactiques* » qui sont venus ensemencer la Terre, c'est nous, c'est nos enfants ; à ce moment-là, nous ne sommes pas vraiment colonisés, nous sommes fiers de notre passé qui est notre avenir. Il y a même un thème qui est tellement utilisé qu'il est devenu une obsession chez *"l'auteur américain"* : c'est l'homme qui s'aperçoit, [une fois] la fin du monde arrivée, [qu'il est] le dernier survivant sur Terre : [il réalise alors] avec lassitude – [cette situation] arrive très très souvent – qu'il reste justement une bonne femme qu'il rencontre dans un endroit des plus biscornus et il lui demande son nom, et elle dit : on m'appelle Eve. Il y a même une histoire qui se termine par les lignes suivantes : « *avec l'aisance que donne une longue habitude, le héros s'aperçut qu'une fois de plus il était le premier homme sur Terre.* »

Bon, il reste un dernier thème qui est disons la quintessence [du thèmes des paradoxes temporels] : je vais le dessiner... Vous comprendrez tout de suite, [c'est] les univers parallèles [...]. Tout ce que je viens de dessiner ici, [ce sont] des variations topologiques sur le temps. L'une des topologies consiste à dire que l'univers est composé de beaucoup d'univers parallèles et qu'avec un peu de chance il y a moyen de passer de l'un à l'autre. Il y a même des gens qui écrivent – puisqu'il existe une infinité d'univers – que tous les univers imaginables sont réalisés, ce qui est quand même un peu dur pour un mathématicien. Là, bien entendu, vous pouvez imaginer beaucoup de choses. En particulier le voyage dans le temps n'est plus vraiment le voyage dans le temps : qui sait si en passant d'un point à un autre ici on ne trouvera pas une planète qui ressemblera tout à fait à la nôtre, mais dans laquelle il y a un décalage temporel. Il y a des récits [...] mémorables de Poul Anderson [19] qui raconte un univers parallèle où par hasard, il y a aussi une modification temporelle : [par exemple un univers parallèle où] c'étaient les Gaulois qui avaient gagné : leur monde avait une technologie bizarre avec un air français [complètement dépassée].

Il y a une chose curieuse [à propos de] la dernière interaction entre science-fiction et science que je vous ai signalée. C'est que cette idée, qui [...] est vraiment une idée de fiction, une idée d'auteur en quête de thème, a été reprise récemment par des physiciens très sérieux [sous forme d'une théorie]. Cette théorie a d'ailleurs été esquissée à Marseille par Jean-Marc Lévi-Leblond. Chacun sait qu'il reste encore dans la mécanique quantique des paradoxes, des difficultés d'interprétation que l'école de Copenhague n'arrive pas encore à surmonter. [Lévi-Leblond montre qu'on peut] supposer effectivement qu'il y a des univers parallèles et que toute transition quantique consiste à passer d'un univers à l'autre. On saute d'un univers à l'autre, mais à ce moment-là, chaque point de l'univers est une bifurcation, c'est-à-dire que

19. N.D.L.R. : « *La patrouille du temps* » (1960).

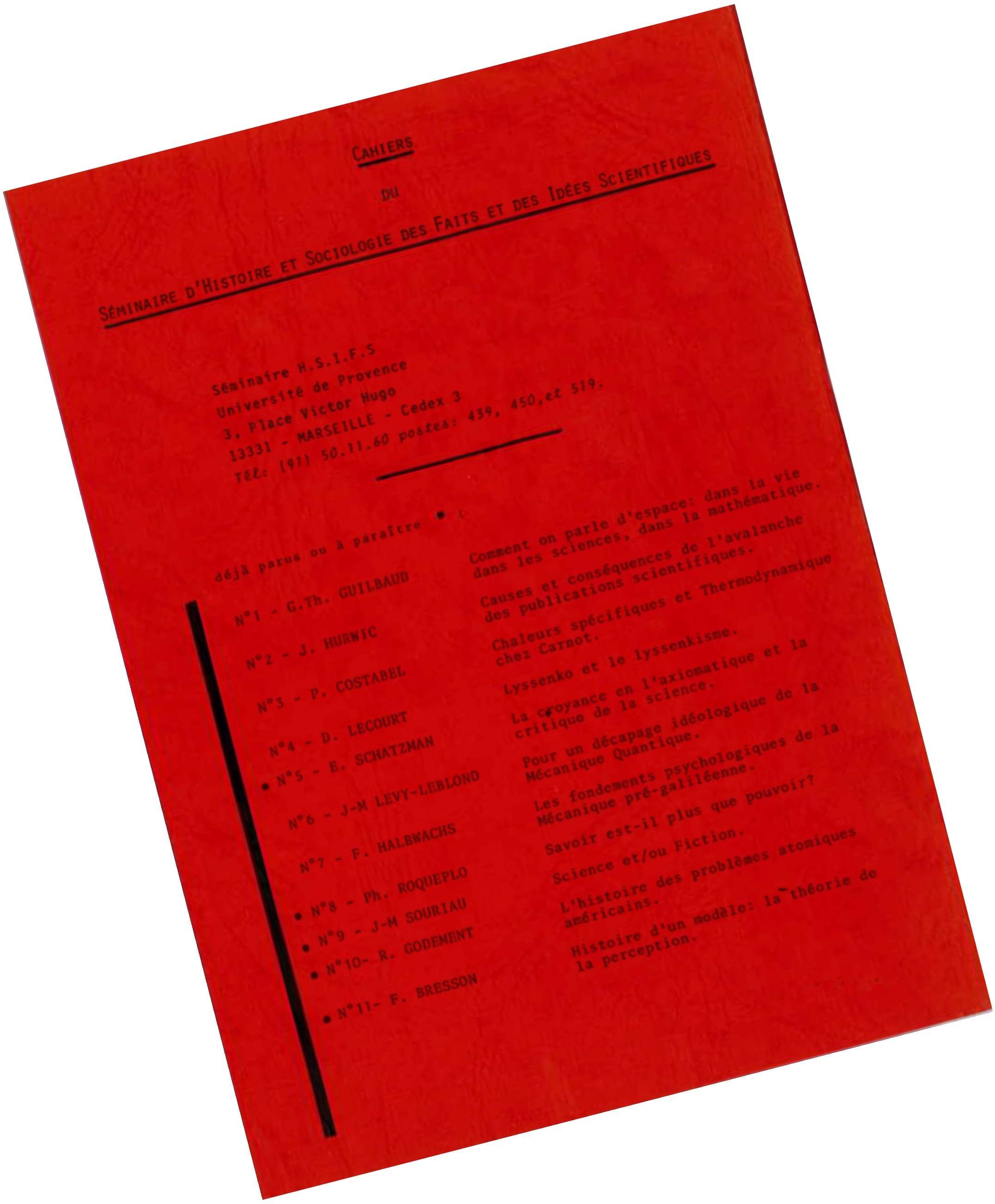
c'est un faisceau de faisceaux, quelque chose d'extrêmement touffu. [Lévi-Leblond] qui est venu ici même il y a quelques mois, nous a dit qu'à son avis c'était une explication du paradoxe du hasard et de la nécessité, du paradoxe de l'évolution. L'évolution sur Terre s'est faite par des hasards dont on n'arrive pas à comprendre [le mécanisme], [ou plus précisément, à imaginer un] mécanisme qui auraient pu tenir depuis les quatre [milliards] d'années que la Terre existe. Et là il y a un paradoxe effectif : on n'a pas d'explications possibles, [c'est-à-dire] que la sélection naturelle ne peut pas expliquer [cela] pour l'instant [...] Si nous existons c'est que nous sommes dans une ligne univers sélectionnée ; c'est assez extraordinaire, mais ce n'est pas extraordinaire que nous y soyons puisque [cette ligne d'univers] a été choisi. Autrement dit [...] il y a des chances qu'il existe une ligne d'univers qui conduit à [notre] évolution et par conséquent – puisque nous vivons – c'est que nous sommes dessus [...] : le hasard c'est [donc] simplement la sélection d'une ligne de l'univers, enfin d'une ligne d'inter-univers parmi de très nombreuses.

Bon, je crois qu'on ne peut guère aller plus loin dans l'état actuel dans ces thèmes et par conséquent je serais content qu'une discussion commence. [Mais avant cela], je voudrais dire [...] que c'est Monsieur Meyer [20] qui [...] a attiré mon attention sur Gassendi et je l'en remercie.

-o-o-o-o-o-o-o-o-o-o-o-o-

(le 5-6-74).

20. N.D.L.R. : François Meyer auteur de « *Gassendi et Descartes* » publié dans les actes du Congrès « *Tricentenaire de Pierre Gassendi, 1655-1955* ». (Actes du Congrès publiés chez PUF - Paris - 1957.)



CAHIERS

DU

SÉMINAIRE D'HISTOIRE ET SOCIOLOGIE DES FAITS ET DES IDÉES SCIENTIFIQUES

Séminaire H.S.I.F.S
Université de Provence
3, Place Victor Hugo
13331 - MARSEILLE - Cedex 3
Tél: (91) 50.11.60 postes: 439, 450, et 519.

déjà parus ou à paraître • :

N°1 - G.Th. GUILBAUD	Comment on parle d'espace: dans la vie dans les sciences, dans la mathématique.
N°2 - J. HURWIC	Causes et conséquences de l'avalanche des publications scientifiques.
N°3 - P. COSTABEL	Chaleurs spécifiques et Thermodynamique chez Carnot.
N°4 - D. LECOURT	Lyssenko et le lyssenkisme.
• N°5 - E. SCHATZMAN	La croyance en l'axiomatique et la critique de la science.
N°6 - J-M LEVY-LEBLOND	Pour un décapage idéologique de la Mécanique Quantique.
N°7 - F. HALBWACHS	Les fondements psychologiques de la Mécanique pré-galiléenne.
• N°8 - Ph. ROQUEPLO	Savoir est-il plus que pouvoir?
• N°9 - J-M SOURIAU	Science et/ou Fiction.
• N°10- R. GODEMENT	L'histoire des problèmes atomiques américains.
• N°11- F. BRESSON	Histoire d'un modèle: la théorie de la perception.

VOUZZAVEDIBISAR : H.S.I.F.S. (par Université de Provence)

(2016) B.I.A.A. no **104** BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
ISSN 0291-5413 (Université d'Aix-Marseille)

Mécanique quantique II : opérateurs non bornés

Eric OLIVIER [1] [2]

Résumé. – Le « *quantification* » du modèle (hamiltonien) d'un système mécanique classique se fait grâce « *au principe de correspondance* » de Schrödinger : celui-ci permet d'associer à chaque quantité (classiquement) mesurable du système, une « *observable quantique* », c'est-à-dire un opérateur autoadjoint agissant sur un espace de Hilbert. Du point de vue quantique, les valeurs mesurables sont – *en principe* – les valeurs spectrales des observables. Cependant – et même dans les cas les plus simples – ces opérateurs ne sont pas définis sur l'espace tout entier, c'est-à-dire qu'ils ne sont pas bornés : un des buts de cette note est de voir comment aborder l'analyse spectrale de ces opérateurs. L'observable énergie, c'est-à-dire le hamiltonien quantique H, joue un rôle particulièrement important : ses valeurs spectrales déterminent les niveaux d'énergie auxquels peut accéder le système : la théorie des quanta commence lorsque Planck réalise que pour certains systèmes, les niveaux d'énergie sont nécessairement discrets. En général un système quantique possède un spectre d'énergie qui peut être discret, continu ou mixte : cet aspect est déterminant quant à la résolution de l'équation linéaire d'évolution quantique associées à H – i.e. « *l'équation de Schrödinger* : nous verrons comment le « *théorème de Stone* » permet une résolution abstraite de cette équation indépendamment de la nature spectrale de H.

1. **Introduction**

1.1. Le *« point matériel »* de la mécanique classique (hamiltonienne) est représenté comme un point mathématique évoluant dans un espace de dimension finie appelé *« espace des phases »*. La dimension de cet espace peut être très grande (problème à n corps), mais elle est toujours finie. Si le système étudié est instable – ou disons, s'il y a *« sensibilité aux conditions initiales »* – alors on s'intéresse à l'évolution d'une petite portion de l'espace des phases et l'objet étudié devient de dimension infinie. Cette procédure revient à faire tendre le nombre de points matériels vers l'infini, c'est-à-dire – en un sens – à remplacer la mécanique du point par une *« mécanique des fluides »*. L'électromagnétisme de Maxwell est une théorie des courants électriques construite sur le modèle de la mécanique des fluides : mais, malgré son efficacité à décrire les phénomènes électromagnétiques (et sa beauté mathématique) elle a aussi abouti à deux contradictions. L'une d'elle – le rayonnement du corps noir – a amené Planck à introduire le quantum d'action (l'autre contradiction a donné la relativité restreinte d'Einstein).

1.2. Soit X (sous-ensemble de l'espace euclidien $\mathbb{R}^n$), l'espace des positions d'un système mécanique classique. Dans la description quantique de ce système, les points de X (supposés être les positions classiques possibles) sont remplacés par des *« états quantiques »* que nous appellerons *« Q-états »*. Ce sont des éléments de la sphère unité de l'espace de Hilbert complexe $\mathscr{H} = L^2(X)$ des fonctions $\psi : X \to \mathbb{C}$ mesurables dont le

1. GDAC-I2M UMR 7373 CNRS Université d'Aix-Marseille
2. eric.olivier@univ-amu.fr

module est de carré Lebesgues-intégrable. Le produit scalaire de deux éléments ψ et ϕ de $\mathscr{H}$ est alors défini comme la quantité

$$\langle\psi|\phi\rangle = \int \psi(x_1,\ldots,x_n)\overline{\phi(x_1,\ldots,x_n)}dx_1\cdots dx_n = \int \psi(x)\overline{\phi(x)}dx \tag{1}$$

(où $x = (x_1,\ldots,x_n)$ et $dx = dx_1\cdots dx_n$) de sorte que $\|\psi\| = \langle\psi|\psi\rangle^{1/2}$ est la norme hilbertienne de ψ. En utilisant les notations de Dirac, un Q-état est représenté par un *« ket »* $|\psi\rangle$, ce qui donne l'identification $\psi = |\psi\rangle$. L'utilisation de la notation en kets ne prend son sens qu'avec la notation *« bra »* en désignant par $\langle\psi|$ la forme linéaire canoniquement associée à ψ et telle que $\langle\psi|(\phi) = \langle\psi|\phi\rangle$.

1.3. Ainsi, les Q-états sont-ils des éléments ψ de $\mathscr{H}$ dont la norme vaut $\|\psi\| = 1$: lorsque cela a un sens, la valeur (complexe) d'un tel Q-état en un point de l'espace[3] s'appelle une *« amplitude de probabilité »*. Si le Q-état ψ représente un système quantique à un instant t de son évolution (gouvernée par *« l'équation de Schrödinger »* : c.f. [Oli16b]), alors la probabilité de présence de ce système dans une partie Y de X est égale à l'intégrale

$$\int_Y |\psi(x)|^2dx$$

La théorie des moments d'une distribution de probabilité permet alors – lorsque c'est possible – de définir la position suivant la i-ème direction cartésienne du système se trouvant dans le Q-état ψ par l'intégrale

$$\int x_i|\psi(x)|^2dx = \langle Q_i\psi|\psi\rangle \tag{2}$$

où nous avons posé $Q_i\psi(x) := x_i\psi(x)$. Nous reviendrons en détail sur cette interprétation probabiliste des Q-états (interprétation due à Born). Cependant l'équation (2) contient déjà le problème qui va nous occuper tout au long de cette première note. Remarquons d'abord qu'en posant « *naïvement* » $Q_i\psi(x) := x_i\psi(x)$, nous définissons un opérateur linéaire Q_i sur le sous-espace $\mathscr{G}_i$ de $\mathscr{H}$ formé des $\psi(x) \in \mathscr{H}$ t.q. $x_i\psi(x) \in \mathscr{H}$. Le sous-espace $\mathscr{G}_i$ étant dense dans $\mathscr{H}$, on dit que $(Q_i,\mathscr{G}_i)$ est *« un opérateur densément défini sur $\mathscr{H}$ »* ; de plus, lorsque X n'est pas une partie de $\mathbb{R}^n$ bornée suivant la i-ème composante cartésienne, l'application linéaire $Q_i : \mathscr{G}_i \to \mathscr{H}$, ne peut être prolongée en un endomorphisme borné[4] de $\mathscr{H}$. En effet, si ψ est un Q-état dans le domaine $\mathscr{G}_i$ de définition de Q_i, alors le produit scalaire $\langle Q_i\psi|\psi\rangle$ en (2) est par définition la positions du Q-état ψ suivant la i-ème direction cartésienne ; or il est assez simple de se convaincre que pour tout point x dans l'espace X, il existe un Q-état ψ t.q.[5] $\langle Q_i\psi|\psi\rangle = x_i$: par suite, si l'espace X n'est pas borné suivant la i-ème direction cartésienne, l'opérateur $(Q_i,\mathscr{G}_i)$ est lui-même nécessairement non borné.

1.4. Les grandeurs mesurables d'un système quantique décrit dans l'espace de Hilbert $\mathscr{H}$, sont représentées par des *« observables »*, c'est-à-dire par des opérateurs densément définis sur $\mathscr{H}$ et *« autoadjoints »*. La notion d'adjoint est classique pour les endomorphismes bornés de $\mathscr{H}$: nous verrons à la Section 3 comment étendre cette notion aux

3. Par exemple lorsqu'il existe une version continue dans de ψ dans $L^2(X)$.

4. C'est-à-dire que $\|Q_if\|$ n'est pas borné pour $\psi \in \mathscr{G}_i$ avec $\|f\| = 1$: ceci est équivalent à dire que l'application linéaire $Q_i : \mathscr{G}_i \to \mathscr{H}$ n'est pas continue (pour $\mathscr{G}_i$ muni de la norme induite par la norme hilbertienne de $\mathscr{H}$).

5. On peut même trouver ψ t.q. $\langle Q_i\psi|\psi\rangle = x_i$ pour tout $i = 1,\ldots,n$.

cas des opérateurs non bornés. Sans prétendre à une justification, remarquons simplement que les grandeurs qu'il est possible de mesurer sur un système physique sont des nombres réels : la définition d'une observable est ainsi cohérente avec le fait que les valeurs spectrales des opérateurs autoadjoints (non nécessairement bornés) sont réelles (c.f. Théorème 3.16). Nous verrons à la Section 4 (avec $X = \mathbb{R}$) que l'opérateur position $(Q, \mathscr{G}) := (Q_1, \mathscr{G}_1)$ défini ci-dessus est effectivement autoadjoint et que ses valeurs spectrales (ce seront dans ce cas des valeurs propres approchées) coïncident avec toutes les positions possibles c'est-à-dire – dans ce cas – $\mathbb{R}$ (on parle de spectre continu).

1.5. Le problème de *« la corde vibrante »* étudié par d'Alembert est directement lié au système quantique constitué d'un *« corpuscule (massif) dans une boîte unidimensionnelle »*. Dans la suite, nous identifions la boîte à l'intervalle $X = [0;1]$, l'espace de Hilbert étant égal à $\mathscr{H} = L^2[0;1]$ (pour la mesure de Lebesgue). Lorsque le corpuscule n'est soumis à aucune force (corpuscule libre) le hamiltonien quantique (Q-hamiltonien) c'est-à-dire l'observable représentant l'énergie du système est (dans un système d'unités adéquates)

$$Hf(x) = -\frac{d^2 f}{dx^2} \tag{3}$$

L'opérateur H est bien défini sur le sous-espace dense $\mathscr{C}^2[0;1]$ formé des $f \in L^2[0;1]$ qui sont de classe C^2 sur $[0;1]$; mais cet espace ne fait pas de $(H, \mathscr{C}^2[0;1])$ un opérateur autoadjoint. En effet, pour $f, g \in \mathscr{C}^2[0;1]$ nous pouvons effectuer *« une double intégration par parties »* ce qui donne :

$$\langle f | Hg \rangle = \left[f(x)\overline{g'(x)} \right]_0^1 - \left[f'(x)\overline{g(x)} \right]_0^1 + \langle Hf | g \rangle \tag{4}$$

Il est alors immédiat qu'en général $\langle f | Hg \rangle \neq \langle Hf | g \rangle$; mais si nous notons $\mathscr{C}^2_P[0;1]$, l'espace des restrictions à l'intervalle $[0;1]$ des fonctions 1-périodiques et de classe C^2 sur $\mathbb{R}$, alors il découle de (4) que $(H, \mathscr{C}^2_P[0;1])$ est un opérateur symétrique, en ce sens que $\langle f | Hg \rangle = \langle Hf | g \rangle$ dès que $f, g \in \mathscr{C}^2_P[0;1]$. Cependant $\mathscr{C}^2_P[0;1]$ n'est pas l'espace *optimal* sur lequel H est symétrique, ce qui indique que $(H, \mathscr{C}^2_P[0;1])$ n'est pas autoadjoint (nous préciserons les notions mises en jeu au § 3). La recherche du domaine du Q-hamiltonien H (i.e. du sous-espace dense $\mathscr{D}$ rendant le couple $(H, \mathscr{D})$ autoadjoint) est une étape déterminante de l'étude du système quantique étudié : pour *« le corpuscule dans la boîte »*, le domaine de H en (3) est l'espace $\mathscr{H}^2_P[0;1]$ des restrictions à l'intervalle $[0;1]$ des fonctions 1-périodiques deux fois dérivables au sens faible et de dérivée seconde dans $L^2[0;1]$ (c.f. § 6.1). Dans ce cas, il est très facile de trouver les niveaux d'énergie du système, c'est-à-dire les valeurs spectrales de $(H, \mathscr{H}^2_P[0;1])$. En effet, pour $\alpha \in \mathbb{R}$ nous avons $He^{i\alpha x} = \alpha^2 e^{i\alpha x}$: mais si nous prenons en compte le fait que $e^{i\alpha x}$ soit dans $\mathscr{H}^2_P[0;1]$, alors il est nécessaire que $1 = e^{i\alpha}$, ce qui équivaut à la condition $\alpha = 2k\pi$, avec $k \in \mathbb{Z}$. Or, $(e^{2ik\pi x}\ ;\ k \in \mathbb{Z})$ étant une base hilbertienne de $L^2[0;1]$, nous pouvons conclure que $\left\{ (2k\pi)^2\ ;\ k \in \mathbb{Z} \right\}$ décrit l'ensemble complet des niveaux d'énergie possibles pour le corpuscule dans une boîte. Nous avons ici un Q-hamiltonien $(H, \mathscr{H}^2_P[0;1])$ à spectre discret : ceci traduit exactement la quantification des niveaux d'énergie du système en question. Notons que les niveaux d'énergie ne sont pas bornés ce qui donne une justification évidente du fait que l'opérateur densément défini $(H, \mathscr{H}^2_P[0;1])$ est lui même non borné.

2. Analyse spectrale des opérateurs non bornés

2.1. Soit $\mathscr{D}$ un sous-espace dense d'un espace de Banach $(\mathscr{E}, \|\cdot\|)$; si A est une application linéaire de $\mathscr{D}$ dans $\mathscr{E}$ alors le couple *« $(A, \mathscr{D})$ est appelé un opérateur densément défini »*. Si A est borné sur la sphère unité de $\mathscr{D}$, alors, par *« le théorème de prolongement par densité »*, l'opérateur A se prolonge en un unique endomorphisme borné de $\mathscr{E}$ (et de même norme que A) : si ce n'est pas le cas *« l'opérateur $(A, \mathscr{D})$ est dit non borné sur $\mathscr{E}$ »*. Dans le cas général (i.e. borné lorsque $\mathscr{D} = \mathscr{E}$ ou non borné lorsque $\mathscr{D} \subsetneq \mathscr{E}$) *« l'ensemble résolvante $\Omega(A, \mathscr{D})$ »* est constitué des nombres complexes ξ pour lesquels $\xi I - A$ est un élément inversible de $L(\mathscr{D}, \mathscr{E})$ et dont l'inverse $R_\xi := (A - \xi I)^{-1}$ est un endomorphisme borné de $\mathscr{E}$: le nombre ξ est appelé une *« valeur régulière »* de $(A, \mathscr{D})$ et l'opérateur R_ξ (borné sur $\mathscr{E}$) est la *« résolvante (propre) en ξ »*[6].

Proposition 2.1. *Soit $(A, \mathscr{D})$ un opérateur densément défini sur $\mathscr{E}$; si $R_\lambda := (A - \lambda I)^{-1}$ et $R_\mu := (A - \mu I)^{-1}$ sont les résolvantes associées à λ et μ dans $\Omega(A, \mathscr{D})$, alors*

$$R_\lambda - R_\mu = (\lambda - \mu) R_\lambda R_\mu$$

en particulier nous avons la commutation $R_\lambda \leftrightarrow R_\mu$.

Preuve. Supposons que $(P, \mathscr{D})$ et $(Q, \mathscr{D})$ soient deux opérateurs densément définis sur $\mathscr{E}$ et supposons que P et Q soient aussi des éléments inversibles de $L(\mathscr{D}, \mathscr{E})$ d'inverses respectifs P^{-1} et Q^{-1} dans $\mathcal{L}(\mathscr{E})$; alors nous vérifions directement que

$$P^{-1} - Q^{-1} = P^{-1} Q Q^{-1} - P^{-1} P Q^{-1} = P^{-1}(Q - P) Q^{-1} \tag{5}$$

Le fait que $\lambda, \mu \in \Omega(A, \mathscr{D})$ assure (par définition) que $A - \lambda I$ et $A - \mu I$ sont deux éléments inversibles de $L(\mathscr{D}, \mathscr{E})$ dont les inverses $R_\lambda := (A - \lambda I)^{-1}$ et $R_\mu := (A - \mu I)^{-1}$ sont des endomorphismes bornés de $\mathscr{E}$; par suite en appliquant (5) avec $P = R_\lambda$ et $Q = R_\mu$ il vient

$$R_\lambda - R_\mu = R_\lambda \Big((A - \mu I) - (A - \lambda I) \Big) R_\mu = (\lambda - \mu) R_\lambda R_\mu$$

$\square$

2.2. Le *« spectre »* de $(A, \mathscr{D})$ est le sous-ensemble de $\mathbb{C}$ correspondant au complémentaire de l'ensemble résolvant, soit $\Sigma(A, \mathscr{D}) = \mathbb{C} \setminus \Omega(A, \mathscr{D})$: les éléments de $\Sigma(A, \mathscr{D})$ sont appelés les *« valeurs spectrales »*. Le *« spectre ponctuel »* est l'ensemble $\mathrm{Vp}(A, \mathscr{D})$ des *« valeurs propres »*, i.e. l'ensemble des nombres complexes ξ pour lesquels il existe $X \in \mathscr{D} \setminus \{0\}$ t.q. $AX = \xi X$: un tel X est appelé un *« vecteur propre »* de $(A, \mathscr{D})$. Il est immédiat que $\mathrm{Vp}(A, \mathscr{D}) \subset \Sigma(A, \mathscr{D})$, l'inclusion pouvant être stricte en dimension infinie. Le nombre complexe ξ est appelé une *« valeur propre approchée »* de $(A, \mathscr{D})$ s'il existe une suite $X_1, X_2, \ldots$ de la sphère unité de $\mathscr{D}$ (i.e. $X_n \in \mathscr{D}$ et $\|X_n\| = 1$), t.q. $\|AX_n - \xi X_n\| \to 0$ quand $n \to +\infty$. Nous noterons $\mathrm{Vp}'(A, \mathscr{D})$ l'ensemble des valeurs propres approchées de $(A, \mathscr{D})$.

Proposition 2.2. *Soit $(A, \mathscr{D})$ un opérateur densément défini sur un espace de Banach $\mathscr{E}$; alors le spectre $\Sigma(A, \mathscr{D})$ est un sous-ensemble fermé de $\mathbb{C}$ tel que*

$$\mathrm{Vp}(A, \mathscr{D}) \subset \overline{\mathrm{Vp}(A, \mathscr{D})} \subset \mathrm{Vp}'(A, \mathscr{D}) \subset \Sigma(A, \mathscr{D})$$

6. Pour $\xi \in \Omega(A, \mathscr{D}) \setminus \{0\}$ nous appelons $(I - 1/\xi A)^{-1}$ la *« résultante caractéristique en ξ »* : nous n'utiliserons pas ici cette notion (liée à la théorie des opérateurs de Fredholm).

Preuve. L'inclusion de $\mathrm{Vp}(A, \mathscr{D})$ dans son adhérence est triviale. Maintenant, supposons que $\xi_0, \xi_1, \ldots$ soit une suite de $\mathrm{Vp}(A, \mathscr{D})$ qui converge vers ξ ; si X_n est un vecteur unitaire t.q. $AX_n = \xi_n X_n$, alors $\|AX_n - \xi X_n\| = \|\xi_n X_n - \xi X_n\| = |\xi_n - \xi| \to 0$ quand $n \to 0$; cela démontre que l'adhérence de $\mathrm{Vp}(A, \mathscr{D})$ est un sous-ensemble de $\mathrm{Vp}'(A, \mathscr{D})$. Pour l'inclusion de $\mathrm{Vp}'(A, \mathscr{D})$ dans $\Sigma(A, \mathscr{D})$, supposons que ξ soit une valeur propre approchée $(A, \mathscr{D})$, i.e. qu'il existe une suite $X_0, X_1, \ldots$ dans la sphère unité de $\mathscr{D}$, t.q. $\|AX_n - \xi X_n\| \to 0$ quand $n \to +\infty$. Alors la condition $\xi \in \Omega(A, \mathscr{D})$ entraînerai par la continuité de R_ξ que $1 = \|X_n\| = \|R_\xi(AX_n - \xi X_n)\| \leq \|R_\xi\| \cdot \|AX_n - \xi X_n\|$, en contradiction avec le fait $\|AX_n - \xi X_n\| \to 0$ quand $n \to +\infty$: l'inclusion $\mathrm{Vp}'(A, \mathscr{D}) \subset \Sigma(A, \mathscr{D})$ est ainsi démontrée. Il nous reste à établir que le spectre $\Sigma(A, \mathscr{D})$ est fermé : pour cela nous allons vérifier que toute valeur régulière de $(A, \mathscr{D})$ est un point intérieur à la résolvante de $\Omega(A, \mathscr{D})$. Remarquons d'abord que ξ_0 est une valeur régulière que $(A, \mathscr{D})$ ssi 0 est une valeur régulière de $(A - \xi_0 I, \mathscr{D})$: comme le spectre (et la résolvante) de $(A, \mathscr{D})$ et de $(A - \xi_0 I, \mathscr{D})$ diffèrent l'un de l'autre par une translation de $\mathbb{C}$, nous aurons établi le résultat général annoncé si, partant de l'hypothèse que 0 est une valeur régulière de $(A, \mathscr{D})$, nous montrons que 0 est un point intérieur à $\Omega(A, \mathscr{D})$. Dans ce cas $(A, \mathscr{D})$ est une bijection de $\mathscr{D}$ sur $\mathscr{E}$ dont l'inverse $R_0 = A^{-1}$ est un endomorphisme borné de $\mathscr{E}$: il s'agit de démontrer qu'il existe $r_0 > 0$ tel que tout nombre complexe ξ est une valeur régulière de $(A, \mathscr{D})$ dès que $|\xi| < r_0$. Soit $r_0 = 1/\|R_0\|$ et soit $\xi \in \mathbb{C}$ donné et t.q $|\xi| < r_0$: alors $\|\xi R_0\| < 1$ et donc (formule d'inversion de von Neumann) $I - \xi R_0$ est un endomorphisme inversible de $\mathcal{L}(\mathscr{E})$ dont l'inverse $S_\xi := (I - \xi R_0)^{-1}$ est aussi dans $\mathcal{L}(\mathscr{E})$. Comme $R_\xi := R_0 S_\xi = S_\xi R_0 \in \mathcal{L}(\mathscr{E})$, il nous reste à démontrer que R_ξ est l'inverse de $A - \xi I$. En effet, si d'une part Y est pris arbitrairement dans $\mathscr{E}$ alors $R_\xi Y = R_0 S_\xi Y$ est dans l'image de R_0, c'est-à-dire dans $\mathscr{D}$ et de plus

$$(A - \xi I)R_\xi Y = (A - \xi I)R_0 S_\xi Y = S_\xi Y - \xi R_0 S_\xi Y = (I - \xi R_0)S_\xi Y = Y$$

Mais d'autre part, si $X \in \mathscr{D}$, alors :

$$R_\xi(A - \xi I)X = S_\xi R_0(A - \xi I)X = S_\xi X - \xi S_\xi R_0 X = S_\xi(I - \xi R_0)X = X$$

ce qui montre que $A - \xi I$ est une application bijective de $L(\mathscr{D}, \mathscr{E})$ dont l'inverse d'inverse $R_\xi = (A - \xi I)^{-1}$ est une endomorphisme borné de $\mathscr{E}$.

□

Notons $\ell^2(\mathbb{N})$ l'espace constitué des suites $X = (X_0, X_1, \ldots)$ de nombre complexes t.q. $\sum_k |X_k|^2 < +\infty$; alors l'application $\langle \cdot | \cdot \rangle$ de $\ell^2 \times \ell^2 \to \mathbb{K}$ et t.q. $\langle X|Y \rangle := \sum_{k=0}^\infty x_k \overline{y_k}$ définit un produit scalaire sur ℓ^2 associé à la norme $X \mapsto \|X\| := (\sum_k |x_k|^2)^{1/2}$ et faisant de $(\ell^2, \|\cdot\|)$ un espace de Hilbert (c.f. [Oli16a, § 5.6]).

Proposition 2.3. *Etant donnée $(\mu_n)_{n=0}^\infty$ une suite de nombres complexes, soit $(A, \mathscr{D})$ l'opérateur densément défini sur $\ell^2(\mathbb{N})$, de domaine*

$$\mathscr{D} = \left\{ X \in \ell^2(\mathbb{N}) \; ; \; \textstyle\sum_{n=0}^\infty |\mu_n x_n|^2 < +\infty \right\}$$

et t.q $(AX)_n = \mu_n x_n$; alors le spectre $\Sigma(A, \mathscr{D})$ est l'adhérence dans $\mathbb{C}$ du support de $(\mu_n)_{n=0}^\infty$.

Preuve. Comme nous savons (par définition de A), que $\mathrm{Vp}(A, \mathscr{D}) = \{\mu_n\}_n$ et que d'autre part, l'adhérence de $\mathrm{Vp}(A, \mathscr{D})$ est constituée de valeurs propres approchées, il nous reste à démontrer que tout nombre complexe ξ qui n'est pas dans l'adhérence de $\{\mu_n\}_n$ est une valeur régulière de $(A,\mathscr{D})$. Pour voir cela, soit Y arbitrairement donnée dans $\ell^2(\mathbb{N})$: alors, l'équation $\xi X - AX = Y$ entraîne que $\xi x_n - \mu_n x_n = y_n$ et comme $\xi \neq \mu_n$, nous obtenons que $x_n = y_n/(\xi - \mu_n)$. Par suite $R_\xi = (\xi I - A)^{-1} = B$, où $Y \mapsto BY$ est l'application t.q. $(BY)_n = y_n/(\xi - \mu_n)$. Ainsi ξ est une valeur régulière de $(A,\mathscr{D})$ si B est un endomorphisme borné de $\ell^2(\mathbb{N})$: mais cela est immédiat puisque, ξ n'étant pas adhérent à $\{\mu_n\}_n$, l'infimum r_0 des $|\xi - \mu_n|$ (pour $n \geq 0$) est strictement positif, ce qui assure que B est borné de norme $1/r_0$.

□

Une conséquence immédiate de la Proposition 2.3 est que pour toute partie F de $\mathbb{C}$ fermée et non vide, il existe une opérateur $(A,\mathscr{D})$ densément défini sur $\ell^2(\mathbb{N})$ t.q. $\Omega(A,\mathscr{D}) = \mathrm{Vp}'(A,\mathscr{D}) = F$ (considérer une suite $(\mu_n)_n$ de nombre complexes dense dans F).

2.3. Le graphe d'un opérateur $(A,\mathscr{D})$ densément défini sur un espace de Banach $\mathscr{E}$ est par définition $\mathrm{Gr}(A,\mathscr{D}) := \{(X,Y) \in \mathscr{D} \times \mathscr{E} \ ; \ Y = AX\}$. Nous dirons que $(A,\mathscr{D})$ est fermé lorsque son graphe $\mathrm{Gr}(A,\mathscr{D})$ est fermé dans $\mathscr{E} \times \mathscr{E}$; il sera dit « fermable » s'il est la restriction d'un opérateur $(A',\mathscr{D}')$ (i.e. $\mathscr{D} \subset \mathscr{D}'$ et $AX = A'X$, pour tout $X \in \mathscr{D}$) qui est lui-même fermé (on dira aussi que $(A,\mathscr{D})$ possède une extension $(A',\mathscr{D}')$ fermé). Nous notons simplement $\mathrm{Gr}(A)$ lorsque $A \in L(\mathscr{E})$ et dans ce cas, nous savons (*« théorème du graphe fermé »* : c.f. Appendice C) que A est borné ssi $\mathrm{Gr}(A)$ est fermé dans $\mathscr{E} \times \mathscr{E}$ (pour la topologie produit sur $\mathscr{E} \times \mathscr{E}$). Dans le cas où $\mathscr{D}$ est un sous-espace (dense et) strict de $\mathscr{E}$, alors $(A,\mathscr{D})$ est fermé ssi pour toute suite $X_0, X_1, \ldots$ de points de $\mathscr{D}$ convergeant (dans $\mathscr{E}$) vers X_*, alors la suite $AX_0, AX_1, \ldots$ converge vers Y dans $\mathscr{E}$ ssi $X_* \in \mathscr{D}$ et $AX_* = Y$. En un sens, pour les opérateurs densément définis, la notion de fermeture remplace la notion de continuité.

Proposition 2.4. *Soit $(A,\mathscr{D})$ un opérateur densément défini sur un espace de Banach $\mathscr{E}$: si la résolvante $\Omega(A,\mathscr{D})$ n'est pas vide alors $(A,\mathscr{D})$ est fermé.*

Preuve. Supposons que $(A,\mathscr{D})$ possède au moins une valeur régulière ξ : alors la résolvante $R_\xi := (A - \xi I)^{-1}$ est (par définition d'une valeur régulière) un élément de $\mathcal{L}(\mathscr{E})$ et donc son graphe $\mathrm{Gr}(R_\xi)$ est fermé. Il s'en suit immédiatement que $(A - \xi I, \mathscr{D})$ est fermé, puisque $\mathrm{Gr}(A - \xi I, \mathscr{D}) = \Phi^{-1}\big(\mathrm{Gr}(R_\xi)\big)$, où Φ est l'homéomorphisme de $\mathscr{E} \times \mathscr{E}$ t.q. $\Phi(X,Y) = (Y,X)$: enfin $(A,\mathscr{D})$ est fermé car $\mathrm{Gr}(A,\mathscr{D}) = \Psi^{-1}\big(\mathrm{Gr}(A - \xi I,\mathscr{D})\big)$ où Ψ est l'homéomorphisme de $\mathscr{E} \times \mathscr{E}$ t.q. $\Psi(X,Y) = (X, Y + \xi X)$.

□

En pratique, l'analyse spectrale d'un opérateur densément défini commence par la détermination (si elle existe !) de la fermeture de l'opérateur, c'est-à-dire la plus petite extension de l'opérateur qui est fermée.

Proposition 2.5. *Supposons que 0 soit une valeur régulière d'un opérateur $(A,\mathscr{D})$ densément défini et fermé sur $\mathscr{E}$, de sorte que la résolvante $R_0 = A^{-1}$ est un endomorphisme borné de $\mathscr{E}$; si ξ est une valeur régulière non nulle de $(A,\mathscr{D})$, alors $1/\xi$ est une valeur régulière de R_0 et de plus*

$$(1/\xi I - R_0)^{-1} = -\xi A R_\xi = \xi I - \xi^2 R_\xi$$

Preuve. Etant donné $\xi \in \Omega(A,\mathscr{D})$, il est facile de vérifier que $1/\xi I - R_0$ est un endomorphisme injectif de $\mathscr{E}$: en effet l'équation $1/\xi X - R_0 X = 0$ équivaut à $AX - \xi X = 0$, ce qui donne $X = 0$ (puisque $A - \xi I$ est bijectif par hypothèse). Afin de montrer que $1/\xi I - R_0$ est surjectif, soit $Y \in \mathscr{E}$ arbitrairement donné : nous utilisons alors le fait que $(\xi I - A)R_0$ est une bijection de $\mathscr{E}$, de sorte qu'il existe $Z \in \mathscr{E}$ tel que $Y = (\xi I - A)R_0 Z$ et par suite :

$$Y = (\xi I - A)(R_0 Z) = \xi R_0 Z - Z = 1/\xi(-\xi Z) - R_0(-\xi Z)$$

En d'autres termes, $Y = (1/\xi I - R_0)X$ pour $X = -\xi A R_\xi Y$. Ainsi $1/\xi I - R_0$ est un isomorphisme de $\mathscr{E}$ t.q. $(1/\xi I - R_0)^{-1} = -\xi A R_\xi$. Il nous reste à démontrer que AR_ξ est borné : mais partant de l'identité $(\xi I - A)R_\xi = I$ il vient $AR_\xi = \xi R_\xi - I$: or R_ξ est un endomorphisme borné de $\mathscr{E}$, car nous avons supposé que ξ est une valeur régulière de $(A, \mathscr{D})$.

□

3. **Opérateurs autoadjoints généralisés**

3.1. Soit $\mathscr{H}$ un $\mathbb{C}$-espace de Hilbert (toujours supposé séparable). Dans l'algèbre $\mathcal{L}(\mathscr{H})$ des endomorphismes bornés d'un espace de Hilbert $\mathscr{H}$, le passage à l'adjoint est l'analogue de la conjugaison des nombres complexes : cela fait des opérateurs autoadjoints et unitaires (c.f. § 3.2 infra) les analogues respectifs des nombres réels et des nombres complexes de module 1. (La notion de C*-algèbres – introduite par Von Neumann et se successeurs – est construite comme une extension axiomatique de l'adjoint dans le cadre des espaces de Hilbert.)

Proposition 3.1. *Pour tout $A \in \mathcal{L}(\mathscr{H})$ il existe un unique $A^\star \in \mathcal{L}(\mathscr{H})$ appelé l'adjoint de A t.q. $\langle AX|Y\rangle = \langle X|A^\star Y\rangle$, pour tout $X, Y \in \mathscr{H}$; de plus nous avons*

(6) $$(i) \ : \ \|A\| = \|A^\star\| \qquad \text{et} \qquad (ii) \ : \ \|A^\star A\| = \|A\|^2$$

Lemme 3.2. $\forall A \in L(\mathscr{H})$, $\sup\left\{\|AX\| \ ; \ \|X\| = 1\right\} = \sup\{\langle AX|Y\rangle \ ; \ \|X\| = \|Y\| = 1\}$.

Preuve. Pour $X \in \mathscr{H}$ fixé, l'inégalité de Cauchy-Schwarz nous donne

$$\sup\left\{|\langle AX|Y\rangle| \ ; \ \|Y\| = 1\right\} \leq \|AX\|$$

en posant $Y = AX/\|AX\|$ (et sous réserve que $AX \neq 0$), nous obtenons alors

$$\sup\left\{|\langle AX|Y\rangle| \ ; \ \|Y\| = 1\right\} = \|AX\|$$

La concusion vient du fait que cette dernière identité est aussi évidente lorsque $AX = 0$.

□

Preuve de la Proposition 3.1. Etant donné $Y \in \mathscr{H}$ l'application $X \mapsto \langle AX|Y\rangle$ est une forme linéaire sur $\mathscr{H}$. Il existe donc $BY \in \mathscr{H}$ t.q. $\langle AX|Y\rangle = \langle X|BY\rangle$ et il est facile de vérifier que $B \in L(\mathscr{H})$. D'une part, il est immédiat que B est le seul endomorphisme à vérifier l'identité $\langle AX|Y\rangle = \langle X|BY\rangle$, pour tout $X, Y \in \mathscr{H}$; d'autre part, le fait que $A^\star := B$ est borné sur $\mathscr{H}$ avec $\|A^\star\| = \|A\|$, découle directement du Lemme 3.2[7]. Enfin, pour démontrer l'identité (ii) de (6), notons que la sous-multiplicativité de la norme

7. Deuxième preuve de l'égalité $\|A\| = \|A^\star\|$: par définition de la norme opérateur, il existe une suite $X_1, X_2, \ldots$ dans la sphère unité $\{\|\cdot\| = 1\}$ de $\mathscr{H}$ t.q. $\lim_n \|AX_n\| = \|A\|$: ici, nous utilisons le fait que la boule fermée $\{\|\cdot\| \leq 1\}$ of $\mathscr{H}$ est faiblement compacte, ce qui assure l'existence d'une sous-suite $X_{n_1}, X_{n_2}, \ldots$ qui tend faiblement vers une limite X_* (avec $\|X_*\| = 1$) : par suite, $\|A\| = \lim_k \|AX_{n_k}\| = \|AX_*\|$ ($X \mapsto \|AX\|$ est fortement donc faiblement continue) et donc $\|A\|^2 = |\langle AX_*|AX_*\rangle| = |\langle X_*|A^\star AX_*\rangle| \leq \|A^\star\|\|A\|$; ainsi $\|A\| \leq \|A^\star\|$ et finalement $\|A^\star\| = \|A\|$, du fait que $(A^\star)^\star = A$.

opérateur assure toujours que $\|A^\star A\| \le \|A^\star\| \cdot \|A\|$; la conclusion vient en utilisant l'égalité $\|A\| = \|A^\star\|$, de sorte que (nouvelle application du Lemme 3.2) :

$$\begin{aligned} \|A^\star A\| &= \sup\Big\{ |\langle AX|AY\rangle| \,;\, \|X\| = \|Y\| = 1 \Big\} \\ &\ge \sup\Big\{ |\langle AX|AX\rangle| \,;\, \|X\| = 1 \Big\} = \|A\|^2 = \|A^\star\| \cdot \|A\|. \end{aligned}$$

□

Corollaire 3.3. *L'application $A \mapsto A^\star$ est une involution isométrique de $\mathcal{L}(\mathscr{H})$, en ce sens que $(A^\star)^\star = A$ avec $\|A\| = \|A^\star\|$: nous appelons cette application « la $\star$-involution de $\mathcal{L}(\mathscr{H})$ ».*

Preuve. Nous savons déjà (c.f. Proposition 3.1) que $\|A\| = \|A^\star\|$, ce qui signifie que $A \mapsto A^\star$ est un isométrie de l'espaces de Banach $(\mathcal{L}(\mathscr{H}), \|\cdot\|)$. Pour vérifier que $A \mapsto A^\star$ est une involution, notons que l'opérateur $B = (A^\star)^\star$ est caractérisé par le fait que $\langle A^\star X|Y\rangle = \langle X|BY\rangle$, pour tout $X, Y \in \mathscr{H}$: comme $\langle A^\star X|Y\rangle = \langle X|AY\rangle$, il vient $\langle X|(A-B)Y\rangle = 0$: cette identité étant valide pour tout $X, Y \in \mathscr{H}$, cela assure que $A = B$.

□

3.2. La $\star$-involution $A \mapsto A^\star$ du passage à l'adjoint dans $\mathcal{L}(\mathscr{H})$ nous amène à définir plusieurs sous-classes très importantes d'opérateurs. Nous dirons d'abord que A est *« autoadjoint »* si $A^\star = A$. Pour $A \in \mathcal{L}(\mathscr{H})$ quelconque, les deux opérateurs $\mathfrak{re}(A) := (A + A^\star)/2$ et $\mathfrak{im}(A) := \boldsymbol{i}(A^\star - A)/2$ sont tous deux autoadjoints et de plus $A = \mathfrak{re}(A) + \boldsymbol{i}\,\mathfrak{im}(A)$. Nous dirons que A est *« positif »* si $\langle AX|X\rangle \ge 0$, pour tout $X \in \mathscr{H}$; comme $\langle AX|X\rangle = \overline{\langle AX|X\rangle} = \langle X|AX\rangle$, nous déduisons que $\langle (A - A^\star)X|X\rangle = 0$, ce qui n'est possible (pour tout $X \in \mathscr{H}$) que si $A = A^\star$: par suite, un opérateur positif est nécessairement autoadjoint.

Lorsque $\langle AX|AY\rangle = \langle X|Y\rangle$, pour tout $X, Y \in \mathscr{H}$, l'opérateur A est appelé une *« isométrie »* : une isométrie d'un espace de Hilbert de dimension infinie est injective, mais pas nécessairement surjective[8] : si l'opérateur isométrique A est surjectif (et donc bijectif) il est dit *« unitaire »*. A l'instar d'un nombre complexe de module 1, un opérateur U est unitaire ssi il est bijectif et si $U^{-1} = U^\star$; il est aussi facile de montrer que U est unitaire ssi pour toute base hilbertienne $\{E_i\}_i$ de $\mathscr{H}$ la famille $\{UE_i\}_i$ est aussi une base hilbertienne. Pour conclure ce paragraphe notons que si $A \in \mathcal{L}(\mathscr{H})$ est un autoadjoint, alors :

$$\exp(\boldsymbol{i}A)^\star = \sum_{k=0}^{+\infty} \frac{1}{k!}(\boldsymbol{i}A)^\star = \sum_{k=0}^{+\infty} -\frac{\boldsymbol{i}}{k!}A = \exp(-\boldsymbol{i}A) = \exp(\boldsymbol{i}A)^{-1}$$

Nous pouvons ainsi affirmer que

Proposition 3.4. *Si $A \in \mathcal{L}(\mathscr{H})$ est autoadjoint, alors $\exp(\boldsymbol{i}A)$ est unitaire.*

---

8. Si $U : \mathscr{H} \to \mathscr{H}$ est une isométrie, alors pour tout $X, Y \in \mathscr{H}$ nous avons$\langle X|Y\rangle = \langle UX|UY\rangle = \langle X|U^\star UY\rangle$: cela implique que $U^\star U = I$ (l'opérateur identité) : cependant, quand $\mathscr{H}$ est de dimension infinie, cela n'implique pas que $U^\star$ soit inversible. Pour ce faire une idée on peut par exemple imaginer la version en dimension infinie des identités matricielles d'ordres finies telles que

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

On est alors amené à considérer l'isométrie non surjective $U : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ t.q. $U(x_0, x_1, \dots) = (0, x_0, x_1, \dots)$ pour laquelle nous avons en effet $U^\star U = I$.

3.3. La formalisation de la mécanique quantique nécessite d'étendre la notion d'endomorphisme autoadjoint d'un espace de Hilbert $\mathscr{H}$ au cas des opérateurs densément définis et possiblement non bornés (c.f. Définition 3.11 infra). Il ne s'agit pas d'une généralisation nécessaire afin de traiter des situations physiques marginales ou exceptionnelles, mais bien pour traiter les cas les plus élémentaires. Pour voir cela, nous devons d'abord étendre la définition classique de l'adjoint d'un opérateur donné ci-dessus au cas où l'opérateur est densément défini et non borné. Considérons donc que $(A, \mathscr{D})$ est densément défini sur $\mathscr{H}$ et définissons $\mathscr{D}^\star$ comme le sous-espace de $\mathscr{H}$ formé des $Y \in \mathscr{H}$ pour lesquels la forme linéaire $\varphi_Y : X \mapsto \langle Y|AX\rangle$ est continue sur $\mathscr{D}$ (munie de la norme hilbertienne). Pour tout $Y \in \mathscr{D}^\star$, la forme linéaire φ_Y étant continue sur $\mathscr{D}$ elle se prolonge en une forme linéaire continue sur tout $\mathscr{H}$ (toujours noté φ_Y) : par suite il existe un $Y^\star \in \mathscr{H}$ t.q. $\varphi_Y(X) = \langle Y^\star|X\rangle$, pour tout $X \in \mathscr{H}$: l'application $A^\star : Y \mapsto Y^\star$ est définie et linéaire sur $\mathscr{D}^\star$ (par définition de $\mathscr{D}^\star$) et de plus elle vérifie

$$(\forall X \in \mathscr{D},\ \forall Y \in \mathscr{D}^\star) \qquad \langle Y|AX\rangle = \langle A^\star Y|X\rangle \tag{7}$$

L'adjoint de $(A, \mathscr{D})$ est par définition, l'opérateur $(A^\star, \mathscr{D}^\star)$ (qui n'est pas nécessairement densément défini). Notons au passage que $Y \in \operatorname{Ker}(A^\star, \mathscr{D}^\star)$ si et seulement si $0 = \langle 0, X\rangle = \langle Y, AX\rangle$ pour tout $X \in \mathscr{D}$: cela signifie que $Y \in \operatorname{Im}(A, \mathscr{D})^\perp$.

Proposition 3.5. *Si $(A, \mathscr{D})$ un opérateur densément défini sur $\mathscr{H}$, alors*

$$\operatorname{Ker}(A^\star, \mathscr{D}^\star) = \operatorname{Im}(A, \mathscr{D})^\perp$$

3.4. Nous allons voir que pour tout opérateur $(A, \mathscr{D})$ densément défini sur $\mathscr{H}$ (c.f. Proposition 3.7 infra) que le graphe de l'adjoint $(A^\star, \mathscr{D}^\star)$ est fermé dans l'espace produit $\mathscr{H} \times \mathscr{H}$. Une manière simple d'aborder la question est d'utiliser le fait que l'orthogonal d'une partie quelconque d'un espace de Hilbert est un sous-espace fermé de cet espace ; cependant ici, l'espace de Hilbert à considérer est le produit cartésien $\mathscr{H} \times \mathscr{H}$ muni de sa structure naturelle d'espace vectoriel et du produit scalaire

$$\langle\!\langle (X, Y)|(U, V)\rangle\!\rangle = \langle X|U\rangle + \langle Y|V\rangle \tag{8}$$

Lemme 3.6. *L'application $(X, Y) \mapsto \Phi(X, Y) = (-Y, X)$ est un endomorphisme unitaire de l'espace de Hilbert $\mathscr{H} \times \mathscr{H}$ muni du produit scalaire $\langle\!\langle \cdot|\cdot \rangle\!\rangle$ défini en (8).*

Preuve. Il est immédiat que Φ est un isomorphisme borné de $\mathscr{H} \times \mathscr{H}$ dont l'inverse est $\Phi^{-1} = -\Phi$; l'adjoint de Φ se calcule simplement en écrivant pour $X, Y, U, V \in \mathscr{H}$:

$$\begin{aligned}
\langle\!\langle (X, Y)|\Phi(U, V)\rangle\!\rangle &= \langle\!\langle (X, Y)|(-V, U)\rangle\!\rangle \\
&= \langle -X, V\rangle + \langle Y|U\rangle = \langle Y|U\rangle + \langle -X, V\rangle = \langle\!\langle (Y, -X)|(U, V)\rangle\!\rangle
\end{aligned}$$

de sorte que $\Phi^\star(X, Y) = (Y, -X) = -\Phi(X, Y)$: Φ est unitaire puisque $\Phi^\star = \Phi^{-1}$.

□

Proposition 3.7. *Si $(A, \mathscr{H})$ un opérateur densément défini sur $\mathscr{H}$, alors*

$$\operatorname{Gr}(A^\star, \mathscr{D}^\star) = \Phi\big(\operatorname{Gr}(A, \mathscr{D})\big)^\perp \tag{9}$$

En particulier le graphe de $(A^\star, \mathscr{D}^\star)$ est fermé.

Preuve. Le couple (Y,X) appartient à $\mathrm{Gr}(A^\star,\mathscr{D}^\star)$ si et seulement si

$$(\forall Z \in \mathscr{D}) \qquad 0 = \langle Y| - AZ\rangle + \langle X|Z\rangle = \langle\!\langle (Y,X)|(-AZ,Z)\rangle\!\rangle$$

ce qui signifie que (Y,X) est orthogonal $\Phi\big(\mathrm{Gr}(A,\mathscr{D})\big)$.

□

Remarque 3.8. *D'après le Lemme 3.6, Φ étant un endomorphisme unitaire de $\mathscr{H}\times\mathscr{H}$, il découle de (9) que* $\mathrm{Gr}(A^\star,\mathscr{D}^\star) = \Phi\big(\mathrm{Gr}(A,\mathscr{D})\big)^\perp = \Phi\big(\mathrm{Gr}(A,\mathscr{D})^\perp\big)$ *; mais comme $\Phi^{-1} = -\Phi$, il vient :*

$$\Phi\big(\mathrm{Gr}(A^\star,\mathscr{D}^\star)\big) = \mathrm{Gr}(A,\mathscr{D})^\perp \tag{10}$$

Proposition 3.9. *Si $(A,\mathscr{D})$ est densément défini sur $\mathscr{H}$ et fermé, alors (i) : $(A^\star,\mathscr{D}^\star)$ est densément défini ; (ii) : $(A^{\star\star},\mathscr{D}^{\star\star}) = (A,\mathscr{D})$ et (iii) : si Φ est l'isomorphisme de $\mathscr{H}\times\mathscr{H}$ t.q. $\Phi(X,Y) = (-Y,X)$, alors nous avons la somme directe orthogonale*

$$\mathscr{H}\times\mathscr{H} = \mathrm{Gr}(A,\mathscr{D}) \boxplus \Phi(\mathrm{Gr}(A^\star,\mathscr{D}^\star)) \tag{11}$$

Preuve. (i) : Pour $(A,\mathscr{D})$ densément défini et fermé sur $\mathscr{H}$, nous allons montrer que si $Y\in\mathscr{H}$ est orthogonal à $\mathscr{D}^\star$, alors $Y = 0$. En effet, si Y est orthogonal à $\mathscr{D}^\star$), alors $(Y,0)$ est orthogonal à $\mathrm{Gr}(A^\star,\mathscr{D}^\star)$ (dans $\mathscr{H}\times\mathscr{H}$) : par suite $(0,Y) = \Phi(Y,0)$ est orthogonal à $\Phi(\mathrm{Gr}(A^\star,\mathscr{D}^\star)) = \mathrm{Gr}(A,\mathscr{D})^\perp$. D'autre part, $(0,Y)\in\mathrm{Gr}(A,\mathscr{D})^{\perp\perp}$ puisque $(A,\mathscr{D})$ est fermé. Enfin $\mathrm{Gr}(A,\mathscr{D})$ est un sous-espace fermé, donc $(\mathrm{Gr}(A,\mathscr{D}))^{\perp\perp} = \mathrm{Gr}(A,\mathscr{D})$: par suite, $(0,Y)\in\mathrm{Gr}(A,\mathscr{D})$ et donc $Y = A0 = 0$.

(ii)–(iii) : Comme $(A,\mathscr{D})$ est densément défini et fermé sur $\mathscr{H}$, nous savons (c.f. Proposition 3.9) que $\mathscr{D}$ est une sous-espace dense de $\mathscr{H}$: ceci justifie que $(A,\mathscr{D})$ est un opérateur densément défini sur $\mathscr{H}$. Mais en introduisant l'endomorphisme unitaire Φ de $\mathscr{H}\times\mathscr{H}$ t.q. $\Phi(X,Y) = (-Y,X)$ nous savons (c.f. Proposition 3.7 et Remarque 3.8) que $\Phi(\mathrm{Gr}(A^\star,\mathscr{D}^\star)) = \mathrm{Gr}(A,\mathscr{D})^\perp$; par suite, si $\mathrm{Gr}(A,\mathscr{D})$ est fermé alors

$$\mathrm{Gr}(A,\mathscr{D}) = \Phi(\mathrm{Gr}(A^\star,\mathscr{D}^\star))^\perp \tag{12}$$

Mais nous savons d'autre part d'après (i) nous savons que $(A^\star,\mathscr{D}^\star)$ est densément défini : par une nouvelle application de la Proposition 3.7, il vient donc $\mathrm{Gr}((A^{\star\star},\mathscr{D}^{\star\star})) = \Phi(\mathrm{Gr}(A^\star,\mathscr{D}^\star))^\perp$, de sorte que par (12) nous obtenons

$$\mathrm{Gr}((A^{\star\star},\mathscr{D}^{\star\star})) = \mathrm{Gr}(A,\mathscr{D})$$

cela signifie que $(A^{\star\star},\mathscr{D}^{\star\star}) = (A,\mathscr{D})$. La somme directe orthogonale en (11) découle donc du fait que dans un espace de Hilbert le supplémentaire orthogonal d'un sous-espace fermé coïncide avec son orthogonal.

□

L'opérateur $(A,\mathscr{D})$ (densément défini sur $\mathscr{H}$) est dit « *symétrique* » si

$$(\forall X,Y\in\mathscr{D}) \qquad \langle AX|Y\rangle = \langle X|AY\rangle \tag{13}$$

Dans ce cas, pour $Y\in\mathscr{D}$ donné, l'application $X\mapsto\langle Y|AX\rangle$ est une forme linéaire continue sur $\mathscr{D}$ (en effet par symétrie $\langle Y|AX\rangle = \langle AY|X\rangle \le \|AY\|\cdot\|X\|$). Par suite, $\mathscr{D}^\star$ étant défini comme le sous-espace des $Y\in\mathscr{H}$ t.q. $\varphi_Y : X\mapsto\langle Y|AX\rangle$ soit continue sur $\mathscr{D}$, nous avons l'inclusion $\mathscr{D}^\star\supset\mathscr{D}$.

Proposition 3.10. *Si l'opérateur $(A, \mathscr{D})$ densément défini sur $\mathscr{H}$ est symétrique, alors $\mathscr{D} \subset \mathscr{D}^\star$ et de plus $AY = A^\star Y$, pour tout $Y \in \mathscr{D}$.*

Remarquons que lorsque $(A, \mathscr{D})$ est symétrique, la densité de $\mathscr{D}$ dans $\mathscr{H}$ entraîne que l'adjoint $(A^\star, \mathscr{D}^\star)$ est densément défini sur[9] $\mathscr{H}$. Nous posons la définition suivante.

Définition 3.11. *Un opérateur $(A, \mathscr{D})$ densément défini sur un espace de Hilbert $\mathscr{H}$ est dit autoadjoint s'il est égal à son adjoint, i.e. si $(A, \mathscr{D}) = (A^\star, \mathscr{D}^\star)$.*

Proposition 3.12. *Un opérateur densément défini est autoadjoint ssi il est symétrique et fermé.*

Preuve. D'après la Proposition 3.7, si $(A, \mathscr{D})$ est autoadjoint alors il est symétrique et fermé. Réciproquement si $(A, \mathscr{D})$ est symétrique alors nous savons (c.f. Proposition 3.10) que $\mathscr{D} \subset \mathscr{D}^\star$, avec $A^\star Y = AY$, pour tout $Y \in \mathscr{D}$; si de plus $(A, \mathscr{D})$ est fermé alors (Proposition 3.9-(ii)) $(A^{\star\star}, \mathscr{D}^{\star\star}) = (A, \mathscr{D})$ d'où $\mathscr{D}^\star \subset (\mathscr{D}^\star)^\star = \mathscr{D}$ et donc $\mathscr{D}^\star = \mathscr{D}$: en conclusion $(A^\star, \mathscr{D}^\star) = (A, \mathscr{D})$ du fait que $A^\star Y = AY$, pour tout $Y \in \mathscr{D}$ et que d'autre part $\mathscr{D}^\star = \mathscr{D}$.

□

Proposition 3.13. *Soit $(A, \mathscr{D})$ un opérateur autoadjoint sur $\mathscr{H}$ t.q. $0 \in \Omega(A, \mathscr{D})$; alors A est un élément inversible de $L(\mathscr{D}, \mathscr{H})$ dont l'inverse A^{-1} est un endomorphisme autoadjoint de $\mathcal{L}(\mathscr{H})$.*

Preuve. Le fait que $A^{-1} \in \mathcal{L}(\mathscr{H})$ découle directement (par définition) de l'hypothèse $0 \in \Omega(A, \mathscr{D})$. Pour vérifier que A^{-1} est autoadjoint considérons $X = AX'$ et $Y = AY'$ arbitrairement pris dans $\mathscr{H}$; alors $\left\langle A^{-1}X | Y \right\rangle = \langle X' | AY' \rangle = \langle AX' | Y' \rangle = \left\langle X | A^{-1}Y \right\rangle$.

□

Nous déduisons la proposition suivante comme corollaire de la Proposition 3.5.

Proposition 3.14. *Soit $(A, \mathscr{D})$ un opérateur autoadjoint sur $\mathscr{H}$ et pour lequel il existe $C > 0$ t.q. $\|AX\| \geq C\|X\|$, pour tout $X \in \mathscr{D}$; alors A est une application inversible de $L(\mathscr{D}, \mathscr{H})$ dont l'inverse A^{-1} est un endomorphisme autoadjoint de $\mathcal{L}(\mathscr{H})$ t.q. $\left\|A^{-1}\right\| \leq 1/C$.*

Preuve. Le fait que $\|AX\| \geq C\|X\|$, pour tout $X \in \mathscr{D}$, assure que A est une application injective de $L(\mathscr{D}, \mathscr{H})$. En combinant la Proposition 3.5 et le fait que $(A, \mathscr{D})$ est autoadjoint il vient : $\{0\} = \mathrm{Ker}(A, \mathscr{D}) = \mathrm{Im}(A^\star, \mathscr{D}^\star)^\perp = \mathrm{Im}(A, \mathscr{D})^\perp$, ce qui signifie que l'image de $\mathscr{G} := \mathrm{Im}(A, \mathscr{D})$ est dense dans $\mathscr{H}$: il nous reste à démontrer que $\mathscr{G} = \mathscr{H}$. Pour cela nous remarquons est une application bijective de $L(\mathscr{G}, \mathscr{H})$; donc, pour tout $Y \in \mathscr{G}$, nous avons (par hypothèse) $\|Y\| = \|AA^{-1}Y| \geq C\|A^{-1}Y\|$ de sorte que $\|A^{-1}Y\| \leq 1/C\|Y\|$; mais comme $\mathscr{G}$ est dense dans $\mathscr{H}$, nous savons (théorème de prolongement par densité des applications linéaires) qu'il existe $B \in \mathcal{L}(\mathscr{H})$ (avec $\|B\| \leq 1/C$) telle que $A^{-1}Y = BY$, pour tout $Y \in \mathscr{G}$. Soit Y est un point arbitraire de $\mathscr{H}$: la proposition sera établie si nous démontrons que $Y \in \mathscr{G}$ (puisque nous aurions alors $A^{-1} = B$). Mais en effet (densité de $\mathscr{G}$ dans $\mathscr{H}$), il existe une suite $Y_1, Y_2, \ldots$ de point de $\mathscr{G}$ t.q. $Y_n \to Y$ quand $n \to +\infty$. Mais alors, la continuité de B assure que $A^{-1}Y_n = BY_n \to BY$ quand $n \to +\infty$: les couples $(BY_n, Y_n) = (A^{-1}Y_n, Y_n)$ forment ainsi une suite dans $\mathrm{Gr}(A, \mathscr{D})$ qui converge vers (BY, Y) (dans $\mathscr{H} \times \mathscr{H}$) : nous pouvons alors conclure que $(BY, Y) \in \mathrm{Gr}(A, \mathscr{D})$ du fait que le graphe de $(A, \mathscr{D})$ est fermé dans $\mathscr{H} \times \mathscr{H}$, ce qui prouve que $Y \in \mathscr{G}$.

□

9. Cela ne découle pas de la Proposition 3.9 car n'est pas supposé fermé (nous verrons que si $(A, \mathscr{D})$ est symétrique et fermé, alors $(A, \mathscr{D}) = (A^\star, \mathscr{D}^\star)$).

3.5. Les opérateurs densément définis et autoadjoints sur un espace de Hilbert jouent un rôle particulièrement important dans la description quantique du monde physique : cette formalisation s'est produite à la fin des années 20 et au début des années 30 à partir du travail de Stone et Von Neumann sur la généralisation du *« théorème spectral »* tel qu'énoncé par Hilbert (c.f. [Sto30, Sto32] et le livre de synthèse [VN55]). Nous allons voir quelques propriétés spectrales de ces opérateurs qui généralisent celles des opérateurs autoadjoints bornés (c.f. Appendice A).

Proposition 3.15. *Si $(A, \mathscr{D})$ est autoadjoint alors son spectre $\Sigma(A, \mathscr{D})$ est non vide.*

Preuve. Supposons par l'absurde que $\Omega(A, \mathscr{D}) = \mathbb{C}$; alors en particulier $0 \in \Omega(A, \mathscr{D})$ ce qui assure (c.f. Proposition 3.13) que A est un élément inversible de $L(\mathscr{D}, \mathscr{H})$ dont l'inverse A^{-1} est un endomorphisme autoadjoint de $\mathcal{L}(\mathscr{H})$. Nous allons voir qu'il est alors nécessaire que $A^{-1} = 0$, ce qui tiendra lieu de contradiction. Pour cela, nous utilisons le fait que les opérateurs autoadjoints bornés sont spectraloïdes (c.f. Théorème 8.3 in Appendice A) : considérons donc que λ est une valeur spectrale non nulle de A^{-1} ; alors λ est un réel non nul et du fait que $1/\lambda \in \Omega(A, \mathscr{D})$, nous savons que $A - 1/\lambda I$ est un élément inversible de $L(\mathscr{D}, \mathscr{H})$ dont l'inverse $(A - 1/\lambda I)^{-1}$ est dans $\mathcal{L}(\mathscr{H})$. Nous pouvons alors vérifier directement que $A^{-1} - \lambda I$ est un endomorphisme inversible de $\mathcal{L}(\mathscr{H})$ d'inverse $1/\lambda A(1/\lambda I - A)^{-1} = A(I - \lambda A)^{-1}$: en effet, d'une part

$$(A^{-1} - \lambda I)\Big(A(I - \lambda A)^{-1}\Big) = (I - \lambda A)(I - \lambda A)^{-1} = I$$

et d'autre part (en utilisant la commutation $A \leftrightarrow (1/\lambda I - A)^{-1}$= c.f. Proposition 2.1)

$$\Big(A(I - \lambda A)^{-1}\Big)(A^{-1} - \lambda I) = \Big((I - \lambda A)^{-1} A\Big)(A^{-1} - \lambda I) = (I - \lambda A)^{-1}(I - \lambda A) = I$$

Par suite, A^{-1} est un endomorphisme de $\mathcal{L}(\mathscr{H})$ qui est autoadjoint et de spectre $\Sigma(A^{-1}) \subset \{0\}$; la norme de l'endomorphisme (spectraloïde) A^{-1} coïncidant avec son rayon spectral, nous en déduisons que A^{-1} est nécessairement l'endomorphisme nul de $\mathscr{H}$: c'est impossible du fait que $A^{-1}AX = X$ pour tout $X \in \mathscr{D}$.

□

Le théorème suivant étend aux opérateurs autoadjoints non bornés la version classique pour les opérateurs autoadjoints bornés (nous en donnons une démonstration qui n'utilise pas le théorème spectral où le critère de Weyl : c.f. Remarque 3.18 infra).

Théorème 3.16. *Le spectre d'un opérateur densément défini autoadjoint (et donc fermé) d'un espace de Hilbert (séparable) est un sous-ensemble non vide de $\mathbb{R}$.*

Preuve. Soit $(A, \mathscr{G})$ opérateur densément défini autoadjoint (et donc fermé) d'un espace de Hilbert (séparable) $\mathscr{H}$ et soit $\xi = \alpha + i\beta$, avec $\alpha, \beta \in \mathbb{R}$ et $\beta \neq 0$. Nous voulons montrer que $(A - \xi I, \mathscr{G})$ est une application bijective de $L(\mathscr{G}, \mathscr{H})$ dont l'inverse est un endomorphisme borné de $\mathscr{H}$. Pour voir cela, notons d'abord que pour tout $X \in \mathscr{G}$

$$\|(A - \xi I)X\|^2 = \|(A - \alpha I)X\|^2 + \beta^2\|X\|^2 \tag{14}$$

Comme $\beta \neq 0$ l'équation (14) assure que $(A - \xi I, \mathscr{G})$ est injective. Pour conclure que $(A - \xi I, \mathscr{G})$ est une application bijective de $L(\mathscr{G}, \mathscr{H})$, nous devons montrer qu'elle est surjective en s'assurant que son image $(A - \xi I)(\mathscr{G})$ est un sous-espace fermé et dense de $\mathscr{H}$.

Commençons par vérifier que $(A-\xi I)(\mathscr{G})$ est un fermé. Soit $X_0, X_1, \ldots$ une suite de $\mathscr{G}$ et supposons que les $Y_n = (A-\xi I)X_n$ tendent vers $Y_* \in \mathscr{H}$ quand $n \to +\infty$. Alors d'après (14) les X_n convergent nécessairement vers un point X_* de $\mathscr{H}$ et donc les (X_n, Y_n) forment une suite de $\mathrm{Gr}(\xi I - A, \mathscr{G})$ qui converge dans $\mathscr{H} \times \mathscr{H}$. Mais $(A, \mathscr{G})$ étant autoadjoint, il est en particulier fermé ; or $\mathrm{Gr}(A-\xi I, \mathscr{G}) = \Phi^{-1}\big(\mathrm{Gr}(A, \mathscr{G})\big)$, où Φ est l'homéomorphisme de $\mathscr{H} \times \mathscr{H}$ t.q. $\Phi(X, Y) = (X, Y + \xi X)$: cela assure que $\mathrm{Gr}(A-\xi I, \mathscr{G})$ est fermé dans $\mathscr{H} \times \mathscr{H}$ et par suite, $X_* \in \mathscr{G}$ avec $Y_* = (A-\xi I)X_*$. Nous avons montré que $(A-\xi I)(\mathscr{G})$ est fermé dans $\mathscr{H}$. Pour établir la bijectivité de $(A-\xi I, \mathscr{G})$, il nous reste donc à vérifier que $(A-\xi I)(\mathscr{G})$ est dense dans $\mathscr{H}$. Si $Y \in (A-\xi I)(\mathscr{G})^\perp$, alors $\langle Y|AX\rangle = \xi\langle Y|X\rangle$, pour tout $X \in \mathscr{G}$: nous en déduisons que $X \mapsto \langle Y|AX\rangle$ est bornée sur $\mathscr{G}$, de sorte que $Y \in \mathscr{G}^\star = \mathscr{G}$ avec $A^\star Y = AY = \xi Y$. Or $(A-\xi I, \mathscr{G})$ étant injectif, l'identité $AY = \xi Y$ entraîne que $Y = 0$: ainsi l'orthogonal $(A-\xi I)(\mathscr{G})^\perp$ est réduit à $\{0\}$, ce qui fait de $(A-\xi I)(\mathscr{G})$ un sous-espace dense de $\mathscr{H}$. Nous venons de montrer $(A-\xi I, \mathscr{G})$ est un application bijective de $L(\mathscr{G}, \mathscr{H})$. Ainsi, $(A-\xi I)^{-1}$ peut-être considéré comme un endomorphisme de $\mathscr{H}$: pour conclure que $\xi \in \Omega(A, \mathscr{G})$ nous devons vérifier que $(A-\xi I)^{-1}$ est borné. Or ceci découle du théorème du graphe fermé : en effet $\mathrm{GR}((A-\xi I)^{-1}) = \Psi(\mathrm{GR}(A-\xi I, \mathscr{G})$ où Ψ est l'homéomorphisme de $\mathscr{H} \times \mathscr{H}$ t.q. $\Psi(X, Y) = (Y, X)$: mais nous avons déjà vu que $\mathrm{GR}(A-\xi I, \mathscr{G})$ est un fermé de $\mathscr{H} \times \mathscr{H}$.

□

L'opérateurs $(A, \mathscr{D})$ est dit *« positif »* si $\langle AX|X\rangle \geq 0$, pour tout $X \in \mathscr{D}$. Dans les applications (c.f. § 6 à propos du laplacien) il est utile de remarquer que si $(A, \mathscr{D})$ est un opérateur densément défini sur $\mathscr{H}$ symétrique, positif et fermable, alors sa fermeture est un opérateur autoadjoint positif. La proposition suivante complète le Théorème 3.16.

Proposition 3.17. *Si $(A, \mathscr{D})$ est autoadjoint positif sur $\mathscr{H}$, alors $\Sigma(A, \mathscr{D}) \subset [0\,; +\infty[$.*

Preuve. Comme $(A, \mathscr{D})$ est autoadjoint, nous savons (c.f. Théorème 3.16) que $\Sigma(A, \mathscr{D}) \subset \mathbb{R}$. Afin de montrer que $\Sigma(A, \mathscr{D}) \subset [0\,; +\infty[$, soit $\lambda < 0$ et vérifions $\lambda \in \Omega(A, \mathscr{D})$ en ce sens que $A - \lambda I$ est inversible et que son inverse est dans $\mathcal{L}(\mathscr{H})$: pour cela, nous notons que la positivité de $(A, \mathscr{D})$ assure la positivité de $(A-\lambda I, \mathscr{D})$, de sorte que pour tout $X \in \mathscr{D}$: $|\lambda|\|X\|^2 \leq \langle (A-\lambda I)X|X\rangle$, soit encore (d'après l'inégalité de Cauchy Schwarz) :

$$\|(A-\lambda I)X\| \geq |\lambda|\|X\| \tag{15}$$

Comme λ est réel, $(A-\lambda I, \mathscr{D})$ est un opérateur autoadjoint inversible de $L(\mathscr{D}, \mathscr{H})$: par suite, d'après (83) et par application de la Proposition 3.14, son inverse est un endomorphisme borné de $\mathscr{H}$.

□

Remarque 3.18. *Rappelons que « le critère de Weyl » affirme que « si $(A, \mathscr{D})$ est autoadjoint sur $\mathscr{H}$ et si λ est une valeur spectrale, alors c'est nécessairement une valeurs propres approchées : en d'autre termes $\Sigma(A, \mathscr{D}) = \mathrm{Vp}'(A, \mathscr{D})$ ». On peut trouver une démonstration de ce théorème dans* [RSN55, § 133, p. 360] *à partir de la notion de « résolution de l'identité » ; dans ce contexte (que nous n'abordons pas ici) il est possible de définir le spectre discret de $(A, \mathscr{D})$ noté $\Sigma_{\mathrm{d}}(A, \mathscr{D})$ et de prouver que cet ensemble est constitué d'une part des points isolés du spectre $\Sigma(A, \mathscr{D})$ et d'autre part des valeurs propres de multiplicité finie. Le spectre essentiel $\Sigma_{\mathrm{e}}(A, \mathscr{D})$, est alors défini comme* *le complémentaire de $\Sigma_{\rm d}(A,\mathscr{D})$ dans $\Sigma(A,\mathscr{D})$, de sorte que $\Sigma_{\rm e}(A,\mathscr{D}) = \Sigma(A,\mathscr{D}) \smallsetminus \Sigma_{\rm d}(A,\mathscr{D})$. On montre que $\Sigma_{\rm e}(A,\mathscr{D})$ est fermé dans $\mathbb{R}$, une valeur spectrale essentielle étant soit un élément du spectre continu, soit un point limite d'une suite de valeurs propres de multiplicité finie soit encore une valeur propre de multiplicité infinie. Notons enfin que l'énoncé général du critère de Weyl affirme que $\lambda \in \Sigma_{\rm e}(A,\mathscr{D})$ ssi il existe une suite $X_1, X_2, \dots$ de la sphère unité de $\mathscr{D}$ qui converge faiblement [10] dans $\mathscr{H}$ et t.q. $\|AX_n - \lambda X_n\| \to 0$ quand $n \to +\infty$; en particulier un tel λ est une valeur propre approchée de $(A,\mathscr{D})$.*

3.6. Le cas des opérateurs autoadjoints à spectre discret est particulièrement important. Considérons que l'opérateur $(A,\mathscr{D})$ (supposé densément défini sur $\mathscr{H}$) est autoadjoint et que son spectre ponctuel $\mathrm{Vp}(A,\mathscr{D})$ est non vide. Alors (comme dans le cas borné), il est facile de vérifier que $\mathrm{Vp}(A)$ est un sous-ensemble fini ou dénombrable de $\mathbb{R}$. Nous considérons alors que $(\lambda_n)_{n=0}^N$ $(0 \le N \le +\infty)$ est une suite (finie ou infinie) dont le support est $\mathrm{Vp}(A)$ et où chaque valeur propre apparait autant de fois que sa multiplicité (possiblement infini). Comme $(A,\mathscr{D})$ est autoadjoint, il existe une famille $(X_n)_{n=0}^N$ orthonormée de vecteurs de $\mathscr{D}$ t.q. $AX_n = \lambda_n X_n$. Si nous notons $\mathscr{D}'$ l'espace vectoriel engendré par $\{X_n\}_{n=0}^N$ alors la restriction $(A',\mathscr{D}')$ de $(A,\mathscr{D})$ à $\mathscr{D}'$ est un opérateur densément défini sur l'espace de Hilbert $\mathscr{K}$ engendré par $(X_n)_{n=0}^N$. Lorsque $(A',\mathscr{D}') = (A,\mathscr{D})$ (i.e. lorsqu'il existe une base hilbertienne de vecteurs propres) l'opérateur autoadjoint $(A,\mathscr{D})$ est dit *« à spectre discret »*. La proposition suivante décrit une classe importante *« d'opérateurs non bornés autoadjoints à spectre discret »* (c.f. [RSN55, Ch. VIII, § 120, p. 310] pour le cas général).

Théorème 3.19. *Soit $(X_n)_{n=0}^\infty$ une base hilbertienne de $\mathscr{H}$ et $(\lambda_n)_{n=0}^\infty$ une suite de réels tendant vers $+\infty$; alors (i) : l'ensemble $\mathscr{D}$ des $Y \in \mathscr{H}$ t.q. $(\lambda_n \langle X_n|Y\rangle)_{n=0}^\infty \in \ell^2(\mathbb{N})$ est dense dans $\mathscr{H}$; (ii) : $(B,\mathscr{D})$ t.q. $BX = \sum_n \lambda_n \langle X_n|X\rangle X_n$ est un opérateur non borné autoadjoint et (iii) : le spectre $\Sigma(B,\mathscr{D})$ coïncide avec le support de $(\lambda_n)_{n=0}^\infty$ (i.e. $(B,\mathscr{D})$ est à spectre discret).*

Preuve. (i)-(ii) : Comme $\mathscr{D}$ contient l'espace vectoriel engendré par les X_n, c'est un sous-espace dense de $\mathscr{H}$; il découle de la définition de $(B,\mathscr{D})$ que $BX_n = \lambda_n X_n$: par suite $(B,\mathscr{D})$ est un opérateur non borné du fait que les λ_n ne sont pas bornés. Pour montrer que $(B,\mathscr{D})$ est autoadjoint, rappelons d'abord que $\mathscr{D}^\star$ est l'ensemble des $Y \in \mathscr{H}$ t.q. $X \mapsto \langle Y|BX\rangle$ soit continue sur $\mathscr{D}$, l'adjoint de $(B,\mathscr{D})$ étant l'opérateur $(B^\star,\mathscr{D}^\star)$ t.q.

$$\forall X \in \mathscr{D}, \quad \forall Y \in \mathscr{D}^\star, \quad \langle Y|BX\rangle = \langle B^\star Y|X\rangle \tag{16}$$

Il est facile de vérifier que $(B,\mathscr{D})$ est symétrique, ce qui assure que $\mathscr{D} \subset \mathscr{D}^\star$; afin de vérifier que $\mathscr{D} = \mathscr{D}^\star$, nous remarquons simplement que tous les X_n sont dans $\mathscr{D}$, de sorte que pour tout $Y \in \mathscr{D}^\star$ nous tirons de (16) que $\langle B^\star Y|X_n\rangle = \lambda_n \langle Y|X_n\rangle$: comme $B^\star Y$ est dans $\mathscr{H}$, nous pouvons conclure que $(\lambda_n \langle Y|X_n\rangle)_n$ est dans $\ell^2(\mathbb{N})$; cela signifie que $Y \in \mathscr{D}$. Par conséquent, $(B,\mathscr{D})$ est bien autoadjoint du fait que c'est un opérateur densément défini symétrique et t.q. $\mathscr{D}^\star = \mathscr{D}$.

(iii) : Les vecteurs propres X_n associés aux valeurs propres λ_n formant (par hypothèse) une base hilbertienne de $\mathscr{H}$, le spectre $\Sigma(B,\mathscr{D})$ est réduit à l'ensemble $\mathrm{Vp}(B)$ des

10. Une sequence $X_1, X_2, \dots$ d'une espace de Hilbert séparable $\mathscr{H}$ est dite faiblement convergente sur $\mathscr{H}$, si pour tout $Y \in \mathscr{H}$, la suite $\langle X_1|Y\rangle, \langle X_2|Y\rangle, \dots$ est convergente (dans $\mathbb{C}$). Dans ce cas il existe un unique X_* t.q. $\langle X_n|Y\rangle \to \langle X_*|Y\rangle$ (en particulier, il est nécessaire que la suite $\|X_n\|$ soit bornée).

valeurs propres de B que constitue le support de la suite $(\lambda_n)_n$. Enfin, le fait que la suite des λ_n tende vers $+\infty$ assure que chacune des valeurs propres est de multiplicité finie : en d'autres termes, $(B, \mathscr{D})$ est à spectre discret.

□

3.7. Le théorème suivant est très utile dans beaucoup de situations de la mécanique quantique, lorsque le système étudié est à spectre discret (pour la notion d'opérateur compact, voir Appendice B).

Théorème 3.20. *Soit $(A, \mathscr{D})$ un opérateur densément défini sur un espace de Hilbert $\mathscr{H}$ (séparable) : alors les trois conditions suivantes sont équivalentes :*

(i) : $(A, \mathscr{D})$ est autoadjoint et il existe λ dans l'ensemble résolvant t.q. la résolvante $R_\lambda = (A - \lambda I)^{-1}$ soit un opérateur compact de $\mathcal{L}(\mathscr{H})$;

(ii) : Le spectre de $(A, \mathscr{D})$ est un sous-ensemble dénombrable et non borné de $\mathbb{R}$, sans point d'accumulations, uniquement formé de valeurs propres de multiplicité finie ; si $\lambda_0, \lambda_1, \ldots$ est une liste (non injective) des valeurs propres de $(A, \mathscr{D})$ (une valeur propre figurant autant de fois que sa multiplicité), alors il existe une base hilbertienne de $\mathscr{H}$ formée de vecteurs $X_0, X_1, \ldots$ dans $\mathscr{D}$ t.q.

$$Y \in \mathscr{D} \iff \left(\lambda_n \langle X_n|Y\rangle \ ; \ n \in \mathbb{N}\right) \in \ell^2(\mathbb{N}) \tag{17}$$

(iii) : $(A, \mathscr{D})$ est autoadjoint et pour tout λ dans l'ensemble résolvant, la résolvante $R_\lambda(A) = (A - \lambda I)^{-1}$ est un opérateur compact de $\mathcal{L}(\mathscr{H})$.

Preuve (d'après [ST89]**).** L'implication $(iii) \implies (i)$ étant évidente, il nous reste à démontrer que $(i) \implies (ii) \implies (i)$. Pour démontrer $(i) \implies (ii)$, remarquons que l'opérateur $(A, \mathscr{D})$ étant supposé autoadjoint, nous avons $(R_\lambda)^\star = (A - \overline{\lambda} I)^{-1}$. Cela entraîne en particulier que l'opérateur compact R_λ est normal et donc (Théorème spectral pour les opérateurs compacts normaux : c.f. Appendice B), qu'il existe une base hilbertienne $(X_0, X_1, \ldots)$ de $\mathscr{H}$ formée de vecteurs propres de R_λ avec $R_\lambda X_n = \xi_n X_n$. Notons que $\xi_0, \xi_1, \ldots$ forme alors la liste non nécessairement injective des valeurs propres de R_λ, comptées avec leur de multiplicité (finie) ; mais R_λ étant inversible, son rang est nécessairement infini et ses valeurs propres ξ_i sont des nombres complexes non nuls formant un sous-ensemble infinie bornée de $\mathbb{C}$: ce sous-ensemble possède donc au moins un point d'accumulation qui ne peut être que 0 (propriétés des opérateurs compacts), de sorte que $\lim_n \xi_n = 0$. Notons d'une part que l'image de R_λ étant égale à $\mathscr{D}$, l'identité $R_\lambda X_n = \xi_n X_n$ et le fait que $\xi_n \neq 0$ assure que X_n est dans $\mathscr{D}$: ainsi, $(X_0, X_1, \ldots)$ est une base hilbertienne de $\mathscr{H}$ formée de vecteurs dans $\mathscr{D}$. D'autre part, nous pouvons alors écrire la suite des implications suivantes :

$$\xi_n X_n = R_\lambda X_n \implies (A - \lambda I) X_n = \frac{1}{\xi_n} X_n \implies A X_n = \left(\lambda + \frac{1}{\xi_n}\right) X_n$$

soit encore $A X_n = \lambda_n X_n$ avec $\lambda_n = \lambda + 1/\xi_n$. Nous avons donc trouvé une suite $\lambda_0, \lambda_1, \ldots$ de valeurs propres de A (avec $\lambda_n \to +\infty$ quand $n \to +\infty$) respectivement associées à des vecteurs propres $X_1, X_2, \ldots$ dans $\mathscr{D}$ formant une base hilbertienne de $\mathscr{H}$. Le fait que les valeurs propres λ_n soient toutes de multiplicité finie et forment un sous-ensemble dénombrable et non borné de $\mathbb{R}$, sans point d'accumulations, découle d'une part de ce que $(A, \mathscr{D})$ est autoadjoint et d'autre part du fait que les ξ_n forment une partie bornée de $\mathbb{C}$ dont 0 est le seul point d'accumulation (i.e. $\lim_n \xi_n = 0$ entraîne que $\lim_n \lambda_n = +\infty$).

Pour montrer (17), notons que $Y \in \mathscr{D}$ ssi il existe $Z \in \mathscr{H}$ t.q. $R_\lambda Z = Y$. Si nous notons $(a_n) \in \ell^2(\mathbb{N})$ t.q. $Z = \sum_n a_n X_n$, alors $Y = \sum_n a_n \xi_n X_n$, soit encore

$$Y = \sum_n b_n X_n = \quad \text{avec} \quad b_n = \langle X_n | Y \rangle = \frac{a_n}{(\lambda - \lambda_n)}$$

Or comme $\lambda_n \to +\infty$, il existe un rang N t.q. $a_n = \lambda_n b_n(\lambda/\lambda_n - 1)$, pour $n \geq N$: par suite $(a_n)_n \in \ell^2(\mathbb{N})$ ssi $(\lambda_n b_n)_n \in \ell^2(\mathbb{N})$: la conclusion vient du fait que $|b_n \lambda_n| = |\langle A X_n | Y \rangle|$.

Il nous reste à démontrer $(ii) \Longrightarrow (iii)$. Supposons donc qu'il existe une base hilbertienne $(X_0, X_1, \dots)$ de $\mathscr{H}$ formée de vecteurs dans $\mathscr{D}$ t.q. $AX_n = \lambda_n X_n$ et où les λ_n forment une suite de points isolés de $\mathbb{R}$ et tendant vers $+\infty$: alors, d'après la Proposition 3.19 $(A, \mathscr{D})$ est autoadjoint. Etant donné $\lambda \in \Omega(A)$ arbitraire, nous voulons montrer que l'opérateur $(A - \lambda I, \mathscr{D})$ est inversible et que son inverse R_λ est compact. Pour cela nous commençons par résoudre l'équation $(A - \lambda I)X = Y$, d'inconnue X, lorsque $Y \in \mathscr{H}$ est arbitrairement donné. En décomposant l'identité $(A - \lambda I)X = Y$ dans la base hilbertienne formée par les X_n, il vient :

$$(\forall n) \quad \langle X_n | Y \rangle = \langle X_n | AX - \lambda X \rangle = (\lambda_n - \lambda) \langle X_n | X \rangle$$

Or λ étant pris dans $\Omega(A)$, nous avons $\lambda \neq \lambda_n$, pour tout n et par suite,

$$(\forall n) \quad \langle X_n | X \rangle = \frac{1}{\lambda_n - \lambda} \langle X_n | Y \rangle$$

Maintenant, comme $1/(\lambda_n - \lambda) \to 0$ quand $n \to +\infty$, en posant, pour tout $\forall Y \in \mathscr{H}$

$$RY = \sum_n \frac{1}{\lambda_n - \lambda} \langle X_n | Y \rangle X_n \tag{18}$$

nous définissons un opérateur nous $R \in \mathcal{L}(\mathscr{H})$ qui est compact. De plus, d'après (18) nous avons, pour tout $Y \in \mathscr{H}$

$$(A - \lambda I)RY = \sum_n \frac{1}{\lambda_n - \lambda} \langle X_n | Y \rangle (AX_n - \lambda X_n) = \sum_n \frac{1}{\lambda_n - \lambda} \langle X_n | Y \rangle (\lambda_n - \lambda) X_n = Y$$

Cela prouve que l'opérateur compact R est l'inverse de l'opérateur $A - \lambda I$.

□

4. Exemple d'opérateur autoadjoint : position

4.1. Soit $\mathscr{G}$ le sous-espace de $L^2(\mathbb{R})$ formé des $f(x)$ tel que $xf(x)$ soit aussi dans $L^2(\mathbb{R})$. Alors, $(Q, \mathscr{G})$ t.q. $Qf(x) = xf(x)$ est un opérateur densément défini appelé opérateur position. Il y a de multiple façons de vérifier que Q est non borné. Par exemple, si nous considérons f_0 dans la sphère unité de $\mathscr{G}$ (i.e. $\langle f_0 | f_0 \rangle = 1$) et telle que $\langle Q f_0 | f_0 \rangle = 0$ et si, pour tout $\alpha \in \mathbb{R}$, nous définissons $f_\alpha(x) := f_0(x - \alpha)$, alors f_α est dans la sphère unité de $\mathscr{G}$ et de plus, en effectuant le changement de variable $y = x - \alpha$

$$\begin{aligned}
\langle Q f_\alpha | f_\alpha \rangle &= \int x f_0(x - \alpha) \overline{f_0(x - \alpha)} dx \\
&= \int (y + \alpha) f_0(y) \overline{f_0(y)} dy \\
&= \int y f_0(y) \overline{f(y)} dy + \alpha \int |f_0(y)|^2 dy = \langle Q f_0 | f_0 \rangle + \alpha \langle f_0 | f_0 \rangle = \alpha
\end{aligned}$$

Ainsi le produit scalaire $\langle Q f | f \rangle$ n'est pas borné pour f décrivant la sphère unité de $\mathscr{G}$: cela prouve que $(Q, \mathscr{G})$ n'est pas borné.

Proposition 4.1. *Soit $\mathscr{G}$ le sous-espace de $L^2(\mathbb{R})$ des $f(x)$ t.q. $xf(x) \in L^2(\mathbb{R})$ et $(Q,\mathscr{G})$ l'opérateur densément défini (et non borné) t.q. $Qf(x) = xf(x)$; alors (i) : $(Q,\mathscr{G})$ est autoadjoint et (ii) : son spectre coïncide avec la droite réelle et n'est constitué que de valeurs propres approchées.*

Preuve. (i) : Il est immédiat que $(Q,\mathscr{G})$ est symétrique et nous voulons montrer qu'il est autoadjoint, c'est-à-dire que $\mathscr{G} = \mathscr{G}^\star$. L'inclusion $\mathscr{G} \subset \mathscr{G}^\star$ étant toujours vraie (pour un opérateur symétrique), prenons $g \in \mathscr{G}^\star$ et considérons pour tout $n \geq 0$ la fonction $G_n \in \mathscr{G}$ définie par $G_n(x) = xg(x)\mathbf{1}_{[-n\,;n]}(x)$. Le fait que $g \in \mathscr{G}^\star$, assure l'existence d'une constante $C > 0$ telle que $|\langle QG_n|g\rangle| \leq C\|G_n\|$, pour tout $n \geq 0$. Par suite,

$$\int_{-n}^{n} x^2|g(x)|^2dx \leq C\left(\int_{-n}^{n} x^2|g(x)|^2dx\right)^{1/2} \leq C\cdot C^{1/2}\left(\int_{-n}^{n} x^2|g(x)|^2dx\right)^{1/4} \leq \cdots$$

Du fait que $\log\left(C\cdot C^{1/2}\cdot C^{1/4}\right) = (1+1/2+1/4+\cdots)\log(C) = \log(C^2)$ il vient :

$$\int_{-n}^{n} x^2|g(x)|^2dx \leq C^2 < +\infty$$

c'est-à-dire que $g \in \mathscr{G}$.

(ii) : Soit $\lambda \in \mathbb{R}$ et pour tout entier $n > 0$, soit $f_n := \mathbf{1}_{[\lambda-1/n\,;\lambda+1/n]}$. Alors f_n est un élément de $\mathscr{G}$ t.q. $|\lambda f_n(x) - Qf(x)| = |t-\lambda|f_n(x) \leq 1/nf_n(x)$ et par suite :

$$\|\lambda f_n - Qf_n\|^2 \leq 1/n^2\|f_n\|^2 = 2/n^3$$

Le fait que $\|\lambda f_n - Qf_n\| \to 0$ quand $n \to 0$ assure que $\lambda \in \mathrm{Vp}'(Q,\mathscr{G})$: nous venons de démontrer que $\mathbb{R} \subset \mathrm{Vp}'(Q,\mathscr{G})$; or $\mathrm{Vp}'(Q,\mathscr{G}) \subset \Sigma(Q,\mathscr{G})$: la conclusion vient du Théorème 3.16 qui assure que $\Sigma(Q,\mathscr{G}) \subset \mathbb{R}$.

□

5. Exemple d'opérateur autoadjoint : impulsion

5.1. Nous commençons par l'opérateur de dérivation sur l'espace de Hilbert $L^2(I)$ pour $I = [0\,;1]$; pour cela, nous partons du fait (*« théorème d'approximation de Weiestrass »*) que le sous-espace $\mathscr{C}^\infty(I)$ des fonctions de classe C^∞ sur I est dense dans $L^2(I)$: alors la dérivation D est définie sur $\mathscr{C}^\infty(I)$ de sorte que pour tout $f(x)$ dans $\mathscr{C}^\infty(I)$:

$$Df(x) = \frac{df}{dx} = f'(x)$$

Il est évident que $De^{\lambda x} = \lambda e^{\lambda x}$, pour tout $\lambda \in \mathbb{C}$, ce qui assure que $\mathrm{Vp}(D,\mathscr{C}^\infty(I)) = \Sigma(D,\mathscr{C}^\infty(I)) = \mathbb{C}$: en particulier $(D,\mathscr{C}^\infty(I))$ est un opérateur densément défini et non borné sur $L^2(I)$. Nous allons voir qu'il existe un domaine (sous-espace dense de $L^2(I)$) sur lequel l'opérateur de dérivation D est proportionnel à un opérateur autoadjoint. Cela dépend d'une certaine condition au bord de l'intervalle I qui doit être imposée aux fonctions dans le domaine en question. L'idée est de partir d'une *« intégration par parties »*, en écrivant pour $f, g \in L^2(I)$ dérivables :

$$\langle Df|g\rangle = \int_0^1 f'(x)\overline{g}(x)dx = \left[f(x)\overline{g}(x)\right]_0^1 - \int_0^1 f(x)\overline{g}'(x)dx$$

de sorte que

$$\langle Df|g\rangle + \langle f|Dg\rangle = \left[f(x)\overline{g}(x)\right]_0^1 \tag{19}$$

Le membre droit de (19) ne s'annule pas en général. Cependant le sous-espace $\mathscr{C}_P^\infty(I)$ formé des $f \in \mathscr{C}^\infty$ t.q. $f(0) = f(1)$ est dense dans $L^2(I)$, puisque $e^{2i\pi nx} \in \mathscr{C}_P^\infty(I)$, pour tout $n \in \mathbb{Z}$; il découle alors de (19) que $(D, \mathscr{C}_P^\infty(I))$ est un opérateur densément défini sur $L^2(I)$ qui est anti-symétrique en ce sens que, pour tout $f, g \in \mathscr{C}_P^\infty(I)$

$$\langle Df | g \rangle = -\langle f | Dg \rangle \tag{20}$$

Cela fait de $(-iD, \mathscr{C}_P^\infty(I))$ un opérateur symétrique. Mais cet opérateur est trivialement diagonalisable ; en effet, les fonctions $f_n(x) := e^{2i\pi nx}$ appartiennent toutes à $\mathscr{C}_P^\infty(I)$ et de plus $-iDf_n = 2\pi n f_n$. Du fait que $(e^{2i\pi nx}\ ;\ n \in \mathbb{Z})$ est une base hilbertienne de $L^2(I)$, nous déduisons directement du Théorème 3.19 [11] la proposition suivante :

Proposition 5.1. *Soit $I = [0\,;1]$; l'opérateur de dérivation $(-iD, \mathscr{G}_P)$, où $\mathscr{G}_P$ est le sous-espaces de $L^2(I)$ des séries de Fourier $\sum_{n=-\infty}^{+\infty} c_n e^{2i\pi nx}$ avec $\sum_n |nc_n|^2 < +\infty$ est un opérateur non borné sur $L^2(I)$ qui est autoadjoint et de spectre discret, soit : $\Sigma(-iD, \mathscr{G}_P) = \mathrm{Vp}(-iD, \mathscr{G}_P) = 2\pi\mathbb{Z}$.*

5.2. Nous considérons maintenant la cas de l'espace de Hilbert $L^2(\mathbb{R})$ des fonctions complexes dont le carré du module est Lebesgue intégrable sur $\mathbb{R}$. Le sous-espace

$$\mathscr{G} := L^2(\mathbb{R}) \cap \mathscr{C}^1(\mathbb{R})$$

est dense dans $L^2(\mathbb{R})$ et nous nous intéressons ici à l'opérateur de dérivation spatiale $(D, \mathscr{G})$ qui à toute fonction $f(x)$ de $\mathscr{G}$, fait correspondre $Df(x) = df/dx = f'(x)$. L'opérateur $(D, \mathscr{G})$ ainsi obtenu est densément défini sur $L^2(\mathbb{R})$: nous allons vérifier qu'il est fermable dans $L^2(\mathbb{R})$ et identifier sa fermeture. Rappelons que l'espace de Sobolev $\mathscr{H}^1 = \mathscr{H}^1(\mathbb{R})$ est le sous-espace de $L^2(\mathbb{R})$ formé des fonctions $f(x)$ dérivables au sens faible et dont la dérivée $f'(x)$ est elle même dans $L^2(\mathbb{R})$. En d'autres termes, si $\mathscr{D}(\mathbb{R})$ désigne l'espace des fonctions tests sur $\mathbb{R}$ (fonctions C^∞ sur $\mathbb{R}$ et à support compact), alors la dérivée $f'(x)$ est la fonction $L^2(\mathbb{R})$ t.q. pour tout $\varphi \in \mathscr{D}(\mathbb{R})$:

$$\int f'(x)\varphi(x)dx = -\int f(x)\varphi'(x)dx$$

Proposition 5.2. *La fermeture de l'opérateur de dérivation $(D, \mathscr{G})$ est $(D, \mathscr{H}^1)$, où pour toute fonction $f(x) \in \mathscr{H}^1$, la valeur de $Df(x)$ coïncide avec la dérivée $f'(x)$ au sens faible ; de plus $(D, \mathscr{H}^1)$ est un opérateur antisymétrique en ce sens pour tout $f, g \in \mathscr{H}^1$*

$$\langle Df | g \rangle = -\langle f | Dg \rangle$$

Preuve. Notons $G_0 := \mathrm{Gr}(D, \mathscr{G})$ et $G_1 := \mathrm{Gr}(D, \mathscr{H}^1)$. Comme $G_0 \subset G_1$ devons démontrer que G_1 est égal à l'adhérence $\overline{G}_0$ (dans $L^2(\mathbb{R}) \times L^2(\mathbb{R})$) de G_0. Nous montrons d'abord que G_1 est fermé dans $L^2(\mathbb{R}) \times L^2(\mathbb{R})$: pour tout $\varphi \in \mathscr{D}(\mathbb{R})$, nous notons $G_1(\varphi)$ l'ensemble des $(f, g) \in L^2(\mathbb{R}) \times L^2(\mathbb{R})$ t.q.

$$\int \varphi'(x)\overline{f(x)}dx = -\int \varphi(x)\overline{g(x)}dx$$

mais par définition du produit scalaire $\langle\!\langle \cdot | \cdot \rangle\!\rangle$ sur $L^2(\mathbb{R}) \times L^2(\mathbb{R})$ tel que donné en (8), cela signifie simplement que $\langle\!\langle (\varphi', \varphi) | (f, g) \rangle\!\rangle = 0$: en d'autres termes $G_1(\varphi)$ est le noyau de la forme linéaire continue $\langle\!\langle (\varphi', \varphi) | \cdot \rangle\!\rangle$: comme G_1 coïncide avec l'intersection des $G_1(\varphi)$, pour φ décrivant $\mathscr{D}(\mathbb{R})$), nous en déduisons que G_1 est fermé dans $L^2(\mathbb{R}) \times L^2(\mathbb{R})$. Il nous

11. Voir aussi le Théorème 3.20.

reste maintenant à vérifier que $G_1 \subset \overline{G_0}$. Pour cela, considérons une *« approximation de l'unité $\mathscr{C}^1$ à support compact »*, e.g. une suite de fonctions $\theta_1, \theta_2, \dots$, avec $\theta_n(t) = n\theta(nt)$ avec $\theta = \theta_1$ une fonction $\mathscr{C}^1$ à support dans l'intervalle $[-1;1]$, positive et telle que $\int \theta(x)dx = 1$: cela assure que pour toute fonction $h \in L^2(\mathbb{R})$, les convolutions $\theta_n * h$ forment une suite de $\mathscr{G}$ convergeant [12] vers h dans $L^2(\mathbb{R})$. Soit $(f,g) \in G_1$; alors, pour tout rang n, nous savons d'une part qu'à la fois $\theta_n * f$ et $\theta_n * g$ sont dans $\mathscr{G}$; mais d'autre part, $\theta_n * g$ est aussi la dérivée (usuelle) de $\theta_n * f$: en effet, si $\varphi \in \mathscr{D}(\mathbb{R})$, alors grâce à intégration par parties (et par application du théorème de Fubini) :

$$\begin{aligned}
\int (\theta_n * f)'(x)\varphi(x)dx &= -\int \theta_n * f(x)\varphi'(x)dx \\
&= -\int \left(\int \theta_n(y) f(x-y)dy\right)\varphi'(x)dx \\
&= \int \theta_n(y)\left(-\int f(x-y)\varphi'(x)dx\right)dy \\
&= \int \theta_n(y)\left(\int g(x-y)\varphi(x)dx\right)dy \\
&= \int \left(\int \theta_n(y)g(x-y)dy\right)\varphi(x)dy = \int (\theta_n * g)(x)\varphi(x)dx
\end{aligned}$$

ce qui signifie que $(\theta_n * f)' = \theta_n * g$ au sens faible et donc au sens classique puisque $\theta_n * f$ est une fonction C^1 de $L^2(\mathbb{R})$. Ainsi $(\theta_n * f, \theta_n * g) = (\theta_n * f, (\theta_n * f)')$ est une suite de G_0 qui converge vers (f,g) : par suite $G_0 \subset G_1 \subset \overline{G_0}$ et donc $G_1 = \overline{G_0}$. Enfin, l'antisymétrie $\langle D\theta_n * f | \theta_n * g\rangle = -\langle \theta_n * f | D\theta_n * g\rangle$ est valable pour tout $n \geq 0$ comme identité dans $\mathscr{G}$: nous en déduisons (passage à la limite) que $\langle Df|g\rangle = -\langle f|Dg\rangle$ dans $\mathscr{H}^1$.

□

La Proposition 5.2 affirme que l'adjoint – dans $L^2(\mathbb{R})$ – de $(D, \mathscr{H}^1)$ est $(-D, \mathscr{H}^1)$. Nous pouvons reformuler et compléter ce résultat comme suit.

Proposition 5.3. *$(-\boldsymbol{i}D, \mathscr{H}^1)$ est un opérateur non borné autoadjoint de $L^2(\mathbb{R})$ dont le spectre coïncide avec l'axe réel entier.*

Preuve. Comme $(D, \mathscr{H}^1)$ est fermé et antisymétrique, l'opérateur $(-\boldsymbol{i}D, \mathscr{H}^1)$ est autoadjoint : son spectre $\Sigma(-\boldsymbol{i}D, \mathscr{H}^1)$ est donc inclus dans $\mathbb{R}$ (Théorème 3.16). Afin de montrer l'égalité $\Sigma(-\boldsymbol{i}D, \mathscr{H}^1) = \mathbb{R}$, nous allons montrer que tout réel λ est une valeur propre approchée de $(-\boldsymbol{i}D, \mathscr{H}^1)$. Pour voir cela, notons $\gamma(x) := e^{-x^2/2}$, de sorte qu'en effectuant le changement de variable $y = x/n$ il vient pour tout $n \geq 1$

$$\int \gamma\left(\frac{x}{n}\right)^2 dx = \int e^{-x^2/n^2}dx = n\int e^{-y^2}dy = n\sqrt{\pi}$$

Par suite, si $f_n(x) := \gamma(x/n)\, e^{\boldsymbol{i}\lambda x}/\sqrt{n}$, alors d'une part $\|f_n\|^2 = \sqrt{\pi}$ et d'autre part,

$$-\boldsymbol{i}Df_n(x) = \left(\lambda\gamma\left(\frac{x}{n}\right) - \frac{\boldsymbol{i}}{n}\gamma'\left(\frac{x}{n}\right)\right)\frac{e^{\boldsymbol{i}\lambda x}}{\sqrt{n}}$$

de sorte que

$$\| -\boldsymbol{i}Df_n - \lambda f_n\|^2 = \int \frac{1}{n^3}\gamma'\left(\frac{x}{n}\right)^2 dx = \frac{1}{n^2}\int \gamma'(y)^2\, dy = \frac{\|\gamma'\|^2}{n^2}$$

12. Voir e.g. [GW90, Proposition 21.3.6 & Remarque 21.3.7].

Nous en déduisons finalement que $\| -\boldsymbol{i}Df_n - \lambda f_n\| \to 0$ quand $n \to +\infty$.

□

6. Exemple d'opérateur autoadjoint : le laplacien

6.1. Pour $I := [0\,;1]$, l'étude du laplacien sur l'espace de Hilbert $L^2(I)$ est analogue à celle de l'opérateur D de dérivation effectuée au § 5.1. Nous commençons par considérer l'opérateur densément défini $(\Delta, \mathscr{C}^\infty(I))$, où $\mathscr{C}^\infty(I)$ et où

$$\Delta f(x) = d^2 f/dx^2 = f''(x)$$

Il est immédiat que $(\Delta, \mathscr{C}^\infty(I))$ est non borné : en effet, $\Delta e^{\lambda x} = \lambda^2 e^{\lambda x}$, pour tout $\lambda \in \mathbb{C}$, ce qui assure que $\mathrm{Vp}(\Delta, \mathscr{C}^\infty(I)) = \Sigma(\Delta, \mathscr{C}^\infty(I)) = \mathbb{C}$. Mais d'autre part, grâce à une double *« intégration par partie »*, nous avons pour tout $f, g \in \mathscr{C}^\infty$:

$$\langle f|\Delta g\rangle = \int_0^1 f(x)\overline{g''(x)}dx = \Big[f(x)\overline{g'(x)}\Big]_0^1 - \Big[f'(x)\overline{g(x)}\Big]_0^1 + \int_0^1 f''(x)\overline{g}(x)dx$$

de sorte que

$$\langle f|\Delta g\rangle - \langle \Delta f|g\rangle = \Big[f(x)\overline{g'(x)}\Big]_0^1 - \Big[f'(x)\overline{g(x)}\Big]_0^1 \tag{21}$$

Par suite, si nous considérons $\mathscr{C}_P^\infty(I)$ le sous-espace de $\mathscr{C}^\infty(I)$ formé des restrictions à l'intervalle I de fonctions C^∞ sur $\mathbb{R}$ et 1-périodiques, alors $\mathscr{C}_P^\infty(I)$ est toujours dense dans $L^2(I)$ et il découle immédiatement de (21) que $(\Delta, \mathscr{C}_P^\infty(I))$ est un opérateur densément défini sur $L^2(I)$ qui est symétrique. Remarquons d'une part que (malgré la restriction du domaine de $\mathscr{C}^\infty(I)$ à $\mathscr{C}_P^\infty(I)$), l'opérateur densément défini $(\Delta, \mathscr{C}_P^\infty(I))$ demeure non borné : en effet, pour tout $n \in \mathbb{Z}$, les $f_n(x) := e^{2i\pi nx}$ sont des fonctions de $\mathscr{C}_P^\infty(I)$ telles que $\Delta f_n = -(\pi n)^2 f_n$ et donc $\{ -(2\pi n)^2 \;;\; n \geq 1\} \subset \mathrm{Vp}(\Delta, \mathscr{C}_P^\infty(I))$. Remarquons aussi d'autre part, que le domaine $\mathscr{C}_P^\infty(I)$ n'est pas optimal pour Δ, puisque $\mathscr{C}_P^\infty(I)^\star$ contient (par exemple) le sous-espace $\mathscr{C}_P^2(I)$ des restrictions de fonctions de classe C^2 sur $\mathbb{R}$ et 1-périodiques : cela montre que $(\Delta, \mathscr{C}_P^\infty)$ n'est pas autoadjoint sur $L^2(I)$. L'identification de la *« version autoadjointe de Δ »* nécessité – en un sens – de trouver le domaine $\mathscr{C}_P^\infty(I) \subset \mathscr{G} \subset L^2(I)$ optimal t.q. $(\Delta, \mathscr{G})$ soit symétrique et ensuite de montrer que $\mathscr{G}^\star = \mathscr{G}$. Notons $\mathscr{H}^2(I)$ *« l'espace de Sobolev »* des fonctions $f \in L^2(I)$ deux fois différentiables au sens faible avec $f'' \in L^2(I)$; nous considérons aussi le sous-espace $\mathscr{H}_P^2(I)$ de $\mathscr{H}^2(I)$ constitué des restrictions f à l'intervalle I de fonctions 1-périodiques deux fois dérivables au sens faible et t.q. f'' soient dans $L^2(I)$. On peut alors montrer (c.f. [Oli16a, § 9]) que $\mathscr{H}_P^2(I)$ est un sous-espace de l'espaces $\mathscr{C}_P^1(I)$ des restrictions à l'intervalle I de fonctions $f : \mathbb{R} \to \mathbb{R}$ de classe C^1 qui sont 1-périodiques et que

$$\mathscr{H}_P^2(I) := \Big\{ f \in \mathscr{H}^2(I) \;;\; \big(f(0), f'(0)\big) = \big(f(1), f'(1)\big)\Big\} \tag{22}$$

Remarquons alors que d'après (21) l'opérateur $(\Delta, \mathscr{H}_P^2)$ est symétrique.

Proposition 6.1. *Le laplacien $(\Delta, \mathscr{H}_P^2(I))$ est un opérateur non borné sur $L^2(I)$ qui est autoadjoint et de spectre discret, avec $\Sigma(\Delta, \mathscr{H}_P^2(I)) = \mathrm{Vp}(\Delta, \mathscr{H}_P^2(I)) = \{ -(2\pi n)^2 \;;\; n \geq 1\}$.*

Preuve. C'est une conséquence du Théorème 3.19 et du fait que le laplacien $(\Delta, \mathscr{H}^2_P(I))$ est un opérateur autoadjoint qui possède un spectre discret dont les fonctions propres $e^{2ik\pi x}$ $(k \geq 1)$ forment une base hilbertienne de $L^2(I)$.

□

6.2. Nous allons adapter les calculs effectués au § 5.2 pour l'opérateur de dérivation au cas du laplacien sur $L^2(\mathbb{R})$. Pour cela, nous considérons le sous-espace dense

$$\mathscr{K} := L^2(\mathbb{R}) \cap \mathscr{C}^2(\mathbb{R})$$

de sorte que pour tout $f \in \mathscr{K}$, nous notons Δf la dérivée seconde de f (au sens classique) et qui par définition de $\mathscr{K}$ est encore un élément de $L^2(\mathbb{R})$. L'opérateur $(\Delta, \mathscr{K})$ ainsi défini est densément défini sur $L^2(\mathbb{R})$ et non borné. Afin de déterminer sa fermeture rappelons que l'espace de Sobolev $\mathscr{H}^2 = \mathscr{H}^2(\mathbb{R})$ est le sous-espace de $L^2(\mathbb{R})$ formé des fonctions $f(x)$ deux fois dérivables au sens faible et dont la dérivée seconde $f''(x)$ est elle même dans $L^2(\mathbb{R})$. En d'autres termes, si $\mathscr{D}(\mathbb{R})$ désigne l'espaces des fonctions tests sur $\mathbb{R}$ (fonctions C^∞ sur $\mathbb{R}$ et à support compact), alors $f''(x)$ est la fonction $L^2(\mathbb{R})$ t.q. pour tout $\varphi \in \mathscr{D}(\mathbb{R})$:

$$\int f''(x)\varphi(x)dx = \int f(x)\varphi''(x)dx$$

Proposition 6.2. *La fermeture de $(\Delta, \mathscr{K})$ est l'opérateur autoadjoint $(\Delta, \mathscr{H}^2)$.*

Preuve. Notons $G_0 := \mathrm{Gr}(\Delta, \mathscr{K})$ et $G_1 := \mathrm{Gr}(\Delta, \mathscr{H}^2)$. Comme $G_0 \subset G_1$ il reste à démontrer que G_1 est égal à l'adhérence $\overline{G_0}$ (dans $L^2(\mathbb{R}) \times L^2(\mathbb{R})$) de G_0. Pour cela considérons une suite $\theta_n(t) = 2^n\theta_0(2^n t)$ $(n \geq 0$ et $\theta_0 \in \mathscr{K})$ d'approximation de l'unité dans $\mathscr{K}$. Etant donné $(f, g) \in G_1$, nous savons d'une part que $\theta_n * f$ et $\theta_n * g$ sont dans $\mathscr{K}$; mais d'autre part, $\theta_n * g$ est aussi la dérivée (en distribution) de $\theta_n * f$: en effet, si $\varphi \in \mathscr{D}(\mathbb{R})$, alors (par intégrations par parties puis par une double application du théorème de Fubini) :

$$\begin{aligned}
\int (\theta_n * f)''(x)\varphi(x)dx &= \int \theta_n * f(x)\varphi''(x)dx \\
&= \iint \theta_n(y)f(x-y)\varphi''(x)dxdy \\
&= \int \theta_n(y)\Big(\int f(x-y)\varphi''(x)dx\Big)dy \\
&= \iint \theta_n(y)g(x-y)\varphi(x)dxdy = \int (\theta_n * g)(x)\varphi(x)dx
\end{aligned}$$

ce qui signifie que $(\theta_n * f)'' = \theta_n * g$. Ainsi les $(\theta_n * f, \theta_n * g) = (\theta_n * f, (\theta_n * f)')$ pour $n \geq 0$, forment-ils une suite de G_0 qui converge vers (f, g). Nous avons ainsi démontré la suite d'inclusion $G_0 \subset G_1 \subset \overline{G_0}$, ce qui assure que $G_1 = \overline{G_0}$. Enfin, $(\Delta, \mathscr{H}^2)$ est un opérateur autoadjoint comme fermeture d'un opérateur symétrique.

□

Proposition 6.3. *$(-\Delta, \mathscr{H}^2)$ est un opérateur non borné autoadjoint de $L^2(\mathbb{R})$ dont le spectre coïncide avec le demi-axe des réels positifs ou nuls.*

Preuve. Il est facile de vérifier (double intégration par partie) que $(-\Delta, \mathscr{S}(\mathbb{R}))$ est symétrique et positif, ce qui assure $(-\Delta, \mathscr{H}^2)$ est un opérateur autoadjoint positif : par suite (c.f. Proposition 3.17) nous avons $\Sigma(-\Delta, \mathscr{H}^2) \subset [0\,; +\infty[$. D'après la Proposition 3.17, il nous reste à montrer que tout réel positif ou nul est une valeur propre approchée de $(-\Delta, \mathscr{H}^2)$. Pour voir cela, notons $\gamma(x) := e^{-x^2/2}$, de sorte qu'en effectuant le changement de variable $y = x/n$ il vient pour tout $n \geq 1$

$$\int \gamma\left(\frac{x}{n}\right)^2 dx = \int e^{-x^2/n^2} dx = n \int e^{-y^2} dy = n\sqrt{\pi}$$

Soit λ un réel arbitraire et, pour tout entier $n \geq 0$, soit $f_n(x) := \gamma(x/n)\, e^{-\boldsymbol{i}\lambda x}/\sqrt{n}$; nous avons alors d'une part $\|f_n\| = \sqrt{\pi}$ et d'autre part,

$$\begin{aligned} -\Delta f_n(x) &= -\frac{d}{dx}\left[\left(-\boldsymbol{i}\lambda\gamma\left(\frac{x}{n}\right) + \frac{1}{n}\gamma'\left(\frac{x}{n}\right)\right)\frac{e^{-\boldsymbol{i}\lambda x}}{\sqrt{n}}\right] \\ &= \left(\lambda^2\gamma\left(\frac{x}{n}\right) + \frac{\boldsymbol{i}\lambda}{n}\gamma'\left(\frac{x}{n}\right) + \frac{\boldsymbol{i}\lambda}{n}\gamma'\left(\frac{x}{n}\right) - \frac{1}{n^2}\gamma''\left(\frac{x}{n}\right)\right)\frac{e^{-\boldsymbol{i}\lambda x}}{\sqrt{n}} \end{aligned}$$

Or nous avons $\gamma' \in L^2(\mathbb{R})$ avec $\int 1/n^2\gamma'(x/n)^2\, dx = \|\gamma'\|^2/n$ et de même $\gamma'' \in L^2(\mathbb{R})$ avec $\int 1/n^4\gamma''(x/n)^2\, dx = \|\gamma''\|^2/n^3$ de sorte que

$$\| - \Delta f_n(x) - \lambda^2 f_n(x)\| \leq \frac{|\lambda|\|\gamma'\|}{n} + \frac{\|\gamma''\|}{n^2}$$

Ainsi $\| - \Delta f_n(x) - \lambda^2 f_n(x)\| \to 0$ quand $n \to +\infty$ et donc $\lambda^2 \in \mathrm{Vp}'(-\Delta, \mathscr{H}^2)$.

□

7. Théorème de Stone - Théorème de Hille-Yosida

7.1. Au début des années 30, Stone [Sto30, Sto32] démontre un résultat sur les groupes à un paramètre d'isomorphisme d'espace de Hilbert qui donne un cadre théorique à la résolution abstraite de l'équation de Schrödinger. Plus précisément, étant donné $\mathscr{H}$ un espace de Hilbert, une famille $\{U_t\ ;\ t \in \mathbb{R}\}$ d'opérateurs inversibles de $\mathcal{L}(\mathscr{H})$ est un *« groupe unitaire continu à un paramètre »* si d'une part $U_0 = I$ avec $U_{s+t} = U_sU_t$, pour tout $s, t \in \mathbb{R}$ et si d'autre part, la convergence simple $U_tX \to X$ a lieu, dans la limite où $t \to 0$, pour tout $X \in \mathscr{H}$. Il est alors immédiat de vérifier $U_0 = I$ (opérateur identité) et que $t \mapsto U_tX$ est un chemin continu dans E, pour tout $X \in E$. Si $\mathscr{D}$ est l'espace vectoriel des $X \in \mathscr{H}$ t.q. $\boldsymbol{i}(U_tX - X)/t$ possède une limite AX lorsque $t \to 0$, alors il est immédiat que $\mathscr{D}$ est un espace vectoriel (éventuellement réduit à $\{0\}$) et que $X \mapsto AX$ est une application linéaire de $\mathscr{D}$ dans $\mathscr{H}$; le *« théorème de Stone »* affirme d'une part que $(A, \mathscr{D})$ est un opérateur (densément défini et) autoadjoint sur $\mathscr{H}$ et que, pour tout $X \in \mathscr{D}$, l'unique solution du problème de Cauchy $Y'(t) = \boldsymbol{i}AY(t)$ avec $Y(0) = X$ est $Y(t) = U(t)X$ (on note alors $U_t = \exp(-\boldsymbol{i}At)$) ; réciproquement, pour tout opérateur $(A, \mathscr{D})$ (densément défini et) autoadjoint sur $\mathscr{H}$ et pour tout $X \in \mathscr{D}$, l'unique solution du problème le problème de Cauchy $Y'(t) = \boldsymbol{i}AY(t)$ avec $Y(0) = X$ est $Y(t) = U_tX$, où $\{U_t\ ;\ t \in \mathbb{R}\}$ est un groupe unitaire continu à un paramètre sur $\mathscr{H}$. Nous allons voir comment déduire une version de ce résultat (dans le cas des semi-groupe : c.f. Théorème 7.20 infra), à partir du théorème de Hille-Yosida (pour les espaces de Banach).

7.2. Soit $\mathscr{E}$ un espace de Banach et considérons que $\{S(t)\ ;\ t \geq 0\}$ est un *« semi-groupe »* d'endomorphismes de $\mathcal{L}(\mathscr{E})$: en d'autres termes, $S(t) \in \mathcal{L}(\mathscr{E})$, pour tout $t \geq 0$, avec

$S(0) = I$ et $S(s+t) = S(s)S(t)$, pour tout $s,t \geq 0$. Nous supposerons toujours que le semi-groupe est *« fortement continu »* (i.e. ponctuellement continu en $t = 0$) en ce sens que

$$\forall X \in \mathscr{E}, \quad \lim_{t \to 0} S(t)X = X$$

Lorsque $\|S(t) - I\| \to 0$ quand $t \to 0$ le semi-groupe est dit *« uniformément continu »*.

Proposition 7.1. *$\{S(t)\ ;\ t \geq 0\}$ est un semi-groupe uniformément continu de $\mathcal{L}(\mathscr{E})$ si et seulement si il existe $A \in \mathcal{L}(\mathscr{E})$ tel que $S(t) = \exp(tA) = \sum_{n=0}^{+\infty}(tA)^n/n!$.*

Preuve. Si A, B sont deux endomorphismes de $\mathcal{L}(\mathscr{E})$ qui commutent alors $\exp(A+B) = \exp(A)\exp(B)$: cela entraîne que $t \mapsto \exp(tA)$ est un semi-groupe uniformément continu de $\mathcal{L}(\mathscr{E})$ dès que $A \in \mathcal{L}(\mathscr{E})$. Réciproquement, supposons que $t \mapsto S(t)$ soit un semi-groupe uniformément continu de $\mathcal{L}(\mathscr{E})$ et pour tout $\varepsilon > 0$, définissons

$$M(\varepsilon) := \frac{1}{\varepsilon} \int_0^{\varepsilon} S(s) ds$$

(Nous utilisons ici *« l'intégrale de Bochner »* : c.f. Appendice D.) Alors $M(\varepsilon) \in \mathcal{L}(\mathscr{E})$ et en intégrant de 0 à ε l'identité $S(s+t) = S(s)S(t)$, il vient

$$\int_t^{\varepsilon+t} S(\sigma)d\sigma = \int_0^{\varepsilon} S(s+t)ds = \int_0^{\varepsilon} S(s)S(t)ds = \varepsilon M(\varepsilon)S(t) \tag{23}$$

Mais l'uniforme continuité du semi-groupe $S(t)$ assure que $M(\varepsilon) \to I$ quand $\varepsilon \to 0$: par suite (lemme de l'inverse de von Neumann), lorsque $\varepsilon > 0$ est assez petit, $M(\varepsilon)$ est inversible dans $\mathcal{L}(\mathscr{E})$. Partant de (23) – avec $\varepsilon > 0$ assez petit – nous obtenons

$$S(t) = \frac{1}{\varepsilon} M(\varepsilon)^{-1} \left(\int_t^{\varepsilon+t} S(\sigma)d\sigma \right)$$

ce qui assure que $S(t)$ est un chemin C^1 de $\mathcal{L}(\mathscr{E})$. En particulier $A := S'(0) \in \mathcal{L}(\mathscr{E})$ et en prenant la dérivée de $S(s+t) = S(s)S(t)$ par rapport à s et évaluée en $s = 0$, il vient

$$S'(t) = \left.\frac{d}{ds}\right|_{s=0} S(s+t) = \left(\left.\frac{d}{ds}\right|_{s=0} S(s) \right) S(t) = AS(t)$$

□

7.3. Lorsque $\mathscr{E}$ n'est pas de dimension finie, un semi-groupe fortement continu d'endomorphisme n'est pas nécessairement uniformément continu. Cependant, si $\{S(t)\ ;\ t \geq 0\} \subset \mathcal{L}(\mathcal{E})$ est un semi-groupe fortement continu, la continuité des trajectoires demeure : en effet, pour $X \in \mathscr{E}$ et $t \geq 0$ notons que pour tout h t.q. $t + h \geq 0$:

$$\|S(t+h)X - S(t)X\| \leq \|S(t)\| \cdot \|S(h)X - X\|$$

La continuité forte du semi-groupe entraîne donc la continuité du chemin $t \mapsto S(t)X$.

Proposition 7.2. *Si $\{S(t)\ ;\ t \geq 0\}$ est un semi-groupe fortement continu de $\mathcal{L}(\mathscr{E})$, alors pour tout $X \in \mathscr{E}$, l'application $t \mapsto S(t)X$ est un chemin continu de $\mathscr{E}$.*

Dans le cas où l'espace de Banach $\mathscr{E}$ est l'espace de Hilbert associé à un système quantique, la classe des semi-groupes uniformément continus, n'est généralement pas suffisante pour traiter les questions de dynamique quantique et il est nécessaire d'élargir cette classe aux semi-groupes fortement continu. La difficulté qui se pose alors est la différentiabilité des chemins $t \mapsto S(t)X$: lorsque $\{S(t)\ ;\ t \geq 0\}$ est uniformément continu, nous pouvons écrire (c.f. Proposition 7.1) $S(t) = \exp(tA)$, où A est l'opérateur de $\mathcal{L}(\mathscr{E})$ tel que pour tout $X \in \mathscr{E}$

$$AX = \lim_{t\to 0} \frac{1}{t}\Big(S(t)X - X\Big)$$

Il est alors raisonnable de poser la définition suivante.

Définition 7.3. *Soit $\{S(t)\ ;\ t \geq 0\} \subset \mathcal{L}(\mathscr{E})$ un semi-groupe fortement continu et soit $\mathscr{D}$ le sous-espace des $X \in \mathscr{E}$ t.q. la limite AX de $(S(t)X - X)/t$ lorsque $t \to 0$ existe ; alors A est une application linéaire de $\mathscr{D}$ dans $\mathscr{E}$ et $(A, \mathscr{D})$ est le « générateur infinitésimal » de $\{S(t)\ ;\ t \geq 0\}$.*

Lorsque X appartient au domaine du générateur infinitésimal du semi-groupe fortement continu $\{S(t)\ ;\ t \geq 0\}$, il est alors possible d'interpréter les trajectoires $t \mapsto S(t)X$ comme l'unique solution d'un problème de Cauchy sur $\mathscr{E}$: c'est l'objet du théorème (important) que nous allons maintenant démontrer.

Théorème 7.4. *Soit $(A, \mathscr{D})$ le générateur infinitésimal d'un semi-groupe fortement continu $\{S(t)\ ;\ t \geq 0\} \subset \mathcal{L}(\mathscr{E})$ et soit $X \in \mathscr{D}$: alors (i) : pour tout $t \geq 0$, $S(t)X \in \mathscr{D}$ et $AS(t)X = S(t)AX$; (ii) : le chemin (continu) $t \mapsto Y(t) := S(t)X$ ($t \geq 0$) est continûment dérivable, de support inclus dans $\mathscr{D}$ et satisfait le problème de Cauchy*

$$\text{(C)} :\quad Y'(t) = AY(t) \quad \textit{avec} \quad Y(0) = X$$

(iii) : Si $t \mapsto Z(t)$ ($t \geq 0$) est un chemin dérivable de support dans $\mathscr{D}$ t.q. $Z'(t) = AZ(t)$, alors $Z(0) = X$ entraîne $Z(t) = S(t)X$: autrement dit, $t \mapsto S(t)X$ est l'unique solution de (C).

Pour démontrer ce théorème nous aurons besoin du lemme suivant.

Lemme 7.5. *Si $\{S(t)\ ;\ t \geq 0\}$ est un semi-groupe fortement continu de $\mathcal{L}(\mathscr{E})$, alors il existe $C > 0$ et $\omega \geq 0$ tels que (i) : $\|S(t)\| \leq Ce^{\omega t}$, pour tout $t \geq 0$ et (ii) : pour tout $X \in \mathscr{E}$,*

$$\forall t \geq 0,\ \forall h \in \mathbb{R},\ t + h \geq 0 \implies \|S(t+h)X - S(t)X\| \leq Ce^{\omega t}\|S(h)X - X\|$$

Preuve. Pour $X \in \mathscr{E}$ donné, nous savons (c.f. Proposition 7.2) que $t \mapsto S(t)X$ est un chemin continu de $\mathscr{E}$. Nous en déduisons alors que $\sup\big\{\|S(t)X\|\ ;\ t \in [0;1]\big\} < +\infty$ et donc (*« théorème de la borne uniforme »* : c.f. Appendice A), il existe une constant $C < +\infty$ assurant la borne uniforme des normes opérateurs des $S(t)$ pour $0 \leq t \leq 1$ i.e.

$$\sup\big\{\|S(t)\|\ ;\ t \in [0;1]\big\} < C \tag{24}$$

Soit alors $t = n + \tau$, où l'entier $n \geq 0$ et le réel $0 \leq \tau < 1$ sont arbitrairement donnés ; alors $S(t) = S(\tau)S(n) = S(\tau)S(1)^n$ et donc d'après l'inégalité (24),

$$\|S(t)\| \leq \|S(\tau)\| \cdot \|S(1)\|^n \leq C \cdot C^n \leq Ce^{\omega t}$$

où nous avons noté $\omega = \log C$. Nous avons $\omega \geq 0$, car $C \geq \|S(0)\| = 1$.

(ii) : Soit $X \in \mathscr{E}$; alors, d'une part pour tout $h > 0$, nous pouvons utiliser les constantes C et ω données en (i) pour écrire

$$\|S(t+h)X - S(t)X\| = \|S(t)(S(h)X - X)\| \leq Ce^{\omega t}\|S(h)X - X\| \tag{25}$$

Mais d'autre part si $t - h > 0$, l'inégalité (25) appliquée en $t - h$ (au lieu de t) donne

$$\|S(t-h)X - S(t)X\| \leq Ce^{\omega(t-h)}\|S(h)X - X\| \leq Ce^{\omega t}\|S(h)X - X\| \tag{26}$$

□

Preuve du Théorème 7.4. (i) : Etant donné $X \in \mathscr{D}$ arbitraire et $h > 0$, nous avons :

(27) $$\left(\frac{S(h) - I}{h}\right) S(t)X = S(t)\left(\frac{S(h) - I}{h}\right) X$$

En faisant tendre $h \to 0$, la continuité de $S(t)$ et le fait que $X \in \mathscr{D}$, assure que le membre de droite de (27) converge vers $S(t)AX$: la convergence du membre de gauche signifiant que $S(t)X \in \mathscr{D}$, nous obtenons $AS(t)X = S(t)AX$.

(ii) : Le fait que $X \in \mathscr{D}$ entraîne, par définition même du générateur infinitésimal $(A, \mathscr{D})$, que $Y(t) = S(t)X$ est dérivable en 0^+ avec $Y'(0^+) = AX$. Afin de montrer que $Y'(t) = S(t)AX$, nous écrivons pour tout h t.q. $0 < h < t$:

$$\frac{Y(t) - Y(t-h)}{h} - S(t)AX = S(t-h)\left(\frac{S(h)X - X}{h} - AX\right) + \left(S(t-h)AX - S(t)AX\right)$$

Si C et ω sont des constantes positives t.q. $\|S(t)\| \le Ce^{\omega t}$ (c.f. Lemme 7.5), il vient

$$\left\|\frac{Y(t) - Y(t-h)}{h} - S(t)AX\right\| \le Ce^{t\omega}\left(\left\|\frac{Y(h) - Y(0)}{h} - Y'(0^+)\right\| + \|S(h)AX - AX\|\right)$$

Le fait que $Y'(0^+) = AX$ combiné à la continuité du chemin $h \mapsto S(h)AX$ avec $S(0)AX = AX$ (c.f. Proposition 7.2) nous permet alors de conclure que

$$\lim_{h\to 0}\left\|\frac{Y(t) - Y(t-h)}{h} - S(t)AX\right\| = 0$$

Ainsi $Y(t)$ est dérivable en t avec $Y'(t) = S(t)AX$ et donc continûment dérivable ; mais de plus nous pouvons aussi écrire – par la définition même de la dérivée $Y'(t)$ – que

$$S(t)AX = Y'(t) = \lim_{h\to 0}\frac{S(h)S(t)X - S(t)X}{h}$$

de sorte que $S(t)X \in \mathscr{D}$ avec $S(t)AX = AS(t)X$.

(iii) : Soit $X \in \mathscr{D}$ donné et soit $t \mapsto Z(t)$ $(t \ge 0)$ un chemin continûment dérivable de support dans $\mathscr{D}$ t.q. $Z'(t) = AZ(t)$ avec $Z(0) = X$. Pour $t_0 > 0$ arbitrairement donné et $0 \le s \le t_0$, nous posons $V(t) := S(t_0 - t)Z(t)$; alors, d'après (i) et (ii) :

$$V'(t) = -AS(t_0 - t)Z(t) + S(t_0 - t)AZ(t) = 0$$

Par suite $V(t) \equiv V(0)$ pour tout $0 \le t \le t_0$: en particulier $V(0) = V(t_0)$; or $V(t_0) = Z(t_0)$ et comme $V(t_0) = V(0) = S(t_0)Z(0) = S(t_0)X$, il vient $Z(t_0) = S(t_0)X$.

□

Remarque 7.6. *Soit $(A, \mathscr{D})$ le générateur infinitésimal d'un semi-groupe fortement continu $\{S(t)\ ;\ t \ge 0\} \subset \mathcal{L}(\mathscr{E})$; si $\mathscr{D} = \mathscr{E}$ alors, nous considérons (convention [13]) que A est un endomorphisme de $\mathcal{L}(\mathscr{E})$. Dans ce cas, pour tout $X \in \mathscr{D} = \mathscr{E}$ nous savons que $t \mapsto S(t)X$ et $t \mapsto \exp(tA)X$ (avec $\exp(tA) := \sum_{n=0}^{\infty}(tA)^n/n!$) sont solution du problème de Cauchy $Y'(t) = AY(t)$ avec la condition initiale $Y(0) = X$: par unicité de la solution de ce problème (c.f. (iii)-Théorème 7.4) nous obtenons $S(t)X = \exp(tA)X$, pour tout $X \in \mathscr{E}$, i.e. $S(t) = \exp(tA)$. En particulier, cela signifie que le semi-groupe $\{S(t)\ ;\ t \ge 0\}$ est uniformément continu sur $\mathscr{E}$.*

13. Rappelons que la convention porte sur la définition des endomorphismes densément définis sur $\mathscr{E}$: lorsque $(A, \mathscr{D})$ est densément défini sur $\mathscr{E}$ et que $\mathscr{D} = \mathscr{E}$, nous convenons que $A \in \mathcal{L}(\mathscr{E})$. L'existence d'endomorphismes de $\mathscr{E}$ non bornés dépend (essentiellement) de l'axiome du choix.

Théorème 7.7. *Soit* $(A, \mathscr{D})$ *le générateur infinitésimal d'un semi-groupe fortement continu* $\{S(t)\ ;\ t \geq 0\} \subset \mathcal{L}(\mathscr{E})$ *; alors* $(A, \mathscr{D})$ *est nécessairement densément défini sur* $\mathscr{E}$ *et fermé.*

La démonstration du Théorème 7.7 est basée sur le lemme suivant.

Lemme 7.8. *Soit* $\{S(t)\ ;\ t \geq 0\} \subset \mathcal{L}(\mathscr{E})$ *un semi-groupe fortement continu ; si pour tout* $X \in \mathscr{E}$ *et tout* $t > 0$, *nous notons*

$$M(t)X = \frac{1}{t}\int_0^t S(u)X du \tag{28}$$

alors (i) *: pour tout* $t \geq 0$, *l'application* $X \mapsto M(t)X$ *est un endomorphisme continu de* $\mathscr{E}$ *– i.e.* $M(t) \in \mathcal{L}(\mathscr{E})$ *;* (ii) *:* $t \mapsto M(t)$ *est une approximation de l'identité en* $t = 0$ *en ce sens que pour tout* $X \in \mathscr{E}$, $M(t)X \to X$, *quand* $t \to 0$ *;* (iii) *: pour tout* $s, t > 0$

$$\frac{S(s) - I}{s} M(t) = \frac{S(t) - I}{t} M(s) = M(s)\frac{S(t) - I}{t} \tag{29}$$

(iv) *: si* $(A, \mathscr{D})$ *est le générateur infinitésimal de* $\{S(t)\ ;\ t \geq 0\}$, *alors* $M(t)X \in \mathscr{D}$, *pour tout* $X \in \mathscr{E}$ *(i.e. l'image de* $M(t)$ *est dans* $\mathscr{D}$*) et*

$$AM(t) = \frac{S(t)X - X}{t}$$

De plus, si $X \in \mathscr{D}$ *alors pour tout* $t > 0$, *nous avons*

$$AM(t)X = M(t)AX$$

Preuve. (i) : Il est immédiat (par définition de l'intégrale de Bochner : c.f. Appendice D) que l'application $X \mapsto M(s)X$ définie en (28) est linéaire et que

$$\|M(s)X\| \leq \frac{1}{s}\int_0^s \|S(t)X\| dt \leq Ce^{\omega s}\|X\|$$

où les constantes C et ω sont associées à $\{S(t)\ ;\ t \geq 0\}$ par le Lemme 7.5.

(ii) : Etant donné $X \in \mathscr{E}$, la restriction de l'application $t \mapsto S(t)X$ à l'intervalle $[0\,;1]$ étant une application continue sur un compact [14], il existe un module de continuité (uniforme) $\mu(\eta)$ avec $\mu(\eta) \to 0$ quand $\eta \to 0$ et t.q. $\|S(t)X - S(t+h)X\| \leq \mu(\eta)$ dès que t et $t + h$ sont dans $[0\,;1]$ avec $\|h| \leq \eta$: par suite pour tout $0 \leq s < 1$

$$\|M(s)X - X\| = \frac{1}{s}\int_0^s \|S(t)X - S(0)X\|\, dt \leq \mu(s)$$

Cela prouve que $M(s)X \to X$ quand $s \to 0$.

(iii) : Pour $s, t > 0$ et $X \in \mathscr{E}$, du fait que $(S(s) - I)/s$ appartient à $\mathcal{L}(\mathscr{E})$ il vient :

$$\begin{aligned}
\frac{S(s) - I}{s} M(t)X &= \frac{1}{t}\left(\frac{1}{s}\int_0^t S(s+u)X du - \frac{1}{s}\int_0^t S(u)X du\right) \\
&= \frac{1}{t}\left(\frac{1}{s}\int_s^{s+t} S(v)X dv - \frac{1}{s}\int_0^s S(u)X du - \frac{1}{s}\int_s^t S(u)X du\right) \\
&= \frac{1}{t}\left(\frac{1}{s}\int_t^{s+t} S(v)X dv - \frac{1}{s}\int_0^s S(u)X du\right) \\
&= \frac{1}{t}\left(\frac{1}{s}\int_0^s S(t+u)X du - \frac{1}{s}\int_0^s S(u)X du\right)
\end{aligned}$$

14. Nous utilisons ici le « *théorème de Heine* » généralisé qui affirme que si X un espace métrique compact et Y un espace métrique, alors toute application continue de X dans Y est uniformément continue.

soit encore

$$\frac{S(s)-I}{s}M(t)X = \frac{1}{t}\Big(S(t)M(s)X - M(s)\Big) = \frac{S(t)-I}{t}M(s)X$$

ce qui prouve la première identité de (29). La deuxième identité, se déduit en remarquant que pour tout $s, t > 0$, nous avons (par continuité de $S(t)$)

$$\begin{aligned} S(t)M(s)X &= S(t)\frac{1}{s}\int_0^s S(t)X du \\ &= \frac{1}{s}\int_0^s S(t+u)X du = \frac{1}{s}\int_0^s S(u)\big(S(t)X\big)du = M(s)S(t)X \end{aligned}$$

Cela termine la démonstration de la partie (iii).

(iv) : Pour tout $X \in \mathscr{E}$, d'après la partie (ii) et la première identité de (29) que

$$\lim_{s\to 0}\frac{S(s)-I}{s}M(t)X = \lim_{s\to 0}\frac{S(t)-I}{t}M(s)X = \frac{S(t)-I}{t}\lim_{s\to 0}M(s)X = \frac{S(t)-I}{t}X$$

ce qui signifie que $M(t)X \in \mathscr{D}$ avec $AM(t)X = (S(t)X - X)/t$. Maintenant si nous supposons en plus que $X \in \mathscr{D}$, alors utilisant la deuxième identité de (29) nous avons aussi

$$AM(t) = \lim_{s\to 0}\frac{S(s)-I}{s}M(t)X = \lim_{s\to 0}M(t)\frac{S(s)X-X}{s} = M(t)\lim_{s\to 0}\frac{S(s)X-X}{s} = M(t)AX$$

□

Preuve du Théorème 7.7. Soit $X \in \mathscr{E}$ arbitrairement donné ; alors d'après le Lemme 7.8 et la définition de $M(t)$ en (28) nous savons que $M(t)X$ appartient au domaine $\mathscr{D}$ du générateur infinitésimal A : la conclusion du fait que $\mathscr{D}$ est dense dans $\mathscr{E}$ s'obtient en notant que (c.f. Lemme 7.8) $M(t)X \to X$ quand $t \to 0$. Il s'agit maintenant de montrer que $(A, \mathscr{D})$ est fermé. Pour cela, considérons que $X_1, X_2, \dots$ est une suite de points de $\mathscr{D}$ tels que $X_n \to X$ avec $AX_n \to Y$ quand $n \to +\infty$ (X et Y étant a priori dans $\mathscr{E}$). Par une nouvelle application du Lemme 7.8 avec $X = X_n$, nous avons, pour tout n :

$$\frac{S(s)-I}{s}M(t)X_n = M(s)\frac{S(t)-I}{t}X_n$$

de sorte qu'en faisant tendre $t \to 0$, il vient

$$\frac{S(s)X_n - X_n}{s} = M(s)AX_n$$

soit encore avec $n \to +\infty$

$$\frac{S(s)X - X}{s} = M(s)Y \tag{30}$$

En faisant tendre cette fois $s \to 0$, nous avons $M(s)Y \to Y$ et par suite, d'après (30), nous pouvons conclure que $AX = Y$.

□

La définition suivante joue un rôle clef dans l'énoncer du théorème de Hille-Yosida.

Définition 7.9. *Le semi-groupe fortement continu $\{S(t)\ ;\ t \geq 0\} \subset \mathcal{L}(\mathscr{E})$ est un « semi-groupe fortement continu de contraction (s.g.f.c.c.) » dès que $\|S(t)\| \leq 1$, pour tout $t \geq 0$.*

Remarque 7.10. *Tout semi-groupe fortement continu $\{S(t)\ ;\ t \geq 0\} \subset \mathcal{L}(\mathscr{E})$ peut être associé à un s.g.f.c.c.. En effet (c.f. Lemme 7.5), si $C > 0$ et $\omega \geq 0$ sont des constantes t.q. $\|S(t)\| \leq Ce^{\omega t}$, pour tout $t \geq 0$, alors $\{\hat{S}(t) = e^{-\omega t}S(t)\ ;\ t \geq 0\}$ est encore un semi-groupe fortement continu ; de plus, si $(A, \mathscr{D})$ est le générateur infinitésimal de $S(t)$, alors pour tout $X \in \mathscr{D}$,*

$$\frac{1}{t}\Big(\hat{S}(t)X - X\Big) = \frac{1}{t}\Big(S(t)(e^{-\omega t}X) - (e^{-t\omega}X)\Big) + \frac{1}{t}\Big(e^{-\omega t} - 1\Big)X$$

on en déduit que le générateur infinitésimal de $t \mapsto \hat{S}(t)$ est $(A - \omega I, \mathscr{D})$. Maintenant si nous notons $\|\cdot\|$ la norme de $\mathscr{E}$ alors $X \mapsto \|X\|_C = 1/C\|X\|$ est une norme sur $\mathscr{E}$ qui est équivalente à la norme initiale de $\mathscr{E}$. Or, sur l'espace de Banach $(\mathscr{E}, \|\cdot\|_C)$ nous avons

$$\|\hat{S}(t)\|_C = \frac{1}{C}\|e^{-\omega t}S(t)\| \leq \frac{\|S(t)\|}{Ce^{\omega t}} \leq 1$$

Dans la suite nous noterons $\mathbb{H}$ le demi-plan de Poincaré, i.e. le sous-ensemble de $\mathbb{C}$ formé des z t.q. $\mathfrak{im}(z) > 0$; le théorème suivant fait partie du *« théorème de Hille-Yosida »* :

Théorème 7.11. *Soit $(A, \mathscr{D})$ le générateur infinitésimal d'un s.g.f.c.c. $\{S(t)\ ;\ t \geq 0\} \subset \mathcal{L}(\mathscr{E})$; alors l'ensemble résolvant $\Omega(A, \mathscr{D})$ contient $\boldsymbol{i}\mathbb{H}$ et de plus, pour tout $\lambda \in \boldsymbol{i}\mathbb{H}$ la résolvante $R_\lambda = (A - \lambda I)^{-1}$ vérifie $\|R_\lambda\| \leq 1/\mathfrak{re}(\lambda)$: en particulier*

$$\forall X \in \mathscr{D},\ \forall \lambda > 0, \quad \|AX - \lambda X\| \geq \lambda\|X\|$$

Preuve. Soit $(A, \mathscr{D})$ le générateur infinitésimal d'un semi-groupe continu de contraction $\{S(t)\ ;\ t \geq 0\}$ et soit $\lambda = a + \boldsymbol{i}b \in \mathbb{C}$ tel que $a = \mathfrak{re}(\lambda) > 0$. Comme par hypothèse $S(t)$ est une contraction, nous avons

$$\|e^{-\lambda t}S(t)X\| \leq e^{-at}\|X\| \tag{31}$$

Comme $a > 0$, l'inégalité (31) assure la condition de domination de $e^{-\lambda t}S(t)X$ sur l'intervalle $[0\,;+\infty[$, ce qui nous autorise à définir pour tout $X \in \mathscr{E}$:

$$\mathcal{R}_\lambda X := -\int_0^{+\infty} e^{-\lambda t}S(t)X\,dt, \tag{32}$$

de sorte que $\mathcal{R}_\lambda$ est un endomorphisme linéaire de $\mathscr{E}$ borné, puisque (31) entraîne que

$$\|\mathcal{R}_\lambda X\| \leq \int_0^{+\infty} e^{-at}\|X\|dt = \|X\|/a \tag{33}$$

Par ailleurs, pour tout $h > 0$, nous obtenons d'une part (grâce à la continuité de $S(h)$)

$$\begin{aligned}\left(\frac{S(h) - I}{h}\right)\mathcal{R}_\lambda X &= \frac{1}{h}\left(-\int_0^{+\infty} e^{-\lambda t}S(h+t)X\,dt + \int_0^{\infty} e^{-\lambda t}S(t)X\,dt\right)\\ &= \frac{1}{h}\left(-\int_h^{+\infty} e^{-\lambda(u-h)}S(u)X\,du + \int_h^{+\infty} e^{-\lambda t}S(t)X\,dt - \int_0^h e^{-\lambda t}S(t)X\,dt\right)\end{aligned}$$

soit encore,

$$\left(\frac{S(h) - I}{h}\right)\mathcal{R}_\lambda X = -\frac{e^{\lambda h} - 1}{h}\int_h^{\infty} e^{-\lambda t}S(t)X\,dt + \frac{1}{h}\int_0^h e^{-\lambda t}S(t)X\,dt \tag{34}$$

En faisant tendre $h \to 0$ il vient (en conclusion de cette première étape)

$$\forall X \in \mathscr{E}, \quad \mathcal{R}_\lambda X \in \mathscr{D} \quad \text{et} \quad A\mathcal{R}_\lambda X = \lambda\mathcal{R}_\lambda X + X \tag{35}$$

Mais d'autre part, nous pouvons aussi écrire (grâce à la continuité de $S(h)$)

$$\left(\frac{S(h)-I}{h}\right)\mathcal{R}_\lambda X = -\int_0^{+\infty} e^{-\lambda t}\left(\frac{S(h+t)-S(t)}{h}\right)X\,dt$$

c'est-à-dire

$$\left(\frac{S(h)-I}{h}\right)\mathcal{R}_\lambda X = \mathcal{R}_\lambda\left(\frac{S(h)-I}{h}\right)X$$

Par suite, étant donné $X \in \mathscr{D}$ nous avons $A\mathcal{R}_\lambda X = \mathcal{R}_\lambda AX$: dans ce cas, (35) devient

(36) $$\forall X \in \mathscr{D}, \quad \mathcal{R}_\lambda AX = \lambda\mathcal{R}_\lambda X + X$$

Nous savons d'après (33) que $\mathcal{R}_\lambda$ et un endomorphisme de $\mathcal{L}(\mathscr{E})$ t.q.

(37) $$\|\mathcal{R}_\lambda\| \le \frac{1}{a} = \frac{1}{\mathfrak{re}(\lambda)}$$

Par (35) il vient $\mathcal{R}_\lambda(\mathscr{E}) \subset \mathscr{D}$ avec $(A-\lambda I)\mathcal{R}_\lambda X = X$ pour tout $X \in \mathscr{E}$; enfin, (36) assurant que $\mathcal{R}_\lambda(A-\lambda I)X = X$, pour tout $X \in \mathscr{D}$, nous pouvons conclure que $A-\lambda I$ est un élément inversible de $L(\mathscr{D},\mathscr{E})$ dont l'inverse est $\mathcal{R}_\lambda$. Nous avons démontré que $R_\lambda = \mathcal{R}_\lambda$: lorsque $\lambda > 0$, nous avons $a = \lambda$ et de (37) nous tirons que $\|R_\lambda\| \le 1/\lambda$; par suite, pour tout $X \in \mathscr{D}$

$$\|X\| = \|R_\lambda(A-\lambda I)X\| \le 1/\lambda\|AX-\lambda X\|$$

Ceci achève la démonstration du théorème.

□

Le Théorème 7.11 nous amène à poser la définition suivante.

Définition 7.12. *Nous dirons qu'un opérateur $(A,\mathscr{D})$ densément défini sur l'espace de Banach $\mathscr{E}$ est « dissipatif » si et seulement si*

$$\forall X \in \mathscr{D},\ \forall\lambda > 0,\ \|AX-\lambda X\| \ge \lambda\|X\|$$

Si de plus tout $\lambda > 0$ appartient à $\Omega(A,\mathscr{D})$, alors l'opérateur $(A,\mathscr{D})$ est dit [15] *« m-dissipative ».*

Remarque 7.13. *Notons que si $(A,\mathscr{D})$ est dissipatif sur l'espace de Banach $\mathscr{E}$, alors pour tout $\lambda > 0$, nous avons $(A-\lambda I)(X) = 0$ entraîne que $\lambda\|X\| = 0$, i.e. $X = 0$: cela signifie que $A-\lambda I$ est une application injective de $L(\mathscr{D},\mathscr{E})$.*

Le théorème suivant simplifie grandement la vérification de la propriété de dissipation des opérateurs densément définis

Théorème 7.14. *Soit $(A,\mathscr{D})$ un opérateur densément défini sur $\mathscr{E}$ et dissipatif : alors les trois propositions suivantes sont équivalentes :*
- (i) *: $(A,\mathscr{D})$ est m-dissipatif ;*
- (ii) *: $\forall\lambda > 0,\ \forall Y \in \mathscr{E},\ \exists X \in \mathscr{D},\ AX-\lambda X = Y$;*
- (iii) *: $\exists\lambda_0 > 0,\ \forall Y \in \mathscr{E},\ \exists X \in \mathscr{D},\ AX-\lambda_0 X = Y$.*

Preuve. L'équivalence $(i) \iff (ii)$ correspondant à la définition de la m-dissipativité. L'implication $(ii) \Longrightarrow (iii)$ étant immédiate, il nous reste à démontrer que $(iii) \Longrightarrow (ii)$. Supposons donc l'existence de $\lambda_0 > 0$ tel que pour tout $Y \in \mathscr{E}$, l'équation $AX-\lambda_0 X = Y$ admet une solution X (nécessairement unique) dans $\mathscr{D}$. L'application $R_{\lambda_0} := (A-\lambda_0 I)^{-1}$

15. Le « *m* » de m-dissipatif est pour « maximal »

est un endomorphisme de $\mathscr{E}$ t.q. $R_{\lambda_0}(\mathscr{E}) \subset \mathscr{D}$: ainsi, pour tout $X \in \mathscr{E}$, nous avons $R_{\lambda_0}X \in \mathscr{D}$ et nous pouvons utiliser la dissipativité de $(A, \mathscr{D})$ pour écrire que

$$\|R_{\lambda_0}X\| \leq \frac{1}{\lambda_0}\|(A - \lambda_0 I)R_{\lambda_0}X\| = \frac{1}{\lambda_0}\|X\|$$

En d'autre termes, $R_{\lambda_0} \in \mathcal{L}(\mathscr{E})$ avec $\|R_{\lambda_0}\| \leq 1/\lambda_0$. Pour $\lambda > 0$ et $Y \in \mathscr{E}$ arbitraire donnés, $AX - \lambda X = Y$ ssi $AX - \lambda_0 X = Y + (\lambda - \lambda_0)X$, d'où $X = R_{\lambda_0}\big(Y + (\lambda - \lambda_0)X\big) =: B(X)$. Alors l'application linéaire $X \mapsto B(X)$ (dépendant de λ et Y) est bornée sur $\mathscr{E}$ et vérifie :

$$\|B(X_1) - B(X_2)\| \leq \|R_{\lambda_0}\| \cdot |\lambda - \lambda_0| \cdot \|X_1 - X_2\| \leq \frac{|\lambda - \lambda_0|}{\lambda_0}\|X_1 - X_2\|$$

Si $0 < \lambda \leq (3/2)\lambda_0$, alors B est une contraction linéaire de $\mathscr{E}$: par suite, (théorème du point fixe de Banach-Picard), il existe $X \in \mathscr{E}$ t.q. $BX = X$: pour tout $Y \in \mathscr{E}$ et tout $0 < \lambda \leq (3/2)\lambda_0$, il existe donc $X \in \mathscr{D}$ t.q.

$$Y + (\lambda - \lambda_0)X = (A - \lambda_0 I)BX = AX - \lambda_0 X$$

c'est-à-dire t.q. $Y = AX - \lambda X$. Nous pouvons réitérer cet argument en remplaçant successivement λ_0 par $(3/2)\lambda_0, (3/2)^2\lambda_0, \ldots$ de sorte que l'équation $\lambda X - AX = Y$ possède une solution X, pour tout $0 < \lambda \leq (3/2)^n\lambda_0$ et pour tout $n \geq 1$, i.e. pour tout $\lambda > 0$.

□

La notion m-dissipativité, permet de combiner les Théorèmes 7.7 et 7.11 comme suit :

Théorème 7.15. *Soit $(A, \mathscr{D})$ le générateur infinitésimal d'un semi-groupe continu de contractions $\{S(t)\ ;\ t \geq 0\} \subset \mathcal{L}(\mathscr{E})$; alors $(A, \mathscr{D})$ est densément défini, fermé et m-dissipatif.*

Le Théorème de Hille-Yosida affirme que la réciproque est aussi vraie.

Théorème 7.16 (Hille-Yosida)**.** *Soit $\mathscr{D}$ un sous-espace de l'espace de Banach $\mathscr{E}$ et A un élément de $L(\mathscr{D}, \mathscr{E})$; alors on a équivalence entre les deux propositions suivantes, soient :*
(i) : $(A, \mathscr{D})$ est un opérateur densément défini sur $\mathscr{E}$, fermé et m-dissipatif ;
(ii) : $(A, \mathscr{D})$ est le générateur infinitésimal d'un semi-groupe continu de contraction.

Preuve. Voir [Bre87, Théorème VII.8, p. 116].

□

7.4. Nous allons maintenant spécialiser le contexte du théorème de Hille-Yosida au cas particulier (très important pour la mécanique quantique) des espaces de Hilbert [16]. Pour cela, considérons que $(A, \mathscr{D})$ un opérateur densément défini sur l'espace de Hilbert $\mathscr{H}$; alors il est facile de vérifier que $(A, \mathscr{D})$ est dissipatif ssi $\mathfrak{re}\,\langle AX|X\rangle \leq 0$, pour tout $X \in \mathscr{D}$. En effet, supposons que $(A, \mathscr{D})$ soit dissipatif ; alors pour tout $X \in \mathscr{D}$ et tout $\lambda > 0$

$$\lambda^2\|X\|^2 \leq \|AX - \lambda X\|^2 = \|AX\|^2 + \lambda^2\|X\|^2 - 2\lambda\,\mathfrak{re}\,\langle AX|X\rangle$$

ce qui donne

$$\mathfrak{re}\,\langle AX|X\rangle \leq \frac{\|AX\|^2}{2\lambda}$$

16. C'est un cas spécial du sous cas des espaces de Banach réflexifs que nous n'abordons pas ici.

et $\mathfrak{re}\langle AX|X\rangle \le 0$ en faisant tendre $\lambda \to +\infty$. Réciproquement, si $\mathfrak{re}\langle AX|X\rangle \le 0$ pour tout $X \in \mathscr{D}$, alors pour tout $X \in \mathscr{D}$, l'inégalité de Cauchy-Schwarz assure que pour tout $\lambda > 0$

$$\|AX - \lambda X\| \cdot \|X\| \ge \mathfrak{re}\langle \lambda X - AX|X\rangle = \lambda\|X\|^2 - \mathfrak{re}\langle AX|X\rangle \ge \lambda\|X\|^2$$

cela signifie bien que $(A, \mathscr{D})$ est dissipatif [17].

Proposition 7.17. *$(A, \mathscr{D})$ densément défini sur un espace de Hilbert $\mathscr{H}$ est dissipatif ssi*

$$(\forall X \in \mathscr{D}) \quad \mathfrak{re}\langle AX|X\rangle \le 0$$

Théorème 7.18. *Soit $(A, \mathscr{D})$ un opérateur densément défini fermé et dissipatif sur un espace de Hilbert $\mathscr{H}$. Si $(A^\star, \mathscr{D}^\star)$ est aussi dissipatif alors $(A, \mathscr{D})$ est m-dissipatif.*

Preuve. D'après le Théorème 7.14 nous devons vérifier que $1 \in \Omega(A, \mathscr{D})$, c'est-à-dire que $A - I$ est une application linéaire bijective de $\mathscr{D}$ sur $\mathscr{H}$ dont l'inverse est un endomorphisme borné de $\mathscr{H}$. Notons $\mathscr{I} := \mathrm{Im}(A - I, \mathscr{D})$; alors la dissipativité de $(A, \mathscr{D})$ assure (c.f. Remarque 7.13) que $A-I$ est une application bijective de $L(\mathscr{D}, \mathscr{I})$; de plus, si $R := (A - I)^{-1}$, alors pour tout $(X, Y) \in \mathscr{D} \times \mathscr{I}$ t.q. $AX - X = Y$ nous avons

$$\|RY\| = \|X\| \le \|AX - X\| = \|Y\| \tag{38}$$

Le théorème sera donc établi si nous démontrons que $\mathscr{I} = \mathscr{H}$: pour cela nous allons vérifier que $\mathscr{I}$ est un sous-espace dense et fermé de $\mathscr{H}$. La densité de $\mathscr{I}$ dans $\mathscr{H}$ s'obtient à partir de la dissipativité de $(A^\star, \mathscr{D}^\star)$: celle-ci assure (de même que pour $A - I$) que $A^\star - I$ est un élément injectif de $L(\mathscr{D}^\star, \mathscr{H})$, d'où nous tirons que $\mathrm{Ker}(A^\star - I, \mathscr{D}^\star) = \{0\}$; la densité de $\mathscr{I}$ découle alors du fait (c.f. Proposition 3.5) que $\mathscr{I}^\perp = \mathrm{Im}(A - I, \mathscr{D})^\perp = \mathrm{Ker}(A^\star - I, \mathscr{D}^\star)$. Par suite, le théorème sera démontré si nous pouvons vérifier que $\mathscr{I}$ est un sous-espace fermé de $\mathscr{H}$. Pour voir cela, montrons d'abord que $(A - I, \mathscr{D})$ est un opérateur fermé dans $\mathscr{H}$: en effet, si $(X_1, Y_1), (X_2, Y_2), \ldots$ est une suite de points de $\mathscr{D} \times \mathscr{H}$ qui convergent vers (X, Y) dans $\mathscr{H} \times \mathscr{H}$ et t.q. $Y_n = AX_n - X_n$, alors $AX_n = X_n + Y_n$ et la fermeture de $(A, \mathscr{D})$ assure alors que $X \in \mathscr{D}$ avec $AX = X - Y$, soit encore que $(A_I)X = Y$: le graphe de $(A - I, \mathscr{D})$ est donc bien fermé dans $\mathscr{H} \times \mathscr{H}$. D'autre part, partant du fait que $(R, \mathscr{I})$ est un opérateur densément défini sur $\mathscr{H}$, il découle de (38) qu'il existe un endomorphisme de $\mathscr{H}$ borné que nous notons $\tilde{R}$, de norme $\|\tilde{R}\| \le 1$ et prolongeant R à tout $\mathscr{H}$, en ce sens que $\tilde{R}Y = RY$, pour tout $Y \in \mathscr{I}$. Nous pouvons maintenant terminer la démonstration en montrant que $\mathscr{I}$ est fermé. En effet, si les $Y_1, Y_2, \ldots$ forment une suite de Cauchy de points de $\mathscr{I}$, alors d'une part les $X_n = RY_n = \tilde{R}Y_n$ sont des points de $\mathscr{D}$ (puisque $R(\mathscr{I}) = \mathscr{D}$) et d'autre part, la continuité de $\tilde{R}$ entraîne qu'ils forment suite de Cauchy : si nous notons Y et X les limites respectives des Y_n et des X_n dans $\mathscr{H}$, alors la fermeture de $(A - I, \mathscr{D})$ assure que $(X, Y) \in \mathscr{D} \times \mathscr{I}$, avec $X = RY$: nous pouvons ainsi conclure que $\mathscr{I}$ est bien un sous-espace fermé de $\mathscr{H}$, ce qu'il fallait démontrer.

□

Le théorème suivant est une partie du *« théorème de Lumer et Phillips »*.

17. Dans le cas d'un espace de Hilbert (ou d'un espace de Banach réflexif), il existe une autre terminologie : ainsi $(A, \mathscr{D})$ est dit *« accrétif »* ou *« monotone »* si $\mathfrak{re}\langle AX|X\rangle \ge 0$, pour tout $X \in \mathscr{D}$: en d'autres termes $(A, \mathscr{D})$ est *« accrétif »* ssi $(-A, \mathscr{D})$ est *« dissipatif »*. Si de plus la résolvante $\Omega(A, \mathscr{D})$ contient l'ensemble des $\lambda < 0$ alors $(A, \mathscr{D})$ est dit *« m-accrétif »*, ou encore *« maximal monotone »* (c.f. [Bre87, Chap. VII, p. 101]). Par suite, $(A, \mathscr{D})$ est dissipatif (resp. m-dissipatif) ssi $(-A, \mathscr{D})$ est accrétif (resp. m-accrétif).

Théorème 7.19. *Si* $(A, \mathscr{D})$ *est densément défini fermé et dissipatif sur un espace de Hilbert* $\mathscr{H}$ *et si* $(A^\star, \mathscr{D}^\star)$ *est aussi dissipatif, alors* $(A, \mathscr{D})$ *est le générateur infinitésimal d'un s.g.f.c.c. sur* $\mathscr{H}$.

Preuve. D'après le Théorème de Hille-Yosida et le Théorème 7.18.

□

En vue des applications à la mécanique quantique, nous allons voir une partie du *« théorème de Stone »*, dans une version « *semi-groupe* ». Le résultat en question (c.f. Théorème 7.20) peut se déduire comme corollaire du théorème de Hille-Yosida (historiquement c'est le théorème de Stone qui est à l'origine du théorème de Hille-Yosida). D'après le Théorème 7.4, nous obtenons ainsi le cadre abstrait de la résolution de l'équation de Schrödinger. Notons que dans le cas où le générateur infinitésimal $(A, \mathscr{D})$ du semi-groupe $\{S(t) ; t \geq 0\} \subset \mathcal{L}(\mathscr{E})$ est antisymétrique, en ce sens que $(A, \mathscr{D})^\star = (-A, \mathscr{D})$, il est immédiat de vérifier que l'opérateur $(\boldsymbol{i}A, \mathscr{D})$ est autoadjoint, puisqu'en effet, pour $X, Y \in \mathscr{D}$ nous avons (le produit scalaire est anti-linéaire à droite)

$$\langle \boldsymbol{i}AX|Y\rangle = \boldsymbol{i}\langle AX|Y\rangle = -\boldsymbol{i}\langle X|AY\rangle = \langle X|\boldsymbol{i}AY\rangle$$

Théorème 7.20. *Si* $(A, \mathscr{D})$ *est un opérateur autoadjoint sur l'espace de Hilbert* $\mathscr{H}$, *alors* $(\boldsymbol{i}A, \mathscr{D})$ *est le générateur infinitésimal d'un s.g.f.c.c sur* $\mathscr{H}$.

Preuve. L'opérateur $(A, \mathscr{D})$ étant autoadjoint, c'est par définition un opérateur densément défini qui est nécessairement fermé (c.f. Proposition 3.12). Pour tout $X \in \mathscr{D}$, nous pouvons utiliser la symétrie de $(A, \mathscr{D})$ de sorte que $\langle AX|X\rangle = \langle X|AX\rangle = \overline{\langle AX|X\rangle}$: par suite $\langle AX|X\rangle \in \mathbb{R}$ et donc $\mathfrak{re}\langle \varepsilon \boldsymbol{i}AX|X\rangle = 0$ pour $\varepsilon = \pm 1$: nous venons de vérifier que $(\varepsilon \boldsymbol{i}A, \mathscr{D})$ est un opérateur fermé et dissipatif. La conclusion vient du Théorème 7.19, du fait que l'adjoint de $(\boldsymbol{i}A, \mathscr{D})$ est égal à $(-\boldsymbol{i}A, \mathscr{D})$.

□

Nous verrons dans [Oli16b] que la description quantique d'un système mécanique se réalise par la donnée d'un espace de Hilbert $\mathscr{H}$ ainsi que d'un hamiltonien quantique $(H, \mathscr{G})$ densément défini dans $\mathscr{H}$ et autoadjoint. Un chemin $t \mapsto \Psi(t)$ sur la sphère unité de $\mathscr{G}$ est une évolution quantique possible du système mécanique étudié ssi *« l'équation de Schrödinger »* est satisfaite le long de ce chemin en ce sens que

$$\boldsymbol{i}\hbar \frac{d\Psi}{dt} = H\Psi(t) \tag{39}$$

Ici $\hbar$ est la fameuse constante de Planck égale au quantum d'action. Notons que la forme *« optimale »* du Théorème 7.20, amène naturellement à l'introduction du facteur multiplicatif « $\boldsymbol{i}$ » faisant de $(-\boldsymbol{i}H/\hbar, \mathscr{G})$ le générateur infinitésimal d'un semi-groupe continu de contractions sur $\mathscr{H}$: si (par analogie avec le cas borné) nous notons $\{\exp(\boldsymbol{i}H/\hbar) ; t \geq 0\}$ ce semi-groupe, alors les solutions de (39) sont formellement de la forme $\Psi(t) = \exp(\boldsymbol{i}tH/\hbar)\Psi_0$. Le Théorème 7.20 donne ainsi une justification intéressante de la présence du nombre « $\boldsymbol{i}$ » dans l'équation de Schrödinger.

Remarque 7.21. *Le théorème de Stone original traite de la caractérisation des sous-groupe unitaire à un paramètre* $\{U_t ; t \in \mathbb{R}\} \subset \mathcal{L}(\mathscr{H})$ *qui sont fortement continus : en particulier* $\|U_t\| = 1$. *Dans le Théorème 7.20, nous considérons des semi-groupes* $\{S(t) ; t \geq 0\}$ *fortement continu et formé de contractions, de sorte que* $\|S(t)\| \leq 1$.

8. Appendice A : propriétés spectrales de certains opérateurs bornés

8.1. Un endomorphisme A borné, agissant sur un espace de Hilbert séparable $\mathscr{H}$ est dit *« normal »* si $A^\star A = AA^\star$ (i.e. $[A, A^\star] = 0$). Les isométries (et donc les opérateurs unitaires), ainsi que les opérateurs autoadjoints sont tous des opérateurs normaux.

Proposition 8.1. *Soit $A \in \mathcal{L}(\mathscr{H})$ un opérateur normal ; alors*

(i) : pour tout $X \in \mathscr{H}$ et tout $\xi \in \mathbb{C}$: $AX = \xi X$ ssi $A^\star X = \overline{\xi} X$; de plus, si $X, Y \in \mathscr{H}$ sont deux vecteurs propres respectivement associés à ξ et η, alors $\langle X|Y\rangle = 0$ dès que $\xi \neq \eta$;

(ii) : si A autoadjoint (resp. positif), alors $\mathrm{Vp}(A) \subset \mathbb{R}$ (resp. $\mathrm{Vp}(A) \subset [0\,;+\infty]$) ;

(iii) : si A est isométrique (pas nécessairement surjective), alors $\mathrm{Vp}(A)$ est un sous-ensemble (possiblement vide) du 1-tore $\mathbb{T} = \{|\cdot| = 1\}$ et $\mathrm{Vp}(A^\star) \cap \mathbb{T} = \overline{\mathrm{Vp}(A)}$;

$(iii)'$: si A est unitaire (i.e. isométrique et surjectif), alors $\mathrm{Vp}(A) \subset \mathbb{T}$ et $\mathrm{Vp}(A^\star) = \overline{\mathrm{Vp}(A)}$.

Preuve. (i) : Si A est normal, alors $B := \xi - A$ est aussi normal ; si $BX = 0$ alors

$$0 = \|BX\|^2 = \langle BX|BX\rangle = \langle B^\star BX|X\rangle = \langle BB^\star X|X\rangle = \|B^\star X\|^2$$

et par suite $AX = \xi X$ ssi $A^\star X = \bar{\xi} X$. Si $AX = \xi X$ et $AY = \eta Y$ alors $\xi \langle X|Y\rangle = \langle AX|Y\rangle = \langle X|A^\star Y\rangle = \overline{\overline{\eta}} \langle X|Y\rangle = \eta \langle X|Y\rangle$ et donc $\xi \neq \eta$ implique $\langle X|Y\rangle = 0$.

(ii) : Soit X un vecteur propre de la valeur propre ξ de A. Si A est hermitien, alors

$$\xi\|X\|^2 = \langle AX|X\rangle = \langle X|AX\rangle = \overline{\xi}\|X\|^2$$

Cela entraîne que $\xi = \overline{\xi}$ i.e. $\xi \in \mathbb{R}$. Si A est positif, alors $\xi\|X\|^2 = \langle AX|X\rangle \geq 0$ et donc $\xi \geq 0$.

(iii)-$(iii)'$: Nous démontrons un résultat un peu plus général : soit A un opérateur *« spectraloïde »* de norme $\|A\| = 1$, en ce sens qu'il existe $X \in \mathscr{H}$ avec $\|X\| = 1$ et qui soit un vecteur propre de A associé à une valeur propre $\xi \in \mathbb{T}$. Alors (c.f. Proposition 3.1) nous avons $\|A\|^2 = \|A^\star A\| = \|AA^\star\| = \|A^\star\|^2$: par suite $\|A\| = \|A^\star\| = 1$ et donc :

$$1 = |\xi| = |\langle AX|X\rangle| = |\langle X|A^\star X\rangle| \leq \|X\| \cdot \|A^\star X\| \leq \|X\| \cdot \|A^\star\| \cdot \|X\| = 1$$

de sorte que $|\langle X|A^\star X\rangle| = \|X\| \cdot \|A^\star X\|$: par le cas d'égalité de l'inégalité de Cauchy-Schwartz, il est nécessaire que X et $A^\star X$ soient proportionnels, i.e. $A^\star X = \zeta X$ pour un scalaire $\zeta \in \mathbb{C}$; alors, $\overline{\zeta} = \langle X|A^\star X\rangle = \langle AX|X\rangle = \xi$: nous avons démontré que $A^\star X = \overline{\xi} X$. En appliquant le même argument à $A^\star$ nous déduisons que

$$\mathrm{Vp}(A^\star) \cap \mathbb{T} = \overline{\mathrm{Vp}(A)} \cap \mathbb{T} \tag{40}$$

Quand A est isométrique, nous savons d'une part que $\|A\| = 1$; mais d'autre part, l'identité $AX = \xi X$ (pour $\|X\| = 1$) entraîne que $1 = \langle AX|AX\rangle = |\xi|^2$, d'où l'inclusion $\mathrm{Vp}(A) \subset \mathbb{T}$: par suite, d'après (40) nous avons $\mathrm{Vp}(A^\star) \cap \mathbb{T} = \overline{\mathrm{Vp}(A)}$. Si A est unitaire, alors A et $A^\star$ sont (en particulier) isométriques : par une nouvelle application de (40) nous pouvons conclure que $\mathrm{Vp}(A^\star) = \overline{\mathrm{Vp}(A)}$.

□

8.2. Nous avons noté dans la Proposition 8.1 que lorsque A est une isométrie non surjective d'un espace de Hilbert, il n'existe pas de relations simples entre $\mathrm{Vp}(A)$ et $\mathrm{Vp}(A^\star)$ autre que $\mathrm{Vp}(A) \cap \mathbb{T} = \overline{\mathrm{Vp}(A^\star)} \cap \mathbb{T}$. Pour illustrer cela, nous allons regarder de près les *« opérateur de Koopman-von Neumann »* intervenant dans l'approche spectrale de la classification des *« systèmes dynamiques non inversibles »*. Supposons que $T : X \to X$ soit une transformation continue définie sur l'espace métrique compact X et soit $\mu(dx)$ une mesure de probabilité borélienne sur X supposée T-invariante, en ce sens que $T_*\mu := \mu\circ T^{-1}$ (obtenue par « *push-forward* ») coïncide avec μ (i.e. $\mu = \mu\circ T^{-1}$) : on dit que (X, T, μ) est « *un système dynamique mesuré (s.d.m.)* ». Notons $\eta = f\mu$ la probabilité absolument continue par rapport à μ où $f : X \to [0; +\infty[$ est une fonction de l'espace $L^1(\mu)$ des (classes de) fonctions boréliennes (à valeur complexe) μ-intégrable t.q. $\int f(x)\mu(dx) = 1$; alors, il est immédiat que $T_*\eta = T_*(f\mu) = f\circ T\mu$. Maintenant si nous considérons que f est un élément quelconque de $L^1(\mu)$, alors la probabilité μ étant T-invariante, nous avons $\int |f\circ T(x)|\mu(dx) = \int |f(x)|\mu(dx)$: nous venons de vérifier que $f \mapsto U_T f = f\circ T$ est un endomorphisme de $L^1(\mu)$. D'autre part, l'espace de Hilbert $L^2(\mu)$ est un sous-espace de $L^1(\mu)$ dont il est facile de vérifier qu'il est laissé stable sous l'action de U_T. Par définition « *l'opérateur de Koopman-von Neumann* » est la restriction de U_T (toujours notée U_T) à l'espace $L^2(\mu)$: il est immédiat que $U_T : L^2(\mu) \to L^2(\mu)$ est une isométrie puisque pour tout $f, g \in L^2(\mu)$ la T-invariance de μ assure que

$$\langle U_T f | U_T g\rangle := \int f(T(x))\overline{g(T(x))}\mu(dx) = \int f(x)\overline{g(x)}\mu(dx) = \langle f|g\rangle$$

Le 1-tore $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ est un groupe compact pour l'addition $(x, y) \mapsto x + y \pmod 1$ et nous notons $\lambda(dx) = dx$ « *la mesure de Haar* »[18]. Si T_2 désigne la multiplication par 2 modulo 1 (i.e. $T_2(x) = 2x \pmod 1$), alors il est facile de voir que $(\mathbb{T}, T_2, \lambda)$ est un s.d.m. : on montre que $\mathrm{Vp}(U_{T_2}) = \{1\}$ (cela est dû au fait que le s.d.m. $(\mathbb{T}, T_2, \lambda)$ est « *fortement mélangeant* »). Dans ce cas particulier (simple et important), il est facile d'identifier l'adjoint U_T^* associé à l'opérateur de Koopman-von Neumann ; en effet, pour tout $f, g \in L^2(\lambda)$,

$$\begin{aligned}\langle f|U_{T_2} g\rangle = \int f(x)\overline{g\circ T(x)}dx &= \int_0^{1/2} f(x)\overline{g(2x)}dx + \int_{1/2}^1 f(x)\overline{g(2x-1)}dx \\ &= \int_0^1 \frac{1}{2} f\left(\frac{y}{2}\right)\overline{g(y)}dy + \int_0^1 \frac{1}{2} f\left(\frac{x+1}{2}\right)\overline{g(y)}dy\end{aligned}$$

ce qui nous donne

$$U_{T_2}^* f = \frac{1}{2} f\left(\frac{y}{2}\right) + \frac{1}{2} f\left(\frac{x+1}{2}\right) \tag{41}$$

On peut alors montrer (assez facilement) que $\mathrm{Vp}(U_{T^2}^*) \supset D^\circ \cup 1$, où $D^\circ := \{|\cdot| < 1\}$ est le disque unité ouvert du plan complexe. (L'adjoint $U_{T_2}^* : L^2(\mu) \to L^2(\mu)$ en (41) est traditionnellement appelé « *l'opérateur de Perron-Frobenius* ».) Cette remarque sur les propriétés spectrales des isométries non surjectives (i.e. non unitaires) d'un espace de Hilbert et lié à la question de la brisure de symétrie temporelle observée dans les systèmes dynamiques irréversibles, c'est-à-dire avec le « *Théorème H de Boltzmann* ».

8.3. Mentionnons que le « *théorème spectral* » pour les opérateurs normaux dans $\mathcal{L}(\mathscr{H})$ affirme qu'un tel opérateur A est unitairement conjugué à un opérateur de multiplication : plus précisément, cela signifie qu'il existe une mesure positive et sigma-fini[19] sur $\mathbb{R}$, que nous notons μ, ainsi qu'une fonction $f : \mathbb{R} \to \mathbb{C}$ qui soit μ-essentiellement bornée sur $\mathbb{R}$ t.q. $A = U^{-1}M_f U$, où M_f est l'opérateur de multiplication $g \mapsto M_f g = fg$ définie sur

18. L'unique mesure (de Radon) de probabilité t.q. $\lambda(A) = \lambda(A + x \pmod 1)$, pour tout $x \in \mathbb{T}$ et tout borélien A de $\mathbb{T}$.

19. Cette condition assure que l'espace de Hilbert $L^2_\mu(\mathbb{R})$ est séparable.

l'espace de Hilbert $L^2_\mu(\mathbb{R})$ et où $U : \mathscr{H} \to L^2_\mu(\mathbb{R})$ est une application unitaire (i.e. transforme une base hilbertienne de $\mathscr{H}$ en une base hilbertienne de $L^2_\mu(\mathbb{R})$). On peut alors démontrer que le spectre de M_f coïncide avec les valeurs essentielles de f (i.e. les valeurs $z \in \mathbb{C}$ prises par f et telles que $\mu\big(f^{-1}(V)\big) > 0$ soit strictement positif, pour un voisinage V de z dans $\mathbb{C}$. On peut aussi démontrer que chacune de ces valeurs spectrales sont des valeurs propres approchées de l'opérateur M_f (sans être nécessairement des valeurs propres : c.f. *« critère de Weyl »* et Remarque 3.18 supra) ; en admettant ce résultat, il est facile d'établir la proposition suivante.

Proposition 8.2. *Pour $A \in \mathcal{L}(\mathscr{H})$, les propositions suivantes sont satisfaites : (i) : $\Sigma(A) \subset \{z \in \mathbb{C} \,;\, |z| \leq \|A\|\}$; (ii) : si A est autoadjoint alors $\Sigma(A) \subset [-\|A\| \,;\, \|A\|]$; $(ii)'$: si A est positif alors $\Sigma(A) \subset [0 \,;\, \|A\|]$; (iii) : si A est unitaire alors $\Sigma(A) \subset \{z \in \mathbb{C} \,;\, |z| = 1\}$.*

Preuve. (i) : Si $|\xi| > \|A\|$, alors $\|1/\xi A\| < 1$ et donc (formule d'inversion de Von Neumann) $I - 1/\xi A$ est inversible dans $\mathcal{L}(\mathscr{H})$ d'inverse $B := \sum_{k=0}^{\infty}(1/\xi A)^k$; il est alors immédiat que $\xi I - A$ est inversible dans $\mathcal{L}(\mathscr{H})$ d'inverse $1/\xi B$ et donc $\xi \in \Omega(A)$.

(ii)-$(ii)'$: Soit $A \in \mathcal{L}(\mathscr{H})$ autoajoint et soit ξ une valeur propre spectrale de A, i.e. une valeur propre approchée. Alors, il existe une suite $X_1, X_2, \ldots$ de points de la sphère unité de $\mathscr{H}$ t.q. $\|AX_n - \xi X_n\| \to 0$ quand $n \to +\infty$. Or pour tout rang n

$$\langle AX_n - \xi X_n | X_n\rangle + \xi = \langle AX_n | X_n\rangle = \langle X_n | AX_n\rangle = \langle X_n | AX_n - \xi X_n\rangle + \overline{\xi}$$

et comme (inégalité de Cauchy-Schwarz) $|\langle AX_n - \xi X_n | X_n\rangle| \leq \|AX_n - \xi X_n\|$, nous en déduisons (dans la limite où $n \to 0$) que $\xi = \overline{\xi}$: cela prouve que $\xi \in \mathbb{R}$. Lorsque A est positif, alors pour tout rang n, nous avons

$$0 \leq \langle AX_n | X_n\rangle = \langle AX_n - \xi X_n | X_n\rangle + \xi\|X_n\|^2 = \langle AX_n - \xi X_n | X_n\rangle + \xi$$

ce qui entraîne que $\xi \geq 0$, lorsque $n \to +\infty$

(iii) : Soit $A \in \mathcal{L}(\mathscr{H})$ unitaire et soit ξ une valeur propre spectrale de A, i.e. une valeur propre approchée (*« critère de Weyl »*). Si $X_1, X_2, \ldots$ est toujours une suite de points de la sphère unité de $\mathscr{H}$ t.q. $\|AX_n - \xi X_n\| \to 0$ quand $n \to +\infty$, il vient pour tout rang n :

$$\begin{aligned} 1 = \|X_n\| = \langle AX_n | AX_n\rangle &= \langle AX_n - \xi X_n | AX_n\rangle + \xi \langle X_n | AX_n\rangle \\ &= \langle AX_n - \xi X_n | AX_n\rangle + \xi \langle X_n | AX_n - \xi X_n\rangle + \xi\overline{\xi} \end{aligned}$$

L'inégalité $|\langle AX_n - \xi X_n | X_n\rangle| \leq \|AX_n - \xi X_n\|$, nous permet alors de conclure (dans la limite où $n \to 0$) que $1 = \xi\overline{\xi}$, soit encore que $|\xi| = 1$.

□

8.4. D'une manière générale, si A est un endomorphisme borné d'un espace de Banach $\mathscr{E}$, *« le rayon spectral »* de A est le supremum $\rho(A)$ des modules des éléments du spectre $\Sigma(A)$, de sorte que l'inégalité $\rho(A) \leq \|A\|$ est toujours valide (c.f. Proposition 8.2) ; notons aussi que la sous-multiplicativité de la norme opérateur sur $\mathcal{L}(\mathscr{E})$ justifie l'existence de la limite de la *« formule de Guelfand »* donnant le rayon spectral, soit :

$$\rho(A) := \lim_{n\to+\infty} \sqrt[n]{\|A^n\|} = \inf\left\{\sqrt[n]{\|A^n\|} \,;\, n \geq 1\right\}. \tag{42}$$

Il découle de la formule d'inversion de Von Neumann que pour tout opérateur A borné sur un espace de Banach, le module $|\xi|$ d'une valeur spectrale est majoré par la norme opérateur $\|A\|$. Cependant, le rayon spectral de A, i.e. le supremum $\rho(A)$ des $|\xi|$ pour ξ décrivant $\Sigma(A)$, ne coïncide pas (en général) avec $\|A\|$ (même en dimension finie, on peut considérer le cas des opérateurs nilpotents). Cette remarque nous amène à introduire la classe des opérateurs *« spectraloïdes »*, c'est-à-dire les opérateurs A t.q. $\rho(A) = \|A\|$. La proposition suivante montre que les opérateurs normaux sont spectraloïdes.

Théorème 8.3. *Un opérateur $A \in \mathcal{L}(\mathscr{H})$ normal est « spectraloïde » en ce sens que sa norme coïncide avec sont rayon spectral : en d'autres termes $\|A\| = \max\{|\xi| \,;\, \xi \in \Sigma(A)\} = \rho(A)$.*

Preuve. Soit $X_1, X_2, \ldots$ une suite de point de la sphère unité de $\mathscr{H}$ (i.e. $\|X_i\| = 1$) et telle que $|\langle AX_n|X_n\rangle| \to \|A\|$; alors, quitte à extraire une sous-suite, nous pouvons supposer que $\langle AX_n|X_n\rangle$ converge vers $\xi \in \mathbb{C}$ t.q. $|\xi| = \|A\|$. Alors ξ est une valeur propre approchée (ce qui est suffisant à assurer que $\|A\| = \rho(A)$) : en effet,

$$\begin{aligned} \|(A - \xi I)X_n\|^2 &= \langle AX_n - \xi X_n | AX_n - \xi X_n\rangle \\ &= \|AX_n\|^2 + \|\xi X_n\|^2 - 2\mathfrak{re}(\bar{\xi}\langle AX_n|X_n\rangle) \le 2\|A\|^2 - 2\mathfrak{re}(\bar{\xi}\langle AX_n|X_n\rangle) \end{aligned}$$

la conclusion venant du fait que $\bar{\xi}\langle AX_n|X_n\rangle$ tend vers $\bar{\xi}\xi = |\xi|^2 = \|A\|^2$ quand $n \to +\infty$.

□

9. Appendice B : Opérateurs compacts

9.1. Soit $(\mathscr{E}, \|\cdot\|)$ un espace de Banach ; nous notons $\mathcal{K}(\mathscr{E})$ l'ensemble des opérateurs A de $\mathcal{L}(\mathscr{E})$ qui sont compacts en ce sens l'image $A\{\|\cdot\| \le 1\}$ de la boule unité de $\mathscr{E}$ est relativement compact (i.e possède une adhérence compacte dans $\mathscr{E}$) ; de manière équivalente, cela signifie que pour toute suite $X_0, X_1, \ldots$ de $\mathscr{E}$ bornée, il existe une sous-suite $X_{n_0}, X_{n_1}, \ldots$ t.q. la suite $AX_{n_0}, AX_{n_1}, \ldots$ soit de Cauchy (et donc convergente dans $\mathscr{E}$). Il est facile de vérifier que $\mathcal{K}(\mathscr{E})$ est un *« sous-idéal fermé »* de l'algèbre de Banach $(\mathcal{L}(\mathscr{E}); +, \cdot)$ (munie de la norme opérateur), en ce sens que $\mathcal{K}(\mathscr{E})$ est une sous-espace vectoriel fermé de $(\mathcal{L}(\mathscr{E}), \|\cdot\|)$ absorbant, i.e. vérifiant $\mathcal{L}(\mathscr{E})\mathcal{K}(\mathscr{E}) = \mathcal{K}(\mathscr{E})\mathcal{L}(\mathscr{E}) = \mathcal{K}(\mathscr{E})$.

Théorème 9.1. *Etant donné $\mathscr{E}$ un espace de Banach, $A \in \mathcal{K}(\mathscr{E})$ et $\Sigma(A)$ le spectre de A, les propositions suivantes sont satisfaites : (i) : toute valeur spectrale non nulle (i.e. $\xi \in \Sigma(A) \setminus \{0\}$ est une valeur propre de A (i.e. $\xi \in \mathrm{Vp}(A)$) et sa multiplicité est finie (i.e. il existe un entier $\nu(\xi_i) = \nu_i$ minimal – la multiplicité de ξ_i – t.q $\mathrm{Ker}(\xi_i I - A)^{\nu_i} = \mathrm{Ker}(\xi_i I - A)^{\nu_i+1}$) ; (iii) : 0 est le seul point d'accumulation possible de $\Sigma(A)$ et si $\mathscr{E}$ n'est pas de dimension finie, alors $0 \in \Sigma(A)$.*

9.2. Considérons par exemple que $\mathscr{E} = \mathscr{C}[0\,;1]$ (muni de la norme infini $\|\cdot\|_\infty$) : alors l'opérateur $A \in \mathcal{L}(\mathscr{E})$ t.q. $Af(x) = \int_0^x f(t)dt$ est compact. En effect, considérons que $f_0, f_1, \ldots$ est une suite bornée de la boule unité de $\mathscr{E}$; alors, par une application du *« théorème d'Arzelà-Ascoli »*, l'existence d'une sous-suite de $Af_{n_0}, Af_{n_1}, \ldots$ qui soit de Cauchy est assurée, du moment que nous sommes capables de montrer que $\{Af_n\}_n$ est un sous ensemble équicontinu de $\mathscr{C}[0\,;1]$. Mais ceci est effectivement vrai puisque pour tout n et tout $0 \le x < y \le 1$ nous avons :

$$|Af_n(y) - Af_n(x)| \le \int_x^y |f_n(t)|dt \le |x - y|$$

9.3. Dans le même esprit nous pouvons considérer un *« opérateurs à noyau »* $B : f \mapsto Bf$ de $\mathcal{L}(\mathscr{E})$ défini en posant $Bf(x) = \int_0^1 K(x,t)f(t)dt$ et où la fonction $K : [0;1]\times[0;1] \to \mathbb{R}$ (appelée le noyau de B) est supposée continue [20]. Afin de montrer la compacité de B, considérons une suite $f_0, f_1, \ldots$ est une suite bornée de la boule unité de $\mathscr{E}$. Du fait que $K(x,t)$ est continu sur un compact, alors (uniforme continuité) pour tout $\varepsilon > 0$ il existe $\eta > 0$ t.q. $\max\{|x-x'|, |t-t'|\} \le \eta$ entraîne $|K(x,t) - K(x',t')| \le \varepsilon$. Ainsi, pour tout rang n, nous avons

$$|x-y| \le \eta \Rightarrow |Af_n(y) - Af_n(x)| \le \int_0^1 |K(x,t) - K(y,t)| \cdot |f_n(t)|dt \le \varepsilon$$

Cela signifie que $\{Bf_n\}_n$ est un sous-ensemble équicontinu de $\mathscr{C}[0;1]$: nous en concluons grâce au *« théorème d'Arzelà-Ascoli »* que B est un opérateur compact de $\mathcal{L}(\mathscr{E})$.

9.4. A partir de maintenant nous considérons le cas où l'espace de Banach est un espace de Hilbert (séparable) que nous notons (comme d'habitude) $\mathscr{H}$. Nous savons déjà que $\mathcal{K}(\mathscr{H})$ est un idéal fermé de l'algèbre de Banach $\mathcal{L}(\mathscr{H})$; on peut aussi montrer que $\mathcal{K}(\mathscr{H})$ est laissé invariant par la $\star$-involution [21]. Le théorème suivant est un cas spécial (important) du *« théorème spectral »*.

Théorème 9.2. (i) *: Si $A \in \mathcal{K}(\mathscr{H})$ est normal, alors : $A = 0$ ssi $\Sigma(A) = \{0\}$ et pour $A \neq 0$, il existe une famille orthonormée $\{X_i\}_{k=0}^N$ (avec $1 \le N \le +\infty$) t.q.*

$$A = \textstyle\sum_{i=0}^N \xi_i \langle X_i|\cdot\rangle\, X_i = \sum_{i=0}^N \xi_i\, |X_i\rangle\, X_i. \tag{43}$$

où $(\xi_i)_{i=0}^N$ est la suite des valeurs propres non nulles de A comptée avec leur ordre de multiplicité. (L'expression en (43) est appelée une « décomposition diagonale » de A.)

(ii) *: si $A \in \mathcal{L}(\mathscr{H})$ est normal et possède une décomposition diagonale de la forme (43), où $(X_i)_{k=0}^N$ est une famille orthonormée et où $(\xi_i)_{i=0}^N$ est une suite de nombres complexes non nuls tendant vers 0 lorsque $N = +\infty$, alors $A \in \mathcal{K}(\mathscr{H})$.*

Preuve. Voir [Bre87, § VI.4, Proposition VI.9, Corollaire VI.10 & Théorème VI.11, . p. 96 – 98].

□

10. Appendice C : Théorèmes d'analyse fonctionnelle

Nous commençons par le *« théorème de la borne uniforme »* (*« théorème de Banach-Steinhaus »*).

Théorème 10.1 (de la borne uniforme)**.** *Soit $\mathcal{T} = \{T_i\ ;\ i \in I\}$ une famille quelconque d'applications linéaires bornées d'un espace de Banach E à valeurs dans un espaces vectoriel normé F. Si la famille $\mathcal{T}$ est ponctuellement bornée (i.e., $\sup\{\|T_iX\|\ ;\ i \in I\} < +\infty$ pour tout $X \in E$), alors $\mathcal{T}$ est bornée pour la norme opérateur (i.e., $\sup\{\|T_i\|\ ;\ i \in I\} < +\infty$).*

Nous allons donner une démonstration très simple de ce théorème due à Sokal [Sok11]

Lemme 10.2. *Soient E et F deux espaces vectoriels normés et soit T une application linéaire bornée définie de E dans F. Alors, pour tout $X \in E$ et tout réel $r > 0$, nous avons*

$$\sup\left\{\|TY\|\ ;\ \|X - Y\| < r\right\} \ge \|T\|r$$

20. L'opérateur A donné dans le premier exemple est un opérateur à noyau avec $K(x,t)\mathbf{1}_{[0;x]}(t)$: mais ce noyau n'est pas continu.

21. Cela fait de $\mathcal{K}(\mathscr{H})$ une C*-algèbre d'opérateurs de $\mathcal{L}(\mathscr{H})$

Preuve. Etant donné $Z \in E$, nous avons

$$\max\left\{\|T(X+Z)\|, \|T(X-Z)\|\right\} \geq \frac{1}{2}\Big(\|T(X+Z)\| + \|T(X-Z)\|\Big) \geq \frac{1}{2}\|TZ\|$$

où la dernière minoration est obtenue par l'inégalité triangulaire sous la forme $\|\alpha - \beta\| \leq \|\alpha\| + \|\beta\|$. La conclusion vient en prenant le supremum sur les Z t.q. $\|Z\| < r$.

□

Preuve du Théorème 10.1. Supposons que $\sup\{\|T_i\| \; ; \; i \in I\} = +\infty$ et considérons $T_0, T_1, \ldots$ une suite de $\mathcal{T}$ t.q. $\|T_n\| \geq 4^n$. Si nous posons $X_0 = 0$, alors pour tout $n \geq 1$ le Lemme 10.2 garantit l'existence d'un $X_n \in E$ tel que $\|X_n - X_{n-1}\| < 1/3^n$ et $\|T_n X_n\| \geq 1/(2 \cdot 3^n)\|T_n\|$. La suite $X_0, X_1, \ldots$ est de Cauchy, ce qui assure l'existence d'une limite $X \in E$; il est immédiat de vérifier que $\|X - X_n\| \leq 1/(2 \cdot 3^n)$ de sorte que $\|T_n X\| \geq 1/3^n \|T_n\| \geq 1/6\,(4/3)^n \to +\infty$

□

Théorème 10.3 (de l'application ouverte)**.** *Soient E et F deux espaces de Banach et $T \in L(E, F)$; si l'application T est surjective, alors elle est nécessairement ouverte.*

Preuve. Voir [Bre87, § II.3, Théorème II.5, p. 18 – 20].

□

Si E et F sont deux espaces de Banach et si T est une bijection linéaire continue de E sur F, alors (théorème de l'application ouverte) T est ouverte ; si $U \subset E$ est ouvert alors $(T^{-1})^{-1}(U) = T(U)$ est ouvert et donc T^{-1} est continue.

Corollaire 10.4 (Théorème de l'isomorphisme)**.** *Si E et F sont deux espaces de Banach, alors l'inverse d'une bijection linéaire continue de E sur F est nécessairement elle même continue.*

Corollaire 10.5 (Théorème du graphe fermé)**.** *Soient E et F deux espaces de Banach et $T \in L(E, F)$. Alors T est continue si et seulement si le graphe de T est fermée dans $E \times F$, c'est-à-dire : pour toute suite $X_1, X_2, \ldots$ de E t.q. $X_n \to X$ dans E et si $TX_n \to Y$ dans F, alors $Y = TX$.*

Preuve. L'application T étant linéaire, le graphe $\Gamma = \{(X, TX) \; ; \; X \in E\}$ est un sous-espace vectoriel de $E \times F$. Si T est continue, alors Γ est fermé dans $E \times F$. Pour la réciproque, notons que E et F étant des espaces de Banach, $E \times F$ est aussi un espace de Banach : par suite, si Γ est fermé, alors c'est également un espace de Banach. Soient P et Q les projections (suivant le produit cartésien) de $E \times F$ sur E et F respectivement. L'application P est continue et sa restriction à Γ est une bijection linéaire continue de Γ sur E : sa réciproque R est donc continue (c.f. Théorème de l'isomorphisme), ce qui assure que $T = QR$ est continue.

□

11. Appendice D : Intégrale de Bochner (d'après Wikipédia)

Soit $(X, \mathscr{T}, \mu)$ un espace mesuré et $(\mathscr{E}, \|\cdot\|)$ un espace de Banach. Une fonction simple $s : X \to \mathscr{E}$ est définie de sorte que pour tout $x \in X$:

$$s(x) = \sum_{i=1}^{n} \mathbf{1}_{E_i}(x) b_i$$

où les $E_1, \ldots, E_n$ (resp. $b_1, \ldots, b_n$) sont des éléments deux à deux disjoints de $\mathscr{T}$ (resp. de $\mathscr{E}$). Si $\mu(E_i)$ est finie quel que soit $b_i \neq 0$, alors la fonction simple est Bochner-intégrable et son intégrale est par définition :

$$\int s\, d\mu = \int \left[\sum_{i=1}^{n} \mathbf{1}_{E_i}(x) b_i \right] \mu(dx) = \sum_{i=1}^{n} \mu(E_i) b_i$$

(c'est l'analogue exact de la définition dans la théorie de Lebesgue). Une fonction $f : X \to \mathscr{E}$ est dite *« fortement mesurable »* si elle est la limite μ-presque partout d'une suite de fonctions simples. Une fonction fortement mesurable $f : X \to \mathscr{E}$ est *« Bochner-intégrable »* s'il existe une suite de fonctions simples intégrables s_n telle que :

$$\lim_{n \to \infty} \int \|f - s_n\|\, d\mu = 0,$$

(où l'intégrale dans le membre de gauche est une intégrale *ordinaire* de Lebesgue) ou de manière équivalent ssi : $\int \|f\|\, d\mu < \infty$. Dans ce cas, l'intégrale de Bochner est définie par :

$$\int f\, d\mu = \lim_{n \to \infty} \int s_n\, d\mu$$

Si T est un opérateur linéaire continu et f est intégrable au sens de Bochner, alors Tf est intégrable au sens de Bochner et l'intégration et T peuvent être échangés :

$$\int Tf d\mu = T \int f d\mu$$

Ce résultat est aussi vrai pour des opérateurs fermés à condition que Tf soit aussi intégrable. Une version du *« théorème de convergence dominée »* s'applique à l'intégrale de Bochner. Spécifiquement, si $f_n : X \to \mathscr{E}$ est une suite de fonctions mesurables sur un espace mesuré complet qui converge presque partout vers une fonction limite f et s'il existe $g \in L^1(\mu)$ telle que $\|f_n(x)\| \leq g(x)$, pour presque tout $x \in X$, alors

$$\lim_{n \to +\infty} \int \|f - f_n\|\, d\mu = 0 \quad \text{et} \quad \lim_{n \to +\infty} \int_E f_n\, d\mu = \int_E f\, d\mu \quad (\forall E \in \mathscr{T})$$

Références

[Bre87] H. Brezis. *Analyse fonctionnelle : théorie et applications*. Éditions Masson - Paris, 1987.

[GW90] C. Gasquet and P. Witomski. *Analyse de Fourier et applications : filtrage, calcul numérique, ondelettes*. Masson, 1990.

[Oli16a] E. Olivier. Mécanique quantique I : bases mathématiques. *Bull. d'Inf. App. & App.*, 103 :11–72, 2016.

[Oli16b] E. Olivier. Mécanique quantique III : de de Broglie à Schrödinger. *Bull. d'Inf. App. & App.*, 105 :171–209, 2016.

[RSN55] F. Riesz and B. SZ.-Nagy. *Leçons d'analyse fonctionnelle*. Gauthiers-Villars, 1955.

[Sok11] A. D. Sokal. A Really Simple Elementary Proof of the Uniform Boundedness Theorem. *The Am. Math. Month.*, 118(5) :450–452, 2011.

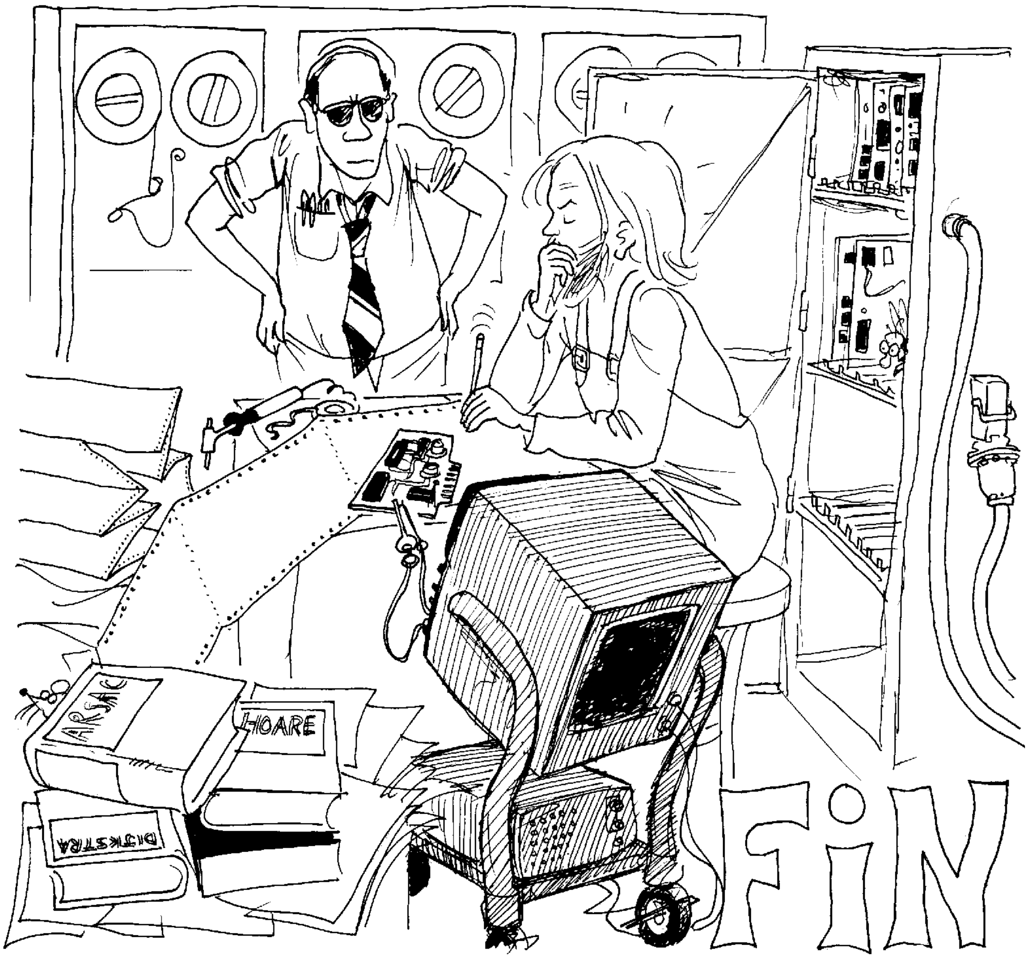
[ST89] M. Samuelides and L. Touzillier. *Analyse fonctionnelle*. Éditions Cépaduès, 1989.

[Sto30] M. H. Stone. Linear transformations in Hilbert space, III : operational methods and group theory. *Proc. Nat. Acad. Sci. U.S.A.*, 16 :172–175, 1930.

[Sto32] M. H. Stone. On one-parameter unitary groups in Hilbert space. *Annals of Mathematics*, 33 (3) :643–648, 1932.

[VN55] J. Von Neumann. *Mathematical Foundations of Quantum Mechanics*. Investigations in physics. Princeton University Press, 1955.

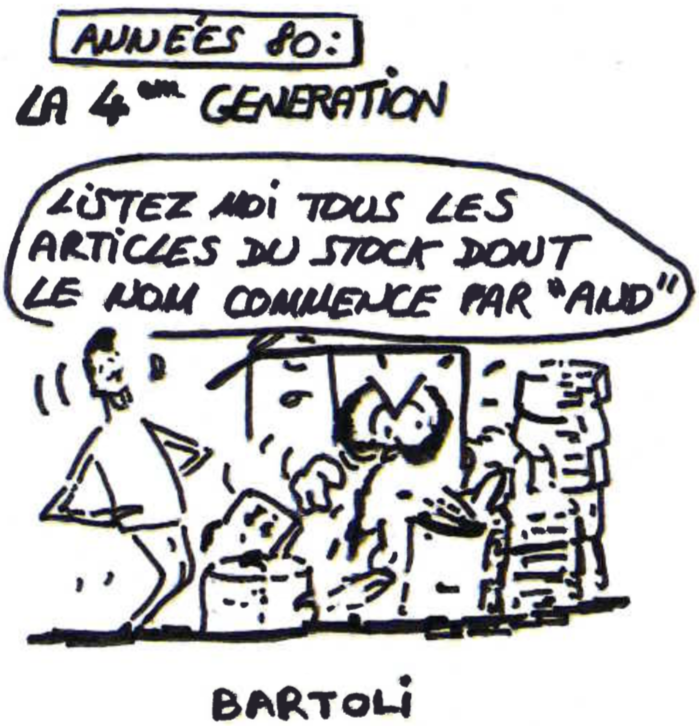
Mais depuis ce jour l'ordinateur du centre a des pannes inexplicables, auxquelles aucun spécialiste n'a su porter remède. Peut-être est-ce la chaussure d'Anselme Lanturlu qui est restée coincée quelque part...



BIAA

Bulletin d'Informatique Approfondie & Applications
Computation - Information
Volumes 2016

103 - 104 - 105



Aix Marseille université

Publication trimestrielle, gratuite, de l'Université d'Aix Marseille
Dépôt légal : décembre 2016 - http://www.biaa.eu - ISSN 0291 - 5413

Couverture : dessin de Michel Avezard (Zevar) interprété par la MMIAGe (1985-1987)

(2016) B.I.A.A. no **105** BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
ISSN 0291-5413 (Université d'Aix-Marseille)

ÉDITORIAL : Un professeur de l'université de Moncton s'inquiète pour sa famille qui vit en Syrie

Jalal ALMHANA, Simon DELATTRE

Résumé. – (N.D.L.R.) : D'origine syrienne, Jalal Almhana est professeur d'informatique à l'Université de Moncton au Canada depuis 26 ans. Il angoisse pour sa famille restée au pays, dévasté par la violence. Alors que ses deux filles sont en sécurité à Moncton, Jalal Almahana craint pour ses proches qui demeurent au cœur du conflit syrien. Nous rapportons un échange entre Jalal Almhana et Simon Delattre journaliste à l'Acadie Nouvelle qui est paru le 8 septembre 2015 dans sa version numérique. Ce journal assure aussi un tirage papier quotidien depuis 1984.

Acadie Nouvelle : Comment avez-vous réagi en voyant la photo du jeune migrant retrouvé sans vie ?

Jalal Almhana : C'est la réaction d'un être humain, c'est triste de voir un enfant mort au bord de la plage. En tant que Syrien ça m'attriste encore plus. Je me sens impuissant, je ne peux rien faire devant cette situation. L'image parle elle-même. C'est un enfant innocent. Si ça réveille les consciences, tant mieux. Mais il y a des horreurs qui ont été commises qui vont au-delà de ce que vous pensez. C'est la pointe de l'iceberg en termes de violence, de tuerie ... Il y a en tous les jours ! On est dans la loi du Moyen-Âge. C'est incroyable ce que l'être humain est capable de faire. Cette image est choquante, mais il y en a mille autres.

Acadie Nouvelle : Vous avez toujours de la famille là-bas ? Comment ça se passe pour eux ?

Jalal Almhana : Toute ma famille pratiquement est en Syrie. J'ai un frère, une sœur, des neveux, des cousins et cousines. On s'inquiète tous les jours ... Chaque appel qui vient de la Syrie peut annoncer un décès. On suit la situation presque tous les jours par Skype, quand on peut parce qu'il y a un rationnement de l'électricité. Quand je peux les joindre parce qu'ils se déplacent entre six appartements.

Acadie Nouvelle : Comment vivent-ils la guerre au quotidien ?

Jalal Almhana : C'est vraiment épeurant. On n'était pas habitué à ça, ma sœur a eu des tirs à quelques mètres de sa maison. Elle était vraiment effrayée, elle ne pouvait pas dormir. Maintenant, ils ont toujours peur de mourir, mais ils essaient de se protéger. On a une maison au sud de Damas, on n'y va plus parce qu'on a reçu des menaces et il y a eu des prises d'otages ... On a eu toutes sortes de problèmes, il n'y a plus de sécurité parce que l'armée est occupée. Il y a deux ans, des gens sont venus chez nous et ont tout pris, l'or, la voiture,... Ils se trouvent dans un quartier chrétien de Damas, à la portée des tirs de la rébellion. La roquette peut arriver n'importe où.

Acadie Nouvelle : Les civils sont autant menacés que les militaires ...

Jalal Almhana : Oui ... J'ai un cousin qui sortait de son travail et il a reçu une balle dans la tête ... C'est une situation aléatoire.

Acadie Nouvelle : Est-ce qu'on devrait faire plus pour accueillir les réfugiés ?

Jalal Almhana : Bien sûr ! Il y a des camps de réfugiés en Turquie et en Jordanie. Quatre millions de déplacés. Mais les gens, les chrétiens ne veulent pas aller dans ces camps, qui sont infestés par les islamistes. Alors ils donnent toutes leurs économies à des passeurs qui prennent environ 11 000 euros, espérant trouver un terrain où on ne reçoit pas de bombe sur la tête. Cette poussée d'immigration, c'est un cri de désespoir. On est passé de la recherche de la démocratie à une situation où on tente d'instaurer un État islamiste, les musulmans non plus ne veulent pas ça. Les États proposent de prendre 10 000 personnes, mais qu'est-ce que ça représente face à des millions ? Passer par les canaux officiels, c'est presque impossible.

Acadie Nouvelle : Il y a trop de règlements qui contrôlent l'entrée au Canada selon vous ?

Jalal Almhana : Mes proches ne remplissent pas les conditions. Généralement pour avoir le statut de réfugié, il faut être enregistré dans un camp. Mais ils refusent d'y aller. Le camp, c'est la prison, il y a des mafias. Aller à l'ambassade, ça ne marche pas non plus. Pour parrainer trois personnes, il faut 12 000 $. Le programme actuel n'est pas fait pour des gens désespérés. La réglementation est tellement impossible qu'ils forcent la porte. Les gens ont besoin de vivre, de bâtir. Il faut une approche réaliste devant la situation de guerre. Il faut ouvrir, on a une approche archaïque. Ma sœur était de passage aux États-Unis, elle n'a pas pu me rendre visite. Il faut arrêter l'hémorragie, donner un espoir de solution. En Syrie, les gens n'ont plus d'espoir.

Jalal Almhana

(2016) B.I.A.A. no **105** BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
ISSN 0291-5413 (Université d'Aix-Marseille)

Science et Science-Fiction. Séminaire H.S.I.F.S. paru dans La Recherche. N°49 – octobre 1974

Jean-Marie SOURIAU [1]

Résumé. – (N.D.L.R.) En juin 1974, notre collègue Jean-Marie Souriau avait fait un exposé dans notre séminaire de l'époque le H.S.I.F.S. : « *Séminaire d'Histoire et Sociologie des Idées et des Faits Scientifiques* ». Cet exposé a trouvé sa place dans le cahier numéro 9 (épuisé) du Séminaire. La partie introductive de la conférence-débat, que nous avons réédité dans le n° 104 du *Bulletin* a servi de *brouillon* pour sa parution dans la revue La Recherche d'octobre 1974 au numéro 49. Nous publions ici le manuscrit original de Souriau, qui contient certaines annotations ne figurant pas dans le numéro 49 de La Recherche. Jean-Marie Souriau, décédé en 2012, évoque notre séminaire à la page 859 de l'article de la revue La Recherche. Les numéros de ces années de La Recherche ne sont pas disponibles en ligne. Nous remercions Madame Sophie Coisne, rédactrice en chef de La Recherche, de nous avoir autorisé l'édition du document source des pages parues dans La Recherche à l'automne 1974.

Science et science-fiction ont fait un long chemin ensemble : aujourd'hui leurs voies divergent. Cette séparation est-elle définitive ?

Et d'abord, qu'est-ce que la science-fiction ? Possède-t-elle une existence autonome, ou s'agit-il d'un compromis entre des activités littéraire et scientifique ?

La science a-t-elle quelque chose à savoir de la science-fiction, en dehors du délassement du scientifique ?

Toutes ces questions ne peuvent guère se traiter abstraitement : une "définition" de la science-fiction est un exercice de style assez inutile ; qui d'ailleurs se risquerait à donner une *définition de la science* ? Nous avons affaire à deux données culturelles et historiques impliquant chacune un certain groupe d'individus, une certaine activité ; et évidemment un certain ésotérisme, une certaine *complicité culturelle*.

En effet, science et science-fiction ont aujourd'hui un aspect structurel commun très net : elles sont toutes deux *allusives* ; obligatoirement allusives même, nous semble-t-il, afin de sauvegarder chacune leur identité.

Un article ou un livre scientifique se doit de comporter une bibliographie abondante surabondante même de fait ; cette bibliographie est destinée évidemment à aider le lecteur, à laisser croire que l'auteur a tout lu, tout compris, tout retenu, à faire quelques politesses indispensables. Mais par-delà ces raisons avouées (ou presque), peut-être s'agit-il aussi de dire au lecteur : je suis des vôtres.

Un roman ou une nouvelle de science-fiction comporte rarement une bibliographie sauf imaginaire (des extraits d'une quelconque Encyclopédie Galactica, par exemple). Mais le plus souvent, transparaissent dans le texte même des allusions à la fois claires et discrètes à des œuvres,

1. Professeur de l'Université de Provence (1923-2012).

2. Un exemple remarquable est la citation par Ivan Efremov, écrivain soviétique (dans la nouvelle « *Cor Serpentis* ») d'un texte de l'américain Murray Leinster, alias Will Jenkis (« *Premier contact* »). Le sujet, exemplaire, est la rencontre de l'humanité avec une civilisation extérieure, mais analogue : et surtout les circonstances de cette rencontre au large, entre navires cosmiques : quelle attitude vont adopter les deux équipages ? Cette rencontre est-elle le début d'une guerre d'extermination (comme dans « *Assassinat des Etats-Unis* », de Will Jenkis), ou d'une symbiose ? Dans cette répétition de la rencontre des Espagnols et des Amérindiens, qui assumera le rôle des Conquistadors, qui celui des victimes ? Existe-t-il une solution différente, ne serait-ce que le double sacrifice des équipages pour faire échapper le reste de la vie à un avenir effrayant ? La citation de Jenkins par Efremov est détaillée et pour cause ; elle est l'occasion d'une confrontation serrée sur le plan éthique et politique.

des thèmes ou des situations de référence ; allusions qui réjouissent les amateurs, et qui permettent à l'auteur de se situer par rapport à telle lignée dont il se réclame ou qu'il conteste [2].

Le plus souvent, la référence se complète d'un *vocabulaire spécifique commun* qui, comme le jargon scientifique, est le plus souvent traduit de l'américain. En science-fiction, chacun sait distinguer un robot d'un androïde [3].

On sait que les *mutants* sont des hommes et des femmes doués de pouvoirs spéciaux au moins télépathes ; que l'on vit dans les fusées grâce à des *cultures hydroponiques* : que les *Grands Galactiques* nous surveillent ; que la monnaie interplanétaire s'appelle le *crédit* ; que notre univers s'insère parmi une foule d'*univers parallèles* : que lorsqu'on est pressé de se rendre aux confins de la galaxie, on met sa fusée en propulsion *hyperluminique* , ou plus simplement que l'on prend le raccourci de l'*hyperespace* (on peut aussi utiliser le *transmetteur de matière*, mais attention à ne pas oublier sa *clef laxienne* ...).

Sitôt débarqué, on se comporte comme tous les marins : sortant de son *croiseur* de combat, le *capitaine* comme l'*astrogateur* se retrouve dans les bouges d'un quelconque *astroport* martien ou *dénébien* [4] en compagnie de quelque créature de rêve, parfois "*étrangère*" mais toujours très fonctionnelle (malgré son origine *extra-terrestre*, son nom a parfois des consonances françaises, signe certain de sa perversité).

Autre remarque : la science-fiction, comme la science, ne véhicule pas une idéologie unique : même si les auteurs progressistes semblent aujourd'hui majoritaires, il existe, et depuis longtemps, de la science-fiction "de droite" et "même fascisante" [5] .

Toutes ces analogies de structure suggèrent que science et science-fiction appartiennent à un même courant culturel : ce que va confirmer une brève analyse historique.

Même si le mot date du XXème siècle, la science-fiction est en fait aussi ancienne que la "science" proprement dite : ainsi Platon nous raconte le récit de l'Atlantide fournissant un thème initial à des écrits innombrables.

Etudions à ce point de vue une période caractéristique, le XVIIème siècle. Au printemps 1600, sur le Campo di Fiori à Rome, Giordano Bruno expie ses hérésies sur le bûcher. Visionnaire inspiré, il soupçonnait une unité de structure sous les divers aspects de la matière ; il avait révélé, dans une œuvre grandiose (le « *Banquet de Cendres* », 1585) la double infinitude du monde : l'univers est illimité et infiniment divers : plus métaphysicien que mathématicien, il ouvrait cependant idéologiquement la voie à la science "moderne".

En 1604, Képler découvre une *supernova* dont l'existence détruit le dogme aristotélicien de l'immuabilité de la sphère céleste [6] et par là même l'existence de toutes ces "sphères" plus ou moins cristallines qui encombraient le ciel. C'est en 1609 qu'il publie ses deux premières *lois de la mécanique céleste* qui, complétées, lui permettront de calculer des éphémérides précises.

En 1610 paraît le « *Message céleste* » où Galilée annonce les découvertes qu'il vient de faire avec sa lunette : le Soleil a des tâches, Jupiter a des lunes, la Lune a des montagnes !

Ce "Message" est partout reçu dans l'enthousiasme ; ainsi, à Aix-en-Provence, Pierre Gassend, Joseph Gaultier, Nicolas Peiresc construisent aussitôt une lunette, découvrent la nébuleuse d'Orion, dressent la première carte

3. Alors que le "robot" est une pure mécanique, ainsi qu'en témoigne le langage courant ou technologique, "l'androïde" contient des composantes biologiques, ce qui le rend particulièrement apte à l'esclavage et à la prostitution. Et pourtant le créateur du mot robot, le tchèque Karel Capek, désignait sous ce nom une créature vivante.

4. Les noms arabes des étoiles ont beaucoup fait rêver, et un lecteur assidu les connaît bien, même s'il est incapable de les situer dans le ciel. La célèbre trilogie « *Dune* » de Franck Herbert (1963-65-69) comporte à la fin un riche glossaire de la « *planète de l'Epice* » ; il s'agit simplement d'arabe parlé.

5. Ainsi certain ouvrage des années 40 qui raconte la résistance de l'Amérique face à une invasion Jaune : certains scientifiques utilisent déjà toutes les ressources de leur art pour découvrir des armes nouvelles : grâce à leur connaissance du calcul tensoriel (sic) ils projettent dans le ciel d'immenses effigies de guerriers blancs qui effraient leurs adversaires bornés. Ils trouvent enfin l'arme absolue : un rayon qui coagule sélectivement les protéines asiatiques : la victoire n'est assombrie que par le décès de quelques Jaunes cuits pour la bonne cause.

6. Cette coïncidence tombe à merveille : depuis lors, on n'a plus observé aucune supernova galactique : et pourtant les statistiques nous permettent d'en espérer quelques unes par siècle. Quelle frustration pour les astronomes !

lunaire ; pour Gassend, qui se fera appeler Gassendi, c'est le début d'une vaste série d'observations et d'expériences [7] .

Képler lui-même manifeste son admiration pour les découvertes de Galilée ; et il rédige un récit de science-fiction « *Le Songe* » (en latin !), où il décrit le voyage dans la Lune, et même ses habitants [8]...

Gassendi réside à Paris à partir de 1625 (il sera nommé professeur de mathématiques au Collège de France). Il devient le maître à penser des "libertins", qui se placent sous l'égide de Giordano Bruno ; il reprend l'atomisme de Démocrite et d'Epicure aux dépens d'Aristote. Son influence sera aussi profonde dans les milieux scientifiques que chez les poètes et les écrivains qu'il fréquente : notamment Tristan l'Hermite, Molière, et Cyrano de Bergerac, auteur des « *Histoires comiques des états et empires de la Lune* » [9].

Lorsqu'on y lit qu'il suffit, pour quitter la Terre, de se laisser aspirer avec la rosée par les premiers rayons du soleil, on pourrait se croire bien loin de la science : mais le dessein de Cyrano n'est pas de décrire une technologie du voyage spatial [10], mais de nous faire comprendre la relativité des situations cosmiques : « *la Lune est un monde comme celui-ci, à qui nôtre Terre sert de Lune* » [11] ; la possibilité future d'un voyage vers la Lune se trouve inscrite dans cette connaissance nouvelle. Ce n'est pas de la "fantaisie" que l'on trouve dans l'œuvre de Cyrano, mais la tranquille assurance de celui qui assisté à la métamorphose du monde des idées, et qui milite pour faire partager cette libération de la pensée.

Même assurance chez Swift quand il rédige les « *Voyages de Gulliver* »(1726) ; dans le monde de Laputa, peuplé de Scientifiques aux mœurs à peine croyables, on a découvert deux satellites de Mars ; Swift indique les diamètres de leurs orbitres, leur durée de révolution, et fait remarquer qu'ils confirment les lois de Képler.

Quelle belle coïncidence si Mars a effectivement deux satellites, Phobos et Deimos, qui furent découverts 150 ans plus tard ! Mais peut-être n'y a-t-il chez Swift qu'un certain sens de la régularité : la Terre a un satellite, Jupiter quatre, on en connaissait déjà cinq à Saturne : il suffisait d'interpoler [12].

D'autres circonstances historiques et culturelles permettront au contraire à Jules Verne de devenir le poète de la technologie. Pourtant son œuvre transmet le même message : nous avons à portée de la main de quoi changer le monde et nous-mêmes ; mais peut-être ne sommes-nous pas capables de le faire.

Chacun connaît l'impact que l'œuvre de Verne a eu sur la technique du siècle qui a suivi ; c'est un « *Nautilus* » qui a percé le premier les glaces du Pôle, comme dans « *Vingt mille lieues sous les mers* » ; la conquête de la Lune semble presque une répétition de l'œuvre de Jules Verne ; dans « *De la Terre à la Lune* » (1865), on assiste à une violente rivalité entre le Texas et la Floride comme point de lancement, c'est finalement la Floride qui est choisie. Un siècle après, dirigé depuis le Texas, le départ est donné en Floride (le point choisi par Jules Verne est à une centaine de kilomètres du Cap Canaveral-Kennedy). Dans les deux cas, trois cosmonautes prennent le départ et reviennent

---

7. Le 7 novembre 1631, il observera le passage de Mercure devant le Soleil, prédiction posthume de Képler dont la confirmation consacrera définitivement la nouvelle mécanique céleste. En 1640, sur une galère devant le port de Marseille, il réalisera une expérience que Galilée plaçait dans un navire imaginaire : à la suite de quoi c'est lui et non Galilée qui formulera le principe de l'inertie (« *De motu impresso a motore translato* », 1642), principe qui sera repris 45 ans plus tard par Newton.

8. Voir l' « *Encyclopédie, de l'Utopie et de la Science-Fiction* » de Pierre Versins. Ed. l'âge d'Homme, Lausanne, 1972.

9. La première édition, posthume et censurée, est de 1657. Le « *Cyrano de Bergerac* » d'Edmond Rostand a tellement trahi son personnage que l'allusion qu'on y trouve au voyage dans la lune devient incompréhensible ; Cyrano n'était ni bretteur ni gascon ; il est l'auteur de « *La mort d'Agrippine* », tragédie dont on dit grand bien et d'une comédie, « *Le pédant joué* », à laquelle Molière a emprunté quelques scènes pour les « *Fourberies de Scapin* ».

10. Pourtant il propose aussi l'utilisation d'une fusée à étages ...

11. Voir « *l'Encyclopédie* » de Pierre Versins, citée plus haut.

12. En 1959, l'astronome soviétique Chklovski, se basant sur des évaluations subtiles, assurait que l'un des satellites était si léger qu'il ne pouvait qu'être creux, et par conséquent, artificiel ! Cette belle idée, déjà exprimée en science-fiction (« *Jupiter V* » d'Arthur Clarke, paru dans les années 50) a malheureusement succombé aux photos précises prises par Mariner : Phobos et Deimos sont d'honnêtes blocs de rochers, parsemés de petits cratères. A moins qu'il ne s'agisse d'un camouflage particulièrement réussi...

plonger grâce à leurs fusées de correction en plein océan, à proximité d'un navire : le canon de Verne s'appelait *Columbiad*, le module luminaire d'Armstrong *Columbia* et pas par hasard [13].

Cette conformité s'explique d'une part par certaines nécessités techniques que Jules Verne avait correctement anticipées [14] ; mais elle manifeste aussi l'influence étonnante qu'une œuvre de fiction a pu garder pendant un siècle sur la pensée et les rêves des scientifiques et des ingénieurs.

Réciprocité, donc, dans les échanges entre science et science-fiction, tout au cours de l'histoire. Y-a-t-il un changement fondamental au XXème siècle, et particulièrement dans le cas de la science-fiction américaine ? Il ne semble pas. Si l'on excepte certaines fortes personnalités, comme H.P. Lovecraft, par exemple, les débuts de la littérature d'anticipation américaine n'ont pas de spécificité particulière ; les premières revues spécialisées se placent résolument dans la tradition européenne et publient beaucoup de traductions [15].

C'est le public lui-même qui a permis à un genre traditionnellement mineur de devenir majeur aux U.S.A. Peut-être grâce à son éducation, qui lui a permis d'échapper à la scolastique des classifications littéraires. Qu'on se souvienne que Jules Verne, malgré l'influence probable qu'il a eue sur Arthur Rimbaud (« *Le bateau ivre* » semble inspiré par « *Vingt mille lieues sous les mers* ») est resté presque un siècle inconnu dans les manuels de littérature française. Sans doute estimait-on que Jules Verne écrivait pour les enfants, alors que les manuels formaient des hommes...

Quand on s'adresse au "public cultivé" français, il est nécessaire de tenir compte de son ignorance ; ainsi, dans la préface de son roman « *Un animal doué de raison* » (1967), Robert Merle a dû, fort honnêtement, prévenir les lecteurs que son thème, l'intelligence des dauphins, était classique [16].

Aux Etats-Unis, par contre, romans, nouvelles, bandes dessinées, cinéma, ont inclus la science-fiction dans l'univers culturel : il n'est pas rare de voir la Science-Fiction inscrite au cursus des étudiants de science, et même dans les académies militaires : analyse des ouvrages, discussions des thèmes et des anticipations sont considérées comme utiles à la formation générale des scientifiques [17].

Dans une telle ambiance, d'innombrables talents ont surgi ; des genres se sont créés. "L'Heroic fantasy" par exemple, héritière directe des romans de cap et d'épée, ne pose guère de problèmes, sinon de comprendre comment le talent de certains auteurs peut rendre passionnants des thèmes insipides chez la plupart des autres comme la guerre interplanétaire ou la science-fiction médiévale.

Le "space opera" a pu prendre parfois des dimensions réellement impressionnantes ; les premières lignes du « *Monde des Ã* » de Van Vogt, par exemple, posent de vastes problèmes métaphysiques : l'identité (le héros découvre rapidement que ses souvenirs sont "truqués"), la vie et la mort, les rapports entre la conscience et le réel (il suffit, pour se transporter instantanément en un lieu, de le "mémoriser" avec une précision suffisante). Peu importe, après tout, que le texte soit émaillé de niaiseries "sémantiques" empruntées à Korzybski (« *La carte n'est pas le territoire* »).

De même, « *Fondation* » d'Isaac Asimov commence dans de si larges perspectives socio-historiques, le récit commence après la chute de l'empire galactique (cadre déjà vermoulu de

13. L'un des trois cosmonautes de Jules Verne, Michel Ardan (« *Autour de la lune* », 1871) était Français : il s'agit de l'anagramme, et sans doute du portrait, de Nadar, célèbre photographe et aéronaute français (de son vrai nom Félix Tournachon). D'une façon générale, le vol de Jules Verne était un peu moins nationaliste que celui de la Nasa ; il était en particulier subventionné par une collecte internationale, occasion pour Verne d'exprimer des jugements sarcastiques sur l'intérêt que les diverses nations portent à la science.

14. Le Texas et la Floride sont les régions les plus méridionales des Etats-Unis, et la latitude d'un point de lancement ne doit pas être trop élevée. A ce titre, la base française de Kourou était bien choisie, même si elle n'a servi à rien.

15. « *Weird Tales* » (1923) ; « *Amazing Stories* » (1926), éditée par Hugo Gernsback, qui reproduit en frontispice la pierre tombale de Jules Verne à Amiens (voir le « *Panorama de la Science-Fiction* » de Jacques Van Herp (Ed. Gérard, Verviers, 1973).

16. Voir par exemple la « *Guerre des salamandres* » de Karel Capek (1935), et une foule d'écrits américains, dont la « *Voix des dauphins* » du physicien Leo Szilard (1961).

17. A notre connaissance, les seules études de science-fiction proposées aux étudiants français sont de type "littéraire".

trop d'écrits antérieurs...) que l'auteur a bien du mal à se dépêtrer de l'écheveau gigantesque dans lequel il s'est lui-même fourré. Pourtant on n'oubliera pas la gloire de Trantor, planète-capitale de l'univers, où l'histoire, devenue enfin science exacte, est la clef du pouvoir...

C'est surtout sur le plan idéologique et sur celui des mœurs que la science-fiction américaine va évoluer.

A côté des œuvres franchement racistes ou xénophobes que nous avons évoquées, la plupart des œuvre de "l'âge d'or" (les années 40-50) se contentent de mettre sereinement en scène l'american way of life, maintenue avec juste ce qu'il faut de force [18].

Cette idéologie sommaire fait place peu à peu à une science-fiction humaniste, souvent humoristique. Certains contes philosophiques de Robert Sheckley sont des chefs-d'oeuvre du genre (*« La dimension des miracles »*, 1968).

La pureté des mœurs était initialement très grande : le flirt auquel le lecteur assistait obligatoirement tout au long du développement et qui souvent tenait lieu de développement, n'aboutissait qu'après l'épilogue.

Pourtant, dès 1935, la *« Passagère clandestine pour Mars »* de John Wyndham Beynon oubliait la pureté de la race dans les bras d'un martien romantique ; il est vrai que l'auteur était anglais.

Les *« Amants étrangers »* de Philip Jose Farmer ont encore fait scandale en 1952. Cette œuvre semble pourtant bien anodine aujourd'hui ne serait-ce qu'en comparaison des fantasmes érotiques, sadiques et mystiques auxquels Farmer lui-même nous a conviés depuis. La pruderie de la science-fiction a complètement disparu aujourd'hui ; elle s'est mise, sur ce plan, à l'unisson du reste de la littérature.

Son évolution idéologique remonte plus haut ; manifestement, aux bombes d'Hiroshima et Nagasaki. Les années qui ont suivi ont vu éclore des textes racontant non seulement l'horreur de la guerre nucléaire, mais les diverses conséquences biologiques, historiques et sociales que l'on pouvait imaginer : destruction de toute vie sur terre, remplacement de l'homme par d'autres espèces, invasion de monstres mutants par exemple ; régression de la vie sociale d'autre part, conduisant le plus souvent à la formation de tribus isolées, en guerre perpétuelle, pour qui le souvenir de la technologie s'estompe et prend l'allure du mythe.

Progressivement, la science-fiction toute entière a basculé ; elle plonge aujourd'hui très profondément dans l'angoisse et l'horreur sous toutes leurs formes actuelles : violence, torture, asservissement, désespoir. Parallèlement, le rôle joué par les thèmes proprement scientifiques diminue ; la science-fiction se transforme progressivement en *"spéculative-fiction"* , selon l'expression de Robert Henlein. Il est remarquable cependant qu'elle préserve son identité à travers cette mutation (le maintien des initiales S-F est significatif à ce point de vue) ; si les auteurs de la "new wave" n'acceptent plus de choisir un sujet scientifique pour leurs anticipations, c'est qu'ils ne croient plus que l'on puisse isoler la science et la technologie du reste de la vie sociale, ils ne croient plus, en particulier, que la science puisse apporter de solution aux problèmes éthiques qui se posent de toute part.

Est-ce à dire qu'il n'y a plus de lien aujourd'hui entre science et science-fiction ? Au contraire : cette évolution est le reflet exact de celle des scientifiques eux-mêmes. Comment échapperaient-ils à l'angoisse et à l'horreur après l'usage des bombes nucléaires, après les méthodes "scientifiques" de guerre au Viet-Nam ? Comment ne craindraient-ils pas les techniques d'asservissement et de dépersonnalisation qui sont effectivement mises en œuvre : lobotomie, stérilisation, drogues psychotropes, conditionnement psychologique et tout simplement l'usage possible des moyens audio-visuels et de l'informatique ? *L'optimisme scientiste a disparu comme mode littéraire en même temps qu'il perdait sa crédibilité dans la conscience des scientifiques.*

Dans cette évolution parallèle, la science-fiction n'est pas restée à la traîne : ainsi le coup de tonnerre écologique qui a éclaté en occident il n'y a pas si longtemps était préparé de longue date par la science-fiction : les *« Chroniques martiennes »* de Ray Bradbury nous décrivaient en 1950 les dégâts faits par les Terriens débarquant sur Mars (ils commencent par exterminer la population en lui communiquant la varicelle) ; les récits des conséquences de la pollution et de la surpopulation sont parmi les plus anciens.

18. Citons l'œuvre ingénieuse de Poul Anderson, créateur du conservatisme de l'avenir : dans la *« Patrouille du temps »*, un corps de police d'élite, venu du futur, veille à ce que le cours de l'histoire ne soit modifié par aucune innovation *intempestive* (au plein sens de ce mot).

Sans doute les auteurs de science-fiction ont-ils plus que les hommes politiques et que certains scientifiques l'habitude d'envisager des effets à long terme ; dans cette voie ils ont souvent montré plus de discernement que les "futurologues" professionnels.

Les rapports entre science-fiction et science peuvent évidemment s'analyser sur un tout autre plan. Que dire par exemple des *erreurs scientifiques* dans les œuvres d'imagination ?

D'abord qu'il ne s'agit pas à proprement parler "*d'erreurs*" mais plutôt de *contradictions* ; la science-fiction commence par un délibéré à côté de la science, donc par une erreur à priori relativement à un état donné de la science, "erreur" qui fait partie de la règle du jeu. Mais la crédibilité du récit exige une certaine cohérence du discours ; cohérence qui conduit fréquemment les auteurs à entrer assez loin dans les explications de type "scientifique", bien qu'elles ne puissent évidemment pas être véritablement scientifiques.

Il y a là un petit jeu entre l'auteur et le lecteur qui s'apparente à la prestidigitation : l'auteur de talent, comme le bon magicien, sait faire oublier certains détails : le bon lecteur est content quand il est honnêtement roulé. Tout au plus le lecteur scientifique éprouve-t-il un léger agacement lorsqu'un menu détail qui aurait pu être corrigé sans rien enlever à l'ensemble, a échappé à l'imagination et à la sagacité de l'auteur.

Ces explications "scientifiques" pourraient rapidement devenir fastidieuses ; heureusement, les règles du genre sont là pour les arrêter quand elles n'amusent plus personne. Malgré tout le talent de Jules Verne – qui d'autre pourrait rendre passionnant un chapitre romanesque intitulé « *Un peu d'algèbre* » – on relève chez lui d'innombrables contradictions scientifiques [19].

N'en étudions qu'une : dans l'obus qui conduit ses cosmonautes de la Terre à la Lune, la pesanteur diminue progressivement, et ne s'annule qu'au "point d'égale attraction" entre la Terre et la Lune ; et cependant les corps de deux chiens, évacués dans le vide, restent constamment à proximité de l'obus [20].

Tout le monde sait bien, *aujourd'hui*, que l'apesanteur règne dans les véhicules célestes qui ont stoppé leur propulsion, et aussi qu'on ne court aucun risque de "tomber" sur Terre en sortant d'un satellite artificiel. Mais c'est sans doute pour l'avoir vu à la télévision : auparavant ces faits étaient refusés par le public. Ainsi, à propos de sa nouvelle « *Jupiter V* » (où un bon scientifique réduit à merci un méchant voleur en faisant mine de le précipiter sur Jupiter du haut d'un satellite artificiel), Arthur Clarke croyait utile de préciser [21] : « *Cet ouvrage, qui comprenait vingt ou trente pages de calculs interplanétaires* [22]*, devrait être dédié au professeur G.C. Mc Vittie, mon maître de mathématiques appliquées. Je souligne ce fait … pour préciser que les faits exposés dans cette histoire, aussi déroutants qu'ils paraissent, son rigoureusement exacts, loin d'être un produit de mon imagination. De plus, ils ne concernent pas seulement l'orbite lointaine de Jupiter, mais affecteront, bien plus près de nous, les satellites artificiels de la prochaine décade* ».

Marcel Boll opposait aux erreurs de Jules Verne la rigueur des anticipations de Herbert George Wells. Et pourtant nous voyons aujourd'hui combien les prévisions de Jules Verne ont mieux collé à la réalité que celle de Wells : « *la machine à explorer le temps* », « *l'homme invisible* », la « *cavorite* » (matériau antigravitionnel des "premiers hommes dans la Lune") sont aussi lointains de nous scientifiquement que du temps où Wells les imaginait (1895-1901).

Mais ce n'est pas la hardiesse de ses anticipations qui diminue la valeur de l'œuvre de Wells, pas plus d'ailleurs que les contradictions

19. Dans certains cas, il ne se soucie même pas d'un début de vraisemblance mécaniste : dans « *Hector Servadac* » (1877), une comète frôle la Terre et emporte des territoires entiers (qui se sont "retournés" pour se coller sur le nouvel astre) sans que leurs habitants s'en aperçoivent immédiatement.

20. L'opinion selon laquelle c'est le vide qui annule la pesanteur est largement répandue ; elle appartient peut-être à un stade génétique, au sens de Piaget. On peut aussi la rapprocher de l'opinion aristotélicienne selon laquelle l'inertie des projectiles est une manifestation de l'atmosphère, qui revient à les pousser par derrière ("antiperistasis").

21. Préface du recueil « *Demain, moisson d'étoiles* » (antérieur à 1960).

22. Sic

23. Un exemple : « *L'homme invisible* » a réussi à disparaître en ramenant à l'unité l'indice de réfraction de ses humeurs (étant albinos, il est censé être dépourvu de tout pigment). Et pourtant, il voit. Comment fonctionnent donc son cristallin et sa rétine dans ces conditions ? Il est probable que cette contradiction n'a pas échappé à Wells (elle a été révélée plusieurs fois par la suite), mais qu'il s'est accordé une "licence poétique".

scientifiques que l'on peut s'amuser à y relever aussi [23].

Même si certains, comme Arthur Clarke, se sont plu à écrire de la *science-fiction didactique* – nous en avons un exemple ci-dessus [24] – ce n'est pas leur unique propos ; Clarke, par exemple, a écrit des œuvres dont le lyrisme fait oublier la référence scientifique : pourquoi diable analyserait-on « *les enfants d'Icare* » comme un traité de physique ?

Dans le même ordre d'idées, mais plus importante socialement, se pose la question de savoir dans quelle mesure la science-fiction peut contribuer à répandre ou accréditer les "fausses sciences".

Le risque d'une simple confusion entre la fiction et la réalité semble minime : il faudrait croire les lecteurs particulièrement débiles pour craindre qu'ils confondent les récits de science-fiction avec des reportages ou des ouvrages scientifiques : à ce compte-là, le *"Père Noël"* et le *"Petit Chaperon Rouge"* détruiraient définitivement tout sens critique dans les notions occidentales. Comme les rêves, les histoires merveilleuses semblent nécessaires à l'équilibre de la pensée : leur universalité est là pour en témoigner.

Un danger existe cependant : certains n'hésitent pas à amalgamer les histoires de science-fiction avec le pire appel à la crédulité, parfois avec un certain succès. Qu'on se souvienne du « *Matin des Magiciens* » et la revue « *Planète* » (1961-1967) dans laquelle Louis Pauwels et Jacques Bergier, servis par une iconographie superbe et un bagout considérable, proposèrent "le réalisme fantastique". L'accolement de ces deux mots caractérise parfaitement l'intention mystificatrice ("réel" et "fantastique" sont donnés comme antonymes dans tout bon dictionnaire). La science-fiction (appelée "littérature différente" y est utilisée systématiquement ; l'étude de textes anciens par ordinateur devient la *"première utilisation d'une machine à remonter le temps"* ; d'excellents textes (Lovecraft, Damon Knight, Borges par exemple), parfois tronqués, sont intercalés dans un étonnant mélange où l'on trouve des extraits de rapports scientifiques, des fables qui semblent reprises aux almanachs du XVIème siècle, des études littéraires, et les élucubrations des génies méconnus qui gravitent dans les zones marginales de la pensée.

Il serait dangereux cependant d'interpréter cette affaire trop sommairement, en se contentant d'opposer la raison et l'obscurantisme : une plaquette collective, intitulée « *le Crépuscule des Magiciens* », tenta en vain d'endiguer le phénomène « *Planète* » par un appel pas toujours adroit au rationalisme.

Ce n'était pas nécessaire : « *Planète* » s'étiola, plus par l'essoufflement de son inspiration que par le manque d'enthousiasme de ses lecteurs, toujours à l'affût de la « *Révélation* » qui devait permettre de tout comprendre sans vraiment se fatiguer ; la révélation tardait à venir, hélas.

A la limite, *"le manichéisme rationaliste"* est aussi contradictoire [25] et dangereux que le *"réalisme fantastique"* lui-même : on ne lutte pas efficacement contre la crédulité en feignant d'oublier que la science est un dialogue entre la créativité de l'imagination d'une part, les faits expérimentaux et le raisonnement déductif d'autre part.

Ce dialogue, intériorisé, est la structure même de la recherche mathématique : la plupart des mathématiciens en sont parfaitement conscients. Les mathématiques s'apparentent ainsi aux activités artistiques (la composition musicale est aussi une dialectique entre le message subjectif à transmettre et la technique des matériaux sonores) ; *elles n'en sont pas moins les garanties de l'objectivité de la science*, selon le point de vue positiviste le plus largement admis.

Si l'on cachait l'un des termes de ce dialogue, l'activité scientifique se réduirait à un inventaire morose de faits et de chiffres, cessant d'être aussi une aventure de l'esprit. Comment s'étonner alors que le public, et particulièrement la jeunesse, s'en détourne et se fourvoie dans des impasses traditionnelles ? Comment ne pas s'apercevoir aussi que la science deviendrait *incompréhensible*, qu'elle se transformerait en un rituel stupide ?

---

24. Voir aussi « *Les sables de Mars* » (1951), et une partie du scénario du film « *2001, odyssée de l'espace* » (1967).

25. Une photographie prise de l'avion "Condorde", au-dessus de l'Afrique, lors de l'éclipse du 30 juin 1973, pouvait passer pour celle d'une "soucoupe volante" ; elle reçut par la suite une autre interprétation, très satisfaisante (il s'agirait de l'impact d'une météorite de l'essaim associé à la comète d'Ericke). Pourquoi donc les auteurs de cette explication avaient-ils déclaré publiquement, quelques jours auparavant, qu'un tel fait *ne pouvait pas* être pris en compte par la science, parce qu'il était isolé ?

L'intention confusionniste est l'exception en science-fiction ; prendre ses hypothèses au-delà de la science, ce n'est pas s'opposer à la science. Est-ce un moyen de la stimuler ?

Nous avons constaté ce rôle stimulant de l'anticipation sur la technologie, la technologie spatiale par exemple. Examinons maintenant le cas de la recherche fondamentale.

Cette étude est à l'évidence difficile. Comment évaluer "objectivement" le rôle que peut jouer la littérature d'anticipation dans le genèse des découvertes fondamentales, alors que ces découvertes se font parfois dans le secret d'une seule conscience, et que leur auteur n'est pas toujours capable d'en expliquer le mécanisme ?

Il est vrai qu'il s'agit souvent de l'aboutissement d'une longue maturation collective, maturation qui laisse des traces écrites. Mais les éléments que nous recherchons ici sont les plus "subjectifs", ceux qui sont le plus souvent censurés dans l'expression écrite de la recherche, même s'ils s'expriment librement dans les conversations privées, par exemple. Cette sorte de pudeur des scientifiques vis-à-vis des éléments intimes de leur motivation est parfois tempérée, notamment par l'habitude anglo-saxonne de placer des textes littéraires en exergue des développements scientifiques ; encore faut-il faire la place du "joke" dans le choix des citations.

Sur ce problème qui nous semble cependant essentiel, nous nous contenterons donc de quelques éléments de réflexion.

Parmi les écrivains de science-fiction, il en existe une proportion notable qui exerce, ou a exercé, une profession scientifique ; plus qu'il ne pourrait sembler à première vue, à cause de l'usage répandu des pseudonymes.

Mais quelle est la proportion de chercheurs qui lisent habituellement de la science-fiction ? Cette question pourrait faire l'objet de sondages dans les divers milieux scientifiques, nous ignorons s'il y en a eu. Ce que nous pouvons rapporter, c'est que lors de discussions collectives évoquant ce thème [26], attirant un public nombreux et passionné venant de disciplines scientifiques et littéraires fort diverses, la plupart des participants venaient beaucoup plus s'*informer* sur la science-fiction, que défendre leurs opinions à ce sujet ; les lecteurs réguliers étaient visiblement très minoritaires, bien qu'ils fussent plus "motivés" que les autres pour participer à de tels débats. Une expérience de ce genre suggère que les "fans" de la science-fiction sont actuellement peu nombreux parmi les chercheurs français, une information plus large et internationale étant évidemment souhaitable.

Mais s'il était montré que la proportion générale de lecteurs de science-fiction soit faible, il ne faudrait pas se hâter d'en conclure que la littérature d'anticipation ne peut jouer qu'un rôle négligeable dans l'élaboration de la recherche ; une analogie va nous le montrer.

Certains mathématiciens puristes, triomphalement installés au sommet de la classification d'Auguste Comte, prétendent que leur art ne doit rien aux problèmes posés par les sciences inférieures, telles que la physique. Ce qui signifie en clair qu'ils ne peuvent accorder leur estime à un collègue que s'ils ne le soupçonnent pas de prostituer les mathématiques en les "appliquant" (nous avons eu l'occasion de recueillir des commentaires édifiants sur la carrière de F. J. Dyson . . .).

Les mathématiciens qui récusent toute analyse de type marxiste de l'histoire des sciences savent bien le rôle joué, comme *poseurs de problèmes*, par de grands mathématiciens qui s'occupaient *aussi* d'autres sciences. Que seraient les mathématiques "pures" si Euler, Lagrange, Gauss, Henri Poincaré, Elie Cartan, David Hilbert, Hermann Weyl, parmi bien d'autres, n'avaient été les *médiateurs* entre les problèmes conceptuels des sciences expérimentales et les problèmes formels des mathématiques ? Remarquons, a contrario, les difficultés que rencontre la mathématisation des sciences humaines, pour lesquelles une telle maturation ne fait que commencer.

De même, dans tout cheminement de la pensée scientifique, on peut soupçonner que le rôle des médiateurs entre disciplines ou ordres de pensée différents soit parfois déterminant, sans que nécessairement leur action ait laissé de traces matérielles (par exemple dans le cas d'une conversation). On en est réduit à

26. Séminaire d'histoire et sociologie des idées et des faits scientifiques, Université de Provence, Marseille.

27. On raconte que le mathématicien Kummer (1810-1893), cherchant depuis longtemps (pour démontrer le dernier théorème de Fermat) à étendre la notion de P.G.C.D. aux cas où ne fonctionne pas l'algorithme d'Euclide, fut illuminé par une conversation avec un chimiste au sujet du *radical* ammonium (NH_4), qui se comporte comme un métal bien qu'on ne puisse pas l'isoler. C'est ainsi qu'il aurait découvert la notion d'*idéal* ; notion aujourd'hui fondamentale, non seulement en théorie des nombres et en algèbre, mais aussi en

quelques témoignages [27] ; quoi de plus naturel si ce témoignage manque, si l'auteur d'une synthèse en oublie de bonne foi la genèse ?

Il est donc possible, en particulier, que des œuvres de fiction aient pu jouer un rôle réel dans l'élaboration de concepts scientifiques nouveaux, alors même que ceux qui ont achevé cette élaboration n'en auraient pas eu connaissance directe ; mais ce genre d'hypothèse est évidemment difficile à étayer par des preuves.

Nous savons par exemple qu'Einstein, à qui l'on racontait en 1920 une histoire de science-fiction – « *Lumen* », de Camille Flammarion, dans laquelle un astronef, voyageant plus vite que la lumière, permet d'observer *la fin*, *puis le début* de la bataille de Waterloo – la réfutait parce qu'elle était en contradiction avec la théorie de la relativité [28]. Einstein savait-il que « *Lumen* » avait été écrit en 1872, et par conséquent devançait plutôt la relativité, en soulignant le caractère paradoxal d'un mouvement matériel plus rapide que la lumière ?

C'est en 1895 que parut le texte suivant : « *Manifestement, tout corps réel doit s'étendre dans quatre directions ; il doit y avoir longueur, largeur, épaisseur, et durée. . .* Il n'y a aucune *différence entre le Temps, quatrième dimension, et l'une quelconque des trois dimensions de l'espace, sinon que notre conscience se meut avec elle* [29]*. Mais quelques imbéciles se sont trompés sur le sens de cette notion.* » C'est un extrait de « *The Time Machine* » ; son auteur, H.G. Wells, était alors un obscur journaliste ; Lorentz proposait cette année-là d'interpréter le résultat de l'expérience de Michelson par une "contradiction" incompréhensible ; Einstein était un lycéen de seize ans.

Il est vraisemblable qu'il n'ait pas eu connaissance de ce texte, « *La machine à explorer le temps* » connut immédiatement un succès considérable.

L'extrait que nous avons cité va nettement plus loin que la simple considération d'un produit cartésien espace x temps, arène du monde physique, déjà attestée au XVIIIème [30], puisqu'il affirme l'identité complète des quatre dimensions ; il est cependant scientifiquement irréprochable, il est par exemple écrit en exergue du chapitre « *Special Relativity* » dans le traité « *Gravitation and Cosmology* » de Steven Weinberg (1972). Comment pourrait-on affirmer qu'il n'a eu aucune influence sur la genèse de la relativité restreinte, dix ans plus tard, et même de la relativité générale, dont il est dans le droit fil ? Même si Wells a exprimé d'autres idées qui n'ont pas eu de postérité scientifique, il appartient évidemment au chercheur de sélectionner dans l'univers des faits *comme dans celui des idées*, les germes de créations nouvelles.

Des exemples analogues, plus récents, sont faciles à trouver, quoique moins convaincants en raison même de leur actualité. Il semble hors de doute, par exemple, que le thème des "univers parallèles", un des poncifs de la science-fiction américaine pendant vingt ans, inspire directement la tentative "non copenhaguienne" d'interprétation de la mécanique quantique dans laquelle chaque transition est effectivement une bifurcation entre plusieurs futurs, chacun réalisé dans un univers différent [31].

La désignation du premier *pulsar* par les initiales L.G.M. (« *Little Green Men* »), en 1967, témoigne que la référence à la science-fiction était immédiatement consciente chez les radios-astronomes auteurs de la découverte. Est-il possible de savoir si elle a joué un rôle positif, aidant par exemple à envisager l'hypothèse que le signal périodique reçu pouvait avoir une origine extra-terrestre, qu'il ne s'agissait pas prosaïquement d'un simple parasite ?

De même, la découverte dans l'espace interstellaire de molécules organiques (précurseurs de certains acides aminés), qui semblent éjectées du noyau de la Galaxie, a-t-elle pu avoir lieu sans évocation naturelle du thème de la panspermie ? Même s'il s'avère que ce ne sont pas les Grands Galactiques qui ensemencent régulièrement leurs Terres, la référence poétique

analyse et même en mécanique quantique (voir la théorie spectrale de Gelfand) ; et donc, par un juste retour des choses, en chimie théorique . . .

28. « *Einstein, Einblicke in seine Gedankenwelt* », par A. Moszkowski (1921).

29. Souligné par l'auteur.

30. *"Les temps et les espaces n'ont pas d'autres lieux qu'eux-mêmes, et ils sont les lieux de toutes choses"* (Newton, Principia, traduction de la Marquise du Chatelet) ; *"- un homme d'esprit de ma connaissance croit qu'on pourrait regarder la durée comme une quatrième dimension et que le produit du temps par la solidité serait en quelque manière un produit de quatre dimensions. Cette idée peut être contestée, mais elle a, il me semble, quelque mérite, quand ce ne serait que celui de la nouveauté"* (D'Alembert, dans l'Encyclopédie, cité par J. Van Herp (loc.cit.)).

31. Hugh Everett, Rev. Mod. Phys. 29.3 p.454 (1957). Voir aussi les "Battelle rencontres" de 1967 (éditées par B. Dewitt et J.A. Wheeler).

aura pu jouer son rôle : il est plus facile de découvrir ce à quoi l'on a déjà pensé ; chacun est aveugle à ce qui est trop radicalement nouveau, l'histoire des occasions manquées de la science le manifeste souvent.

Conclusion

L'histoire des sciences montre aussi que des difficultés apparemment insurmontables finissent par disparaître devant une simple modification d'éclairage, une nouvelle façon de poser les questions ; ce sont souvent ces nouveaux points de vue qui constituent les progrès les plus définitifs.

La science-fiction, malgré les énormes scories qu'elle contient, est une mine quasi-inépuisable de thèmes nouveaux, de rapprochements paradoxaux, de renversements idéologiques ; elle peut donc aider les chercheurs à vaincre les obstacles épistémologiques qu'ils rencontrent, même s'ils ne la lisent que pour succomber au vertige sensuel des idées impossibles.

Serait-il sage de les en détourner ?

-o-o-o-o-o-o-o-o-o-o-o-

(le 5-6-74).

(2016) B.I.A.A. no **105** BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
ISSN 0291-5413 (Université d'Aix-Marseille)

Mécanique quantique III : de de Broglie à Schrödinger

Eric OLIVIER [1] [2]

Résumé. – Les paquets d'ondes associés à l'onde-corpuscule de de Broglie sont solutions d'une équation d'évolution qui est un cas particulier de l'équation de Schrödinger. L'année 1926 voit la publication d'une série d'articles de Schrödinger qui marquent la naissance de la « *mécanique ondulatoire* » : celle-ci s'appellera bientôt la « *mécanique quantique* ». Pour cette nouvelle mécanique, Schrödinger propose un « *principe de correspondance* » qui a chaque quantité classiquement mesurable (comme l'énergie, la position ou l'impulsion) fait correspondre une « *observable* », c'est-à-dire un opérateur autoadjoint (généralement non borné). « *L'interprétation probabiliste de Born* » associée au « *théorème d'Ehrenfest* » sur les moyennes des observables, rend le principe de correspondance accessible à l'intuition. Cependant, la possible non commutativité des couples d'observables (en particulier les observables de position et d'impulsion) implique une différence entre mécanique classique et quantique qui devient essentielle dans le monde microscopique : cette différence s'explicite dans le « *principe d'incertitude de Heisenberg* ». Afin d'illustrer tout cela, nous traitons le cas de « *l'oscillateur harmonique quantique* (o.h.q.), à la fois pour la simplicité du modèle mécanique en question, mais aussi pour son importance théorique dans la découverte du quantum d'action. Nous verrons comment la résolution de l'équation de Schrödinger et le principe d'incertitude de Heisenberg permettent de retrouver (d'expliquer !) l'hypothèse de quantification de l'énergie dans l'heuristique de Planck en montrant – grâce au principe d'incertitude de Heisenberg – que le niveau minimal d'énergie de l'o.h.q. ne peut être nul.

1. Introduction

1.1. La fin du 19-ème siècle est marquée par le retour des structures discrètes (les atomes) comme concept élémentaire des théories physiques. Le changement de paradigme du continue au discret est particulièrement frappant chez Maxwell, qui passe de la théorie de l'électromagnétisme à la théorie cinétique des gaz. Parallèlement et conjointement Boltzmann défend l'hypothèse atomique et amorce le passage de la thermodynamique phénoménologique à la mécanique statistique. Ces deux physiciens donneront leur nom à la « *statistique de Maxwell-Boltzmann* ». La naissance de la « *mécanique quantique* » vient en 1901 lorsque Planck applique les idées de Boltzmann pour élucider une énigme de l'électromagnétisme : le rayonnement du corps noir. Planck est alors amené à faire l'hypothèse de l'existence d'un « *quantum d'action* », une idée qui dérange son intuition de physicien. Puis vient Einstein et la découverte des « *quanta de lumière* » en 1905 [Ein05]. En 1924, de Broglie se base sur le travail d'Einstein (i.e. les deux relativités et la théorie des quanta de lumière) pour introduire dans sa thèse [dBL25] la notion « *d'onde-corpuscule* » (voir aussi la conférence Nobel de de Broglie [dB29]). En redonnant une nature corpusculaire (non massive) aux ondes lumineuses, Einstein amène de Broglie à imaginer un corpuscule massif (typiquement un électron) ayant une nature ondulatoire [3] (c.f. Sections 2 & 4).

1. GDAC-I2M UMR 7373 CNRS Université d'Aix-Marseille
2. eric.olivier@univ-amu.fr
3. De Broglie pense en terme de « *dualité onde-corpuscule* » c'est-à-dire à un objet de la physique de nature corpusculaire « *et* » ondulatoire : il développera cette idée tout au long de sa vie, en particulier en imaginant le corpuscule comme une singularité d'une solution d'une équation d'évolution possiblement non linéaire (théorie de la double solution).

Hamilton et Jacobi avaient déjà développé une formulation ondulatoire de la mécanique (équation de Hamilton-Jacobi) construite par analogie avec la théorie de l'approximation de l'optique ondulatoire par l'optique géométrique classique (optique des rayons lumineux) : mais cette approche est restée formelle à l'époque de Hamilton et Jacobi, parce que trop en avance sur son temps (nous reviendrons sur ce point dans [Oli17]). Grâce au travail théorique de de Broglie, la nature ondulatoire de l'électron est mise en évidence en 1927 par les expériences (indépendantes) de G. P. Thompson [4] et C. J. Davinson (diffraction des électrons par les cristaux). La thèse de de Broglie est donc le point de départ d'une nouvelle approche du phénomène quantique à priori différente de la formulation initialement proposée par Bohr, puis par l'*« école de Copenhagen »* avec *« la mécanique des matrices »* de Heisenberg, Born et Jordan ([Hei25, BJ25, BHJ26]) : en 1926 Schrödinger écrit une série d'articles (voir en particulier [Sch26]) dans lesquels il construit la *« mécanique ondulatoire »*.

1.2. Pour une introduction heuristique de la *« mécanique ondulatoire »*, nous partons du fait que les paquets d'ondes associés à l'onde-corpuscule de de Broglie sont solutions d'une équation d'évolution [5] qui est un cas particulier de *« l'équation de Schrödinger »* (c.f. Section 5). Un des points les plus importants de l'approche de Schrödinger est de proposer un *« principe de correspondance »* [6] qui à chaque quantité classiquement mesurable (comme la position ou l'impulsion) fait correspondre l'ensemble des valeurs spectrales d'une *« observable »*, c'est-à-dire d'un opérateur autoadjoint. Bien que très polémique, l'*« interprétation probabiliste de Born »* (c.f. § 5.5), lorsque celle-ci est associée au *« théorème d'Ehrenfest »* (c.f. Théorème 6.4 in § 6.3), rend le principe de correspondance accessible à l'intuition. La mécanique quantique peut alors être vue – jusqu'à un certain point – comme un raffinement de la mécanique classique. Cependant, la possible non commutativité des couples d'observables (en particulier les observables de position et d'impulsion) est l'indice d'une différence radicale entre le monde microscopique quantique et le monde macroscopique classique. Cette différence amène directement au *« principe d'incertitude de Heisenberg »* qui est la pierre angulaire de la *« mécanique quantique »*. Nous avons déjà rencontré (c.f. [Oli16a]) une version de ce principe dans la théorie du signal, sous la forme de *« l'inégalité de Heisenberg-Gabor »* : en introduisant *« l'échelle de Planck »*, nous retrouvons un cas particulier du principe d'incertitude, i.e. *« l'inégalité de Heisenberg »* (en position-impulsion) pour les paquets d'ondes de de Broglie-Schrödinger unidimensionnel (c.f. Section 5) ; nous verrons aussi la version du principe d'incertitude dans le cadre abstrait de la mécanique ondulatoire, où il exprime une forme déguisée de l'inégalité de Cauchy-Schwarz (c.f. Section 7). Comme illustration, nous traitons le cas de l'oscillateur harmonique (c.f. Section 8). Un aspect théorique important de ce système est liée à son rôle joué dans la découverte du quantum d'action : nous verrons comment le principe d'incertitude de Heisenberg peut être utilisé pour résoudre l'équation de Schrödinger,

4. Fils de J. J. Thompson découvreur de l'électron !
5. C'est bien l'équation qui est déduite de solutions ad hoc !
6. En un sens, le *« principe de correspondance de Schrödinger. »* peut-être interprété comme une version effective du *« principe de complémentarité de Bohr »* dans le cas de la mécanique ondulatoire.

afin de retrouver (et de préciser) les résultats de l'heuristique de Planck ; nous montrerons en particulier que le niveau minimal d'énergie de l'oscillateur harmonique quantique ne peut être nul : ceci a pour conséquence directe la quantifications de l'ensemble des niveaux d'énergie de cet oscillateur.

2. **Ondes et paquets d'ondes – Vitesse de phase, vitesse de groupe**

2.1. L'équation des ondes sur $\mathbb{R}$ s'écrit sous la forme

$$\frac{\partial^2\psi}{\partial x^2} = \frac{1}{u_0^2}\frac{\partial^2\psi}{\partial t^2} \tag{1}$$

où la fonction $\psi(t,x)$ – définie et C^2 sur $\mathbb{R}\times\mathbb{R}$ et à valeurs complexes – dépend des variables temporelle et spatiale (notées respectivement t et x) et où le réel u_0 (positif où négatif) est la *« vitesse de phase »* de l'onde. La résolution de (1) proposée par d'Alembert, consiste à séparer les variables spatiale et temporelle : ici, nous spécialisons cette méthode en recherchant une *« solution vibratoire »* de la forme[7] $\varphi(x)e^{-i\omega_0 t}$, avec $\varphi(x)$ complexe. Le paramètre ω_0 appelée *« pulsation »* est conventionnellement positif. La composante spatiale $\varphi(x)$ est alors nécessairement solution de *« l'équation de Helmoltz »* associée à (1), i.e. :

$$\frac{d^2\varphi}{dx^2} = -\left(\frac{\omega_0}{u_0}\right)^2\varphi(x)$$

Nous obtenons $\varphi(x) = Ae^{\pm ik_0x}$ où A est un nombre complexe et où $k_0 = \omega_0/u_0$ désigne *« le nombre d'onde[8] »* : l'équation reliant k_0 et ω_0 est la *« relation de dispersion »*, soit :

$$\omega_0 = u_0k_0 \tag{2}$$

La pulsation ω_0 étant positive, le nombre d'onde et la vitesse de phase sont des réels de même signe. Dans la suite nous choisirons (conventionnellement[9]) de prendre $\varphi(x) = Ae^{ik_0x}$ pour définir les *« ondes monochromatiques »* comme les solutions de (1) de la forme

$$Ae^{i(k_0x-\omega_0t)} = Ae^{ik_0(x-u_0t)}$$

(avec $A\in\mathbb{C}$). La description ondulatoire d'un corpuscule libre et doté de masse (e.g. l'électron) est initiée dans le travail de thèse de de Broglie au début des années 20 : avec l'introduction de la notion de *« paquet d'ondes »* l'idée consiste à amalgamer – par *« transformée de Fourier »* – un continuum d'ondes monochromatiques de la forme

$$A(k)e^{i(kx-\omega(k)t)}$$

où le nombre d'onde k est variable et où la pulsation $\omega(k)$ (à valeurs positives) ainsi que *« l'amplitude complexe »* $A(k)$ sont des fonctions de k. Nous utiliserons la définition suivante qui introduit l'analogue, pour la théorie du signal, des paquets d'ondes de de Broglie (définis plus bas au § 2.1).

7. Si Γ est le graphe de la fonction réelle $f(x)$ définie pour tout x réel alors, pour tout $a\in\mathbb{R}$ donné, le graphe de la fonction $f(x-a)$ est le translaté $\Gamma+a$ du graphe de $f(x)$: nous allons voir que cette remarque justifie l'introduction du « – » dans $e^{-i\omega_0t}$; en effet, la forme conventionnelle $\varphi(x)e^{-i\omega_0t}$ de la solution vibratoire cherchée par *séparation des variables* nous donne les solutions monochromatiques de la forme $Ae^{i(\pm kx-\omega_0t)}$ où $k_0=\omega_0/u_0$: cette solution s'écrivant sous la forme $f(x\mp tu_0)$, son graphe se déplace à la vitesse $\pm u_0$ (où u_0 est vitesse de phase de l'équation des ondes considérée).

8. En dimension supérieure on parle de vecteur d'onde.

9. Ecrire $\varphi(x) = Ae^{ik_0x}$ au lieu de $\varphi(x) = Ae^{\pm ik_0x}$ revient à intégrer le signe $\pm$ au nombre d'onde k_0.

Définition 2.1. *Soit $k \mapsto A(k)$ dans $L^2(\mathbb{R})$ t.q. $\|A\| = 1$ et $k \mapsto \omega(k)$ une fonction de classe C^1sur $\mathbb{R}$; alors le « paquet d'ondes (général) » d'amplitude complexe $A(k)$ et de pulsation $\omega(k)$ est la fonction $\psi(t,x)$ dépendant du temps t et de l'espace x telle que*

$$\psi(t,x) = \frac{1}{\sqrt{2\pi}} \int A(k) e^{\boldsymbol{i}(kx-\omega(k)t)} dk \tag{3}$$

Lorsque $\omega(k) = \pm k u_0$ nous dirons que $\psi(t,q)$ en (3) est un paquet d'ondes pur.

Remarque 2.2. *(1) : Le facteur $1/\sqrt{2\pi}$ est introduit pour des raisons formelles, en référence à la transformée de Fourier L^2 que nous notons $\mathfrak{F}_{\pm}[\cdot]$ (c.f.* [Oli16a, § 7]*). Ainsi en posant $\hat{\psi}_t(k) := A(k)e^{-\boldsymbol{i}\omega(k)t}$, le paquet d'ondes en (3) s'écrit*

$$\psi(t,x) = \frac{1}{\sqrt{2\pi}} \int \hat{\psi}_t(k) e^{\boldsymbol{i}kx} dk = \mathfrak{F}_{+}[\hat{\psi}_t](x)$$

(2) : Si la fonction de pulsation est de la forme $\omega(k) = \pm k u_0$ (i.e. $\omega(k) = \omega_0 \pm (k-k_0)u_0$ avec $\omega_0 = k_0 u_0$), alors nous pouvons transformer l'expression du paquet d'onde en (3) en écrivant

$$\psi(t,x) = \frac{1}{\sqrt{2\pi}} \int A(k) e^{\boldsymbol{i}k(x \mp u_0 t)} dk = \mathfrak{F}_{+}[A](x \mp t u_0)$$

Dans ce cas, le paquet d'onde est solution de l'équation des ondes (1) associée à la vitesse de phase u_0, d'où la dénomination de « paquet d'onde pur ». Insistons sur le fait que si la fonction de pulsation $\omega(k)$ n'est pas de la forme $\pm k u_0$, alors la fonction $\psi(t,q)$ dans (3) ne satisfait pas l'équation des ondes : il est raisonnable de penser que ce point de la thèse de de Broglie a pu influencer les réflexions de Schrödinger sur la mécanique ondulatoire et l'amener à établir « l'équation qui porte son nom » (c.f. § 5.2 infra).

2.2. L'hétérogénéité des vitesses de phase des composantes monochromatiques rentrant dans la décomposition de $\psi(t,x)$ est associée à une notion de *« vitesse de groupe »*. Cette remarque va permettre à de Broglie de donner une définition cohérente de la vitesse ondulatoire d'un corpuscule massif à partir de la notion de paquets d'ondes. Nous allons définir heuristiquement *« la vitesse de groupe »* du *« paquet d'onde »* $\psi(t,x)$ en (3) lorsque l'amplitude complexe $A(k)$ est concentrée autour d'un nombre d'onde k_0 donné. Pour simplifier, nous supposons que le support $A(k)$ est inclus dans l'intervalle $[k_0 - \varepsilon\,; k_0 + \varepsilon]$ avec ε petit. Soit $\omega_0 := \omega(k_0)$ et définissons la quantité

$$\boldsymbol{v} := \frac{d\omega}{dk}(k_0) \tag{4}$$

(qui est homogène à une vitesse). Nous supposons que ε est suffisamment « *petit* » pour assurer l'approximation $\omega(k_0 + h) \approx \omega_0 + h\boldsymbol{v}$ dès que $|h| \leq \varepsilon$. Alors, il vient :

$$\begin{aligned} \psi(t,x) = \frac{1}{\sqrt{2\pi}} \int \frac{A(k)}{\sqrt{2\pi}} e^{\boldsymbol{i}(kx-\omega(k)t)} dk \quad &\approx \quad \frac{1}{\sqrt{2\pi}} \int A(k) e^{\boldsymbol{i}\left((k-k_0)x + k_0 x - (\omega_0 + (k-k_0)\boldsymbol{v})t\right)} dk \\ &= \left(\frac{1}{\sqrt{2\pi}} \int A(k) e^{\boldsymbol{i}(k-k_0)(x-\boldsymbol{v}t)} dk \right) e^{\boldsymbol{i}(k_0 x - \omega_0 t)} \end{aligned}$$

ce qu'on peut aussi écrire

$$\psi(t,x) \approx \Psi_0(x - \boldsymbol{v}t) e^{\boldsymbol{i}(k_0 x - \omega_0 t)} = \Psi_0(x - \boldsymbol{v}t) e^{\boldsymbol{i}k_0(x - v_0 t)} \tag{5}$$

où nous avons posé $v_0 = \omega_0/k_0$ et où

$$\Psi_0(x) = \frac{1}{\sqrt{2\pi}} \int A(\kappa + k_0)e^{i\kappa x}d\kappa$$

Le facteur $e^{i(k_0 x - \omega_0 t)} = e^{ik_0(x - v_0 t)}$ décrit une *onde de phase monochromatique* de nombre d'onde k_0 et de vitesse de phase $v_0 = \omega_0/k_0$, alors que $\Psi_0(x - \boldsymbol{v}t)$ représente une onde dont la vitesse de phase est égale à $\boldsymbol{v}$. Lorsque $v_0 \neq \boldsymbol{v}$, le signal $\psi(t,x)$ n'est pas une onde en ce sens qu'il n'a pas de vitesse de phase (i.e. $\psi(t,x)$ n'est pas solution d'une équation des ondes) ; cependant nous pouvons déduire de (5) que *« l'énergie du signal »* $\psi(t,x)$ soit $|\psi(t,x)|^2 \approx |\Psi_0(x - t\boldsymbol{v})|^2$ est (approximativement) solution de l'équation des ondes de vitesse de phase $\boldsymbol{v}$: en d'autres termes, l'énergie portée par le paquet d'ondes $\psi(t,x)$ se propage à la vitesse $\boldsymbol{v}$. Par définition, lorsque l'amplitude complexe $A(k)$ d'un paquet d'ondes $\psi(t,x)$ est localisée en k_0, sa vitesse de groupe est la vitesse $\boldsymbol{v}$ définie en (4).

Remarque 2.3. *Si l'amplitude complexe $A(k)$ du paquet d'ondes $\psi(t,x)$ n'est pas localisée, nous pouvons (heuristiquement) écrire $A(k) = \sum_n A_n(k)$ où chaque amplitude complexe*[10] *$A_n(k)$ est concentrée autour d'un nombre d'onde k_n. La linéarité de la transformée de Fourier permet alors d'écrire que $\psi(t,x) = \sum_n \psi_n(t,x)$, où $\psi_n(t,x)$ est le paquet d'ondes associé à l'amplitude complexe $A_n(k)$ et à la pulsation $\omega(k)$: en d'autres termes, $\psi(t,x)$ est la superposition de paquets d'ondes localisés – i.e. les $\psi_k(t,x)$ – chacun d'eux ayant une vitesse de groupe approximativement égale à $\boldsymbol{v}_n := \frac{d\omega}{dk}(k_n)$: le fait que les $\boldsymbol{v}_n$ soient des vitesses possiblement différentes est lié au phénomène de « d'étalement du paquet d'ondes »*[11] *(c.f. § 3).*

Nous verrons que la notion de vitesse de groupe d'un paquet d'ondes se généralise aux fonctions d'ondes de la mécanique ondulatoire grâce au couple d'opérateurs (observables) impulsion/position (c.f. Proposition 5.3) : cela se systématise avec le théorème de Ehrenfest (c.f. § 6.3).

2.3. Nous avons déjà précisé que le facteur $1/\sqrt{2\pi}$ dans la définition du paquet d'ondes en (3) est introduit à des fins de normalisation relativement à l'utilisation de la transformée de Fourier $\mathfrak{F}_\pm[\cdot]$ sur $L^2(\mathbb{R})$ (c.f. [Oli16a]). Cela n'est pas un détail et nous verrons (c.f. [Oli17]) que la transformée de Fourier joue un rôle crucial dans le *« principe de correspondance de Schrödinger »*. Pour l'instant notons que si ψ_t désigne la fonction partielle $x \mapsto \psi(t,x)$, alors $\psi_t(x) = \mathfrak{F}_+[\hat{\psi}_t](x)$, avec $\hat{\psi}_t(k) := A(k)e^{-i\omega(k)t}$ et que

$$\|\hat{\psi}_t\|^2 = \int A(k)e^{-i\omega(k)t}\overline{A(k)e^{-i\omega(k)t}}dk = \int |A(k)|^2 dk = \|A\|^2 = 1$$

Or la transformée de Fourier $\mathfrak{F}_+[\cdot]$ étant une isométrie de $L^2(\mathbb{R})$, nous obtenons que $\|\psi_t\| = \|\hat{\psi}_t\| = 1$. Nous venons de montrer que si $\psi(t,q)$ est un paquet d'ondes au sens de la Définition 2.1, alors l'application $t \mapsto \psi_t$ est un chemin sur la sphère unité de $L^2(\mathbb{R})$.

Proposition 2.4. *Soit $\psi(t,q) = \mathfrak{F}_+[\hat{\psi}_t]$ le paquet d'onde t.q. $\hat{\psi}_t(k) = A(k)^{-i\omega(k)t/\hbar}$; si ψ_t est la fonction partielle $x \mapsto \psi(t,x)$ alors $\|\psi_t\| = \|\hat{\psi}_t\| = \|A\| = 1$ pour tout $t \geq 0$.*

10. Déduite de $A(k)$ en utilisant (par exemple) des partitions de l'unité dont les fonctions poids sont régulières et à support compact.

11. Le phénomène de dispersion ou d'étalement du paquet d'ondes, est un point difficile à interpréter relativement à la question de la localisation des corpuscules (comme l'électron) que le paquet d'ondes est censé représenter dans le modèle quantique.

3. Incertitude et étalement du paquet d'ondes

3.1. Si $\psi_t(x) = \psi(t,x)$ est un paquet d'ondes au sens de la Définition 2.1, alors (c.f. Proposition 2.4) $\mu_t(dx) := |\psi_t(x)|^2 dx$ et $\eta(dk) = |\hat{\psi}_t(k)|^2 dk$ sont deux mesures de probabilités sur $\mathbb{R}$. Nous savons que si μ_t et η_t possèdent des moments d'ordre 1, alors les écarts types

$$\sigma(\mu_t) := \sqrt{\int x^2 |\psi_t(x)|^2 dx - \left(\int x|\psi_t(x)|^2 dx\right)^2}$$

et

$$\sigma(\eta_t) := \sqrt{\int k^2 |\hat{\psi}_t(k)|^2 dk - \left(\int k|\hat{\psi}_t(k)|^2 dk\right)^2}$$

satisfont l'inégalité de Heisenberg-Gabor (c.f. [Oli16a, § 10]), soit :

$$\sigma(\mu_t)\sigma(\eta_t) \geq \frac{1}{2} \tag{6}$$

L'inégalité ainsi obtenue est valable à tout instant t : il est alors légitime de se demander comment les écarts types $\sigma(\mu_t)$ et $\sigma(\eta_t)$, ainsi que le produit $\sigma(\mu_t)\sigma(\eta_t)$ évoluent au cours du temps. Nous avons déjà noté (c.f. Remark 2.3) que l'hétérogénéité des vitesses de phase des composantes monochromatiques d'un paquet d'ondes est à l'origine de *« l'étalement du paquet d'ondes »* : intuitivement, chaque composante monochromatique se propage à sa propre vitesse de phase. Nous allons expliciter cela dans le cas particulier (mais important) des *« paquets d'ondes gaussiens »*, c'est-à-dire du type

$$\psi_t(x) = \psi(t,x) := \frac{K}{\sqrt{2\pi}} \int e^{-\alpha(k-k_0)^2} e^{-\boldsymbol{i}(kx-\omega(k)t)} dk \tag{7}$$

avec $\alpha > 0$ et $k_0 \in \mathbb{R}$ fixés ; ici $K > 0$ est la constante de normalisation assurant que $\|\psi_t\| = 1$ pour $t = 0$ et donc pour tout $t \geq 0$ (c.f. Proposition 2.4). Nous considérons aussi que $\omega(k)$ se réduit à son développement de Taylor en k_0 au second ordre écrit sous la forme

$$\omega(k) = \omega_0 + v_g(k-k_0) + \beta(k-k_0)^2 \tag{8}$$

et où v_g est la vitesse de groupe au voisinage de k_0. Nous allons écrire $\psi(t,x)$ en (7) en fonction des paramètres v_g et β, de sorte qu'en posant $\kappa = k - k_0$ il vient :

$$\begin{aligned}
\psi(t,x) &= \frac{K}{\sqrt{2\pi}} \int e^{-\alpha(k-k_0)^2} e^{-\boldsymbol{i}\left(k_0x+(k-k_0)x-(\omega_0+v_g(k-k_0)+\beta(k-k_0)^2)t\right)} dk \\
&= \frac{K}{\sqrt{2\pi}} e^{\boldsymbol{i}(k_0x-\omega_0 t)} \int e^{-\alpha\kappa^2} e^{\boldsymbol{i}\left(\kappa x-(v_g\kappa+\beta\kappa^2)t\right)} d\kappa
\end{aligned}$$

soit encore

$$\psi(t,x) = \frac{K}{\sqrt{2\pi}} e^{\boldsymbol{i}(k_0x-\omega_0 t)} \int e^{-(\alpha-\boldsymbol{i}\beta)\kappa^2} e^{\boldsymbol{i}\kappa(x-v_g t)} d\kappa \tag{9}$$

Mais en utilisant la valeur des intégrales gaussiennes (à écart type complexe : c.f. [Oli16a, § 6.3]), nous obtenons directement que

$$\psi(t,x) = \frac{K}{\sqrt{2(\alpha-\boldsymbol{i}\beta)}}\, e^{i(k_0x-\omega_0 t)}\, e^{-\frac{1}{4}\frac{\alpha+i\beta t}{\alpha^2+\beta^2t^2}(x-v_g t)^2} \tag{10}$$

Le module carré du paquet gaussien $\psi(t,x)$ s'écrivant

$$|\psi(t,x)|^2 = \sqrt{2\pi\left(\frac{\alpha^2+\beta^2t^2}{\alpha}\right)}\, e^{-\frac{1}{2}\frac{\alpha(x-v_g t)^2}{\alpha^2+\beta^2t^2}} \tag{11}$$

nous reconnaissons directement que $\mu_t(dx) := |\psi(t,x)|^2 dx$ est une *« distribution gaussienne »* centrée en $v_g t$ et d'écart type

$$\sigma(\mu_t) = \sqrt{\frac{\alpha^2 + \beta^2 t^2}{\alpha}} \tag{12}$$

Lorsque $\beta = 0$ l'écart type $\sigma(\mu_t) = \alpha$ est indépendant du temps : dans ce cas le paquet d'onde est pur (i.e. il possède une vitesse de phase) ; par contre, lorsque $\beta \neq 0$, nous avons $\sigma(\mu_t) \sim |\beta| t$ pour $t \to +\infty$ ce qui correspond à une dispersion en position : on retrouve ici le phénomène *« d'étalement du paquet d'ondes »* déjà mentionné.

3.2. Le paquet d'ondes gaussien tel qu'il est décomposé en (9) s'écrit aussi :

$$\psi(t,x) = K e^{\boldsymbol{i}(k_0(x - v_g t) + (k_0 v_g - \omega_0)t} \frac{1}{\sqrt{2\pi}} \int e^{-(\alpha - \boldsymbol{i}\beta t)\kappa^2} e^{\boldsymbol{i}\kappa(x - v_g t)} d\kappa$$

Si nous posons $\gamma(\kappa) := e^{-(\alpha - \boldsymbol{i}\beta)\kappa^2}$, alors en utilisant la transformée de transformée de Fourier (en synthèse) nous obtenons :

$$\psi(t,x) = K e^{\boldsymbol{i}(k_0 v_g - \omega_0)t)} \Big(e^{\boldsymbol{i}k_0(x - v_g t)} \mathfrak{F}_+[\gamma](x - v_g t) \Big)$$

Par suite, si $\hat{\psi}_t(k) = \hat{\psi}(t,k) := \mathfrak{F}_-[\psi_t](k)$ alors nous avons successivement

$$\begin{array}{rcl} \hat{\psi}(t,k) & = & K e^{\boldsymbol{i}(k_0 v_g - \omega_0)t} \mathfrak{F}_-[e^{\boldsymbol{i}k_0(x - v_g t)} \mathfrak{F}_+[\gamma](x - v_g t)](k) \\ (\text{translation en } x) & = & K e^{\boldsymbol{i}(k_0 v_g - \omega_0)t} e^{-\boldsymbol{i}k v_g t} \mathfrak{F}_-[e^{\boldsymbol{i}k_0 x} \mathfrak{F}_+[\gamma](x)](k) \\ (\text{modulation en } k) & = & K e^{\boldsymbol{i}(v_g(k_0 - k) - \omega_0)t)} \mathfrak{F}_-[\mathfrak{F}_+[\gamma]](k - k_0) \end{array}$$

ce qui donne

$$\hat{\psi}(t,k) = K e^{\boldsymbol{i}(v_g(k_0 - k) - \omega_0)t)} \gamma(k - k_0) = K e^{\boldsymbol{i}(v_g(1 - k) - \omega_0)t)} e^{-(\alpha - \boldsymbol{i}\beta)(k - k_0)^2}$$

Par suite,

$$|\hat{\psi}(t,k)|^2 = K^2 |\gamma(k - k_0)|^2 = K^2 e^{-2\alpha(k - k_0)^2} = \sqrt{\frac{\pi}{2\alpha}} e^{-\frac{1}{2} 4\alpha(k - k_0)^2}$$

La distribution gaussienne $\eta_t(dk) = |\hat{\psi}(t,k)|^2 dk$ est donc centrée en k_0 et d'écart type

$$\sigma(\eta_t) = \frac{1}{2} \frac{1}{\sqrt{\alpha}} \tag{13}$$

En combinant (12) et (13) nous obtenons directement la valeur exacte du produit des écarts types $\sigma(\mu_t)$ et $\sigma(\eta_t)$, soit :

$$\sigma(\mu_t)\sigma(\eta_t) = \frac{1}{2} \sqrt{1 + \left(\frac{\beta}{\alpha}\right)^2 t^2} \tag{14}$$

Lorsque le paquet d'ondes $\psi(t,x)$ est gaussien avec $\beta = 0$, le cas d'égalité de l'inégalité de Heisenberg-Gabor en (6) est réalisé pour tout $t \geq 0$ (et il n'y a pas d'étalement). Par contre, lorsque $\beta \neq 0$, alors nous savons d'après (12) que $\sigma(\mu_t)$ est asymptotiquement proportionnel à t quand $t \to +\infty$ (i.e. il y a étalement du paquet d'ondes) et le cas d'égalité de l'inégalité de Heisenberg-Gabor n'a lieu que pour $t = 0$.

4. Les quanta et la dualité onde-corpuscule de de Broglie

4.1. La mécanique quantique commence en 1901 avec la tentative de Planck pour interpoler une série de lois empiriques décrivant le rayonnement du corps noir (loi de Wien, loi de Stefan-Bolztmann) : Planck en déduit que la différence entre deux niveaux quelconques d'énergie d'un oscillateur harmonique de *« fréquence »* ν est nécessairement un multiple entier du *« quantum d'énergie »* $h\nu$, où $h = 6.626070040 \cdot 10^{-34}\,\mathrm{J \cdot s}$ est maintenant appelée *« la constante de Planck »*. Dans la suite nous écrirons $h\nu = \hbar\omega$ où $\omega = 2\pi\nu$ est la *« pulsation »* (fréquence angulaire) associée à la fréquence ν et où $\hbar = h/(2\pi)$ est *« la constante de Planck réduite »* : par suite, étant donnés E et E' deux niveaux d'énergie de l'oscillateur de pulsation ω (avec $E < E'$), il existe un entier $n \geq 1$ t.q.

$$E' - E = n\hbar\omega \tag{15}$$

Si nous notons E_0 le niveau d'énergie de l'état fondamental de l'oscillateur (i.e. le plus petit niveau d'énergie possible), alors tout niveau d'énergie s'écrit $E_n = E_0 + n\hbar\omega$ pour $n \geq 0$ entier. Contrairement à l'intuition classique, le niveau fondamental E_0 n'est pas nul : nous verrons (c.f. Lemme 8.2 infra) que $E_0 = \hbar\omega/2$ comme une conséquence directe du *« principe d'incertitude de Heisenberg »*. En tant qu'objet *« macroscopique »* [12], l'oscillateur harmonique classique possède un continuum de niveau d'énergie allant de 0 à l'infini : par conséquent, le rapport entre son énergie et les fréquences usuelles d'oscillations est énorme devant la constant de Planck.

4.2. Considérons un corpuscule $(\mathcal{P})$ de masse m (positive où nulle) se déplaçant *« librement »* à la vitesse v dans un référentiel $(\mathcal{R})$ lorentzien de l'espace-temps de la relativité restreinte. Le *« facteur de Lorentz »* associé à $(\mathcal{P})$ dans $(\mathcal{R})$ est par définition

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{16}$$

où $\beta = v/c$ est la vitesse réduite. L'énergie $\mathcal{E}$ et l'impulsion p de ce corpuscule s'écrivent :

$$\mathcal{E} = \gamma mc^2 \qquad \text{et} \qquad p = \gamma mv \tag{17}$$

ce qui donne une expression de p indépendante de m, soit :

$$p = \mathcal{E}v/c^2 \tag{18}$$

Proposition 4.1. *Soit $(\mathcal{P})$ un corpuscule libre de la relativité restreinte, de masse $m \geq 0$ et d'énergie positive (finie) ; alors $m = 0$ ssi la vitesse de $(\mathcal{P})$ dans tout référentiel lorentzien vaut c.*

Preuve. Soit $(\mathcal{P})$ un corpuscule libre de masse $m \geq 0$ et d'énergie $\mathcal{E} > 0$ dans un référentiel lorentzien $(\mathcal{R})$ donné. Alors, nous savons d'après l'expression de l'énergie en (17) que

$$m = \frac{1}{\gamma}\frac{\mathcal{E}}{c^2}$$

or la quantité $\mathcal{E}/c^2$ étant (strictement) positive, l'annulation de la masse m équivaux à l'annulation de l'inverse $1/\gamma$ du facteur de Lorentz définie en (16) : or cette dernière condition (i.e. $1/\gamma = 0$) signifie exactement que la vitesse $|v|$ de $(\mathcal{P})$ dans $(\mathcal{R})$ est égale

12. A ma connaissance, il n'y pas de critères objectifs permettant de décider quand un objet est macroscopique. Dans tous les cas, « *un chat* » est considéré par Schrödinger comme une « *entité macroscopique* »

à c (dans ce cas la vitesse de $(\mathcal{P})$ coïncide avec c dans tout référentiel lorentzien – à cause du statut spécial de c dans la relativité restreinte).

□

4.3. Dans son article sur *« l'effet photo-électrique »* [Ein05], Einstein définit un *« quantum de lumière »* [13] comme un corpuscule libre de la relativité restreinte de masse nulle (i.e. de vitesse c) associé – en un sens – à une vibration de pulsation ω et dont l'énergie $\mathcal{E}$ est égale à un *« quantum d'énergie »*, ce qui se traduit par la relation de *« Planck-Einstein »*, soit [14] :

$$\mathcal{E} = \hbar\omega \tag{19}$$

Maintenant, si nous définissons le *« nombre d'onde »*

$$k = \frac{\omega}{c} \tag{20}$$

alors le quantum de lumière d'énergie $\hbar\omega$, vu comme un grain de lumière [15] de vitesse $v = c$, peut (formellement) être associé à une impulsion, puisque d'après (18), (19) et (20)

$$p = \frac{\mathcal{E}}{c} = \frac{\hbar\omega}{c} = \hbar k$$

En résumé, *« l'énergie-impulsion »* du quantum de lumière de pulsation ω s'écrit :

$$(\mathcal{E}, p) = \hbar\,(\omega, k) \tag{21}$$

4.4. Dans trois notes parues en 1923 aux Comptes Rendus de l'Académie des Sciences, puis dans sa thèse en 1924 [dBL25], de Broglie considère (en particulier) une description ondulatoire d'un corpuscule de masse $m > 0$ (typiquement un électron) se déplaçant librement dans l'espace avec une vitesse $v_0 < c$. Par analogie avec la théorie des quanta de lumière (corpuscule non massif et donc de vitesse c), de Broglie associe chaque couple d'énergie-impulsion (E_0, p_0) à une vibration caractérisée par le couple pulsation-nombre d'onde (ω_0, k_0) de sorte à assurer les *« identités de de Broglie »*, soient :

$$(E_0, p_0) = \hbar\,(\omega_0, k_0) \tag{22}$$

Cela l'amène à associer le couple énergie-impulsion (E_0, p_0) à une onde monochratique appelée *« onde de de Broglie »* et que nous écrivons (en dimension 1) :

$$\psi(t,q) = Ae^{-\boldsymbol{i}(k_0 q - \omega_0 t)} = A^{(p_0 q - E_0 t)/\hbar} \tag{23}$$

où A est un nombre complexe arbitraire. Notons au passage qu'une onde de de Broglie est en particulier une onde dont *« la vitesse de phase »* est donnée par la *« relation de dispersion »*

$$u_0 = \frac{\omega_0}{k_0} = \frac{E_0}{p_0}$$

13. Le terme de *« photon »* pour désigner un quantum de lumière a été introduit indépendamment par Wolfer et Lewis en 1926.

14. Ici la pulsation ω est une caractéristique du quantum de lumière relativement au référentiel lorentzien dans lequel celui-ci est observé : c.f. effet Döppler.

15. Newton parlait de corpuscule de lumière : il était en effet un des premiers défenseurs de la théorie corpusculaire de la lumière.

4.5. Pour un corpuscule de masse $m > 0$ *« évoluant librement »* à la vitesse v_0 *« non relativiste »* (i.e. $v_0 \ll c$), une interprétation possible de l'approche de Broglie consiste à identifier l'énergie du couple *« énergie-impulsion »* en (22) à l'énergie cinétique du corpuscule (massif) et d'identifier son impulsion à l'impulsion classique, ce qui donne :

$$(E_0, p_0) = \left(\frac{p_0^2}{2m}, mv_0\right) \tag{24}$$

L'onde monochromatique (23) ne peut décrire le corpuscule non relativiste (et donc massif) pour deux raisons principales. D'une part, il n'est pas possible de localiser l'onde de de Broglie dans l'espace : cela est inconsistant avec l'idée même de corpuscule. D'autre part, l'onde monochromatique (23) est bien attachée à une vitesse ondulatoire – la vitesse de phase $u_0 = \omega_0/k_0 = E_0/p_0$ – mais celle-ci ne coïncide pas avec la *« vitesse (supposée) réelle »* du corpuscule : en effet, pour l'énergie-impulsion en (24) nous obtenons

$$u_0 = \frac{\omega_0}{k_0} = \frac{E_0}{p_0} = \frac{1}{p_0}\frac{p_0^2}{2m} = \frac{p_0}{2m} = \frac{v_0}{2}$$

Pour sortir de cette difficulté de Broglie introduit la notion de paquet d'ondes, i.e. la version quantique des paquets d'ondes de la théorie du signal t.q. définie [16] au § 2.1.

Définition 4.2. *Soit $A(p) \in L^2(\mathbb{R})$ t.q. $\|A\| = 1$; alors le « paquet d'ondes de de Broglie » d'amplitude complexe $A(p)$ est la fonction $\psi(t,q)$ dépendant du temps t et de l'espace q t.q.*

$$\psi(t,q) = \frac{1}{\sqrt{2\pi\hbar}} \int A(p) e^{i(pq - E(p)t)/\hbar} dp \quad \text{où} \quad E(p) = \frac{p^2}{2m} \tag{25}$$

A chaque composante monochromatique d'énergie-impulsion $(E(p), p)$ de $\psi(t,q)$ correspond un couple pulsation-nombre d'onde $(\omega(k), k)$ déterminé par « les relations de de Broglie », soient

$$\big(\omega(k), k\big) = \frac{1}{\hbar}\big(E(p), p\big) \tag{26}$$

ce qui permet d'écrire le paquet d'onde en (25) sous la forme

$$\psi(t,q) = \frac{1}{\sqrt{2\pi\hbar}} \int A(\hbar k) e^{i(kq - \omega(k)t)} dk \tag{27}$$

Par anticipation sur l'interprétation probabiliste de Born (c.f. § 5.5), nous avons aussi imposé au paquet d'ondes de de Broglie de vérifier la condition de normalisation $\|A\| = 1$. A l'instar des paquets d'ondes de la théorie du signal, le facteur $1/\sqrt{2\pi\hbar}$ est introduit afin d'assurer la compatibilité avec la *« transformée de Fourier quantique »* : par définition, la *« Q-transformée de Fourier »* est l'application qui à toute fonction $f \in L^2(\mathbb{R})$ associe la fonction $\mathfrak{F}_\pm[f]$, définie de manière analogue au cas standard (c.f. [Oli16a, §§ 5 & 6]), de sorte que

$$\mathfrak{F}_-[f](p) = \frac{1}{\sqrt{2\pi\hbar}} \int f(q) e^{-ipq/\hbar} dq \quad \text{et} \quad \mathfrak{F}_+[f](q) = \frac{1}{\sqrt{2\pi\hbar}} \int f(p) e^{ipq/\hbar} dp$$

16. Nous insistons sur le fait que la notion de paquet d'ondes de la théorie du signal est postérieure à la notion introduite par de Broglie en mécanique ondulatoire (et dans laquelle intervient – de manière essentielle – l'échelle de Planck) : en commençant par la théorie du signal, nous avons choisi une présentation anachronique afin d'extraire un certain nombre d'idées mathématiques du contexte (parfois très spéculatif) de la représentation quantique du réel.

Il découle du théorème de Plancherel que la Q-transformée de Fourier $\mathfrak{F}_-[\cdot]$ est une isométrie unitaire de $L^2(\mathbb{R})$ dont l'inverse (i.e. son adjoint) est $\mathfrak{F}_+[\cdot]$. Maintenant, si nous définissons le *« paquet d'ondes en impulsion »* associée à $\psi(t,q)$ en (25), soit :

$$\hat{\psi}(t,p) := A(p)e^{-\boldsymbol{i}E(p)t/\hbar}$$

alors les fonctions partielles $\psi_t = \psi(t,\cdot)$ et $\hat{\psi}_t = \hat{\psi}(t,\cdot)$ forment (pour tout t) un couple de Q-transformée de Fourier puisque

(28) $$\psi_t(q) = \frac{1}{\sqrt{2\pi\hbar}}\int \hat{\psi}_t(p)e^{\boldsymbol{i}pq/\hbar}dp \quad \text{et} \quad \hat{\psi}_t(p) = \frac{1}{\sqrt{2\pi\hbar}}\int \psi_t(q)e^{-\boldsymbol{i}pq/\hbar}dx$$

De manière plus condensée nous avons $\psi(t,q) = \mathfrak{F}_+\left[\hat{\psi}_t\right](q)$ et $\hat{\psi}(t,p) = \mathfrak{F}_-[\psi_t](p)$; en particulier, à tout instant t nous avons :

$$\|\psi_t\|^2 = \left\langle \mathfrak{F}_+\left[\hat{\psi}_t\right] \middle| \mathfrak{F}_+\left[\hat{\psi}_t\right]\right\rangle = \left\langle \hat{\psi}_t \middle| \hat{\psi}_t\right\rangle = \|A\|^2$$

Or (définition d'un paquet d'ondes de de Broglie) $\|A\| = 1$ et par suite (de même que pour les paquets d'ondes de la théorie du signal : c.f. Proposition 2.4) nous avons :

Proposition 4.3. *Soit $\psi(t,q) = \mathfrak{F}_+[\hat{\psi}_t]$ le paquet d'onde de de Broglie t.q.*

$$\hat{\psi}_t(p) = A(p)^{-\boldsymbol{i}E(p)t/\hbar} \quad \textit{avec} \quad E(p) = \frac{p^2}{2m}$$

et où $A(p) \in L^2(\mathbb{R})$ (avec $\|A\| = 1$) ; si ψ_t est la fonction partielle $q \mapsto \psi(t,q)$ alors pour tout $t \geq 0$:

$$\|\psi_t\| = \|\hat{\psi}_t\| = \|A\| = 1$$

4.6. A tout instant t, le paquet d'ondes ψ_t est un élément de la sphère unité de $L^2(\mathbb{R})$: cela signifie que la mesure $|\psi_t(q)|^2dq$ est une probabilité sur $\mathbb{R}$. Afin de localiser le corpuscule libre représenté par ψ_t, il faut pouvoir définir *« sa position »* $\boldsymbol{q}(t)$ comme *« le centre d'inertie »* (i.e. *« la moyenne »*) de la probabilité $|\psi_t(q)|^2dq$; nous verrons que le *« théorème d'Ehrenfest »* (c.f. Théorème 6.4) assure en particulier que la dérivée temporelle $\dot{\boldsymbol{q}}(t)$ coïncide avec la vitesse v_0 du corpuscule. Nous reviendrons plus loin (c.f. §§ 5 & 6) sur ce point dans le cadre plus général de la mécanique ondulatoire des fonctions d'ondes. Pour l'instant, nous pouvons faire une remarque intéressante dans le cas des paquets d'ondes gaussiens, c'est-à-dire avec une amplitude complexe de la forme

$$A(p) = e^{-\alpha(p-p_0)^2/\hbar^2}$$

avec $\alpha > 0$. Cela nous donne l'analogue quantique du paquet d'onde en (7) soit [17]

$$\psi(t,q) = \frac{K}{\sqrt{2\pi\hbar}}\int e^{-\alpha(k-k_0)^2}e^{-\boldsymbol{i}(kq-\omega(k)t)}dp \quad \text{avec} \quad \omega(k) = \frac{E(\hbar k)}{\hbar} = \frac{\hbar k^2}{2m}$$

(où K est la constante de normalisation). En développant $\omega(k)$ en k_0 il vient :

$$\omega(k) = \frac{\hbar}{2m}\big(k_0 + (k-k_0)\big)^2 = \frac{\hbar}{2m}k_0^2 + \frac{\hbar}{m}k_0(k-k_0) + \frac{\hbar}{2m}(k-k_0)^2$$

soit encore, avec l'énergie cinétique $E_0 = \frac{(\hbar k_0)^2}{2m}$ et de la vitesse $v_0 = \frac{d\omega}{dk}(k_0) = \frac{\hbar k_0}{m}$:

$$\omega(k) = \frac{E_0}{\hbar} + v_0(k-k_0) + \frac{\hbar}{2m}(k-k_0)^2$$

17. Exemple le plus simple d'un paquet d'onde localisé autour du nombre d'onde $k_0 = p_0/\hbar = mv_0/\hbar$.

Des calculs identiques à ceux faits au § 3.1 donnent

$$\psi(t,q) = \frac{K}{\sqrt{2\pi\hbar}} e^{\boldsymbol{i}(p_0 q - E_0 t)/\hbar} \int e^{-(\alpha - \boldsymbol{i}\frac{\hbar}{2m})k^2} e^{\boldsymbol{i}k(q-v_0 t)} dk$$

d'où nous tirons que

$$|\psi(t,q)|^2 = \sqrt{2\pi\sigma^2(t)}\, e^{\frac{-(q-v_0 t)^2}{2\sigma^2(t)}} \quad \text{avec} \quad \sigma(t) = \sqrt{\alpha + \frac{1}{\alpha}\left(\frac{\hbar t}{2m}\right)^2}$$

Les propriétés statistiques de la distribution de Gauss nous permettent de conclure que *l'énergie* du paquet d'onde gaussien $\psi(t,q)$ (représentée par la probabilité $|\psi(t,q)|^2 dq$) est localisée au point $v_0 t$, c'est-à-dire la position classique au temps t d'un corpuscule de vitesse v_0 initialement situé en 0. (La présence de la constante de Planck dans l'expression de l'écart type $\sigma(t)$ nous donne un ordre de grandeur du phénomène de dispersion du paquet d'onde à l'échelle microscopique).

4.7. Laissant (provisoirement) de côté la question de la position, nous nous intéressons au problème de la vitesse de corpuscule libre modélisée par un paquet d'ondes de de Broglie à partir de la notion de *« vitesse de groupe »*. Pour voir cela, nous partons du fait que les composantes monochromatiques du paquet d'ondes de de Broglie ψ_t en (27) sont caractérisées par le nombre d'onde $k = p/\hbar$ et la pulsation

$$\omega(k) = \frac{\hbar k^2}{2m}$$

Maintenant, si nous supposons que la fonction $k \mapsto A(\hbar k)$ est centrée autour d'une valeur donnée $k_0 = \hbar p_0 = \hbar m v_0$ du nombre d'onde, alors le paquet d'ondes de de Broglie ψ_t possède une vitesse de groupe (c.f. § 2.2) soit :

$$v_g = \left.\frac{d}{dk}\right|_{k=k_0} \omega(k) = \left.\frac{d}{dk}\right|_{k=k_0} \left(\frac{\hbar k^2}{2m}\right) = \frac{\hbar k_0}{m} = \frac{\hbar(p_0/\hbar)}{m} = v_0$$

La notion de paquet d'ondes de de Broglie permet de répondre de manière consistante aux deux obstructions quant à la représentation ondulatoire du corpuscule (non relativiste) de masse m et de vitesse v_0 : en effet, un tel paquet d'ondes est spatialement localisable et sa vitesse de groupe coïncide avec *« la vitesse réelle »* v_0 du corpuscule en question. Le formalisme de l'onde-corpuscule de de Broglie n'est devenu réellement crédible au yeux de la majorité des physiciens qu'à partir du moment où la diffraction des électrons à été effectivement observée [18], exactement comme prédite par de Broglie .

5. **Mécanique ondulatoire – Equation de Schrödinger**

5.1. Les paquets d'ondes tels que considérés par de Broglie vont amener Schrödinger à décrire la *« nouvelle mécanique des quanta »* sous la forme d'une *« mécanique ondulatoire »* à partir de l'approche de Hamilton et Jacobi de la mécanique classique. Dans ce nouveau cadre, la notion de *« fonction d'onde »* intervenant dans *« l'équation de Schrödinger »* (c.f. Définition 5.1 infra) étend la notion de paquet d'ondes et permet d'intégrer les idées/intuitions de de Broglie sur la *« dualité ondes-corpuscules »* dans un système théorique autorisant l'introduction des champs de forces (principe de correspondance : c.f.

18. De manière indépendante par G. P. Thompson and C. J. Davisson en 1927 : pour cette découverte, ils ont partagé le prix Nobel de physique de 1937.

§ 5.2). Schrödinger peut ainsi traiter (c.f. [Sch26]) le cas de l'électron (non relativiste) pris dans le champ coulombien d'un proton et calculer avec une grande précision le spectre d'émission-absortion de l'atome d'hydrogène ; il s'agit là d'une avancée théorique sans précédent depuis la « *révolution* » du modèle « *ad hoc* » de Bohr [19]. La question de la localisation des ondes-corpuscules le long de leurs trajectoires quantiques est un point particulièrement important : nous allons voir que dans la mécanique ondulatoire, la nature unitaire et localisable de la fonction d'onde (généralisation de la Proposition 2.4 pour les paquets d'ondes) peut se déduire du « *théorème d'Ehrenfest* » (nous y reviendrons au § 6.3) : le fait que l'évolution quantique d'une fonction d'onde corresponde à l'évolution d'une distribution de probabilité spatiale, constitue l'interprétation probabiliste de la mécanique quantique due à Born.

5.2. Nous avons déjà mentionné (qu'en général) le paquet d'ondes de de Broglie $\psi(t,q)$ tel que défini en (25) n'est pas une onde, en ce sens qu'il ne possède (généralement) pas de vitesse de phase (il ne peut donc être solution de l'équation des ondes). Cependant $\psi(t,q)$ satisfait une équation où la dérivation temporelle est du premier ordre. Par dérivation sous le signe intégral nous obtenons d'une part

$$\frac{\partial^2\psi}{\partial q^2} = \int \frac{\partial^2}{\partial q^2}\left(\frac{A(p)}{\sqrt{2\pi\hbar}}e^{i(pq-E(p)t)/\hbar}\right)dp = -\frac{1}{\hbar^2}\int p^2\frac{A(p)}{\sqrt{2\pi\hbar}}e^{i(pq-E(p)t)/\hbar}dp \tag{29}$$

et d'autre part (en utilisant le fait que $E(p) = \frac{p^2}{2m}$),

$$\begin{aligned} i\hbar\frac{\partial\psi}{\partial t} &= i\hbar\int\frac{\partial}{\partial t}\left(\frac{A(p)}{\sqrt{2\pi\hbar}}e^{i(pq-E(p)t)/\hbar}\right)dp \\ &= i\hbar\int -\frac{i}{\hbar}\frac{E(p)A(p)}{\sqrt{2\pi\hbar}}e^{i(pq-E(p)t)/\hbar}dp = \int\frac{p^2}{2m}\frac{A(p)}{\sqrt{2\pi\hbar}}e^{i(pq-E(p)t)/\hbar}dp \end{aligned}$$

Au final, en combinant avec (29) il vient :

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial q^2} \tag{30}$$

L'équation (30) est « *l'équation de Schrödinger* » du corpuscule libre de masse m.

5.3. Pour introduire la dynamique quantique, considérons maintenant un *corpuscule* de masse m pris dans un champ de force conservatif dérivant d'un potentiel $V(q)$: si p désigne la variable d'impulsion associée à la variable de position q, alors le hamiltonien du système mécanique est la fonction

$$H(q,p) = \frac{p^2}{2m} + V(q) \tag{31}$$

(somme de « *l'énergie cinétique* » et de « *l'énergie potentielle* » du corpuscule dans le champ de force). Dans le cas considéré, la valeur du hamiltonnien est constante le long d'une trajectoire classique $\boldsymbol{q}(t)$ en ce sens que $H(\boldsymbol{q}(t), m\dot{\boldsymbol{q}}(t))$ est une quantité ne dépendant pas du temps (nous reviendrons sur les formalismes lagrangien et hamiltonien dans [Oli17]). Il est difficile de se faire une idée précise de la manière dont Schrödinger est parvenu à

19. Pour cela, Schrödinger obtient le prix Nobel en 1933 qu'il partage avec Dirac (Dirac traite l'électron relativiste et explique la structure fine du spectre de l'atome d'hydrogène).

l'équation qui porte son nom [20]. La justification théorique de Schrödinger est essentiellement basée sur l'équation de Hamilton-Jacobi ; il est cependant envisageable que cette présentation soit venue justifier à postériori une équation dont la valeur s'est d'abord révélée à Schrödinger par sa prodigieuse efficacité à prédire le spectre de l'atome d'hydrogène. Nous proposons une heuristique obtenue par analogie avec celle du § 5.2 où nous avons obtenu l'équation de Schrödinger (30) satisfaite par les paquets d'ondes associés aux corpuscules libres de de Broglie. Pour cela nous allons remplacer la phase $pq - E(p)t$ de l'onde de Broglie par une *« fonction d'action »* $S(t, q, p)$ issue de l'approche classique de Hamiltons et Jacobi et dont nous admettrons [21] qu'elle vérifie le système d'équations [22]

$$(\mathbf{HJ1}) \; : \; \frac{\partial S}{\partial t} = -H(q,p) \qquad \text{et} \qquad (\mathbf{HJ2}) \; : \; \frac{\partial S}{\partial q} = p$$

(L'équation $(\mathbf{HJ1})$ est précisément une forme de *« l'équation de Hamilton-Jacobi »* proprement dite.) Nous pouvons (formellement) considérer un *« paquet d'ondes de de Broglie-Schrödinger »* écrit sous la forme

$$\psi(t,q) = \frac{1}{\sqrt{2\pi\hbar}} \int A(p) e^{\boldsymbol{i}S(t,q,p)/\hbar} dp \tag{32}$$

D'une part, en dérivant partiellement $\boldsymbol{i}\hbar\psi(t,q)$ par rapport au temps, il vient avec $(\mathbf{HJ1})$:

$$\begin{aligned} \boldsymbol{i}\hbar\frac{\partial\psi}{\partial t} &= \boldsymbol{i}\hbar \int \frac{\partial}{\partial t}\left(\frac{A(p)}{\sqrt{2\pi\hbar}} e^{\boldsymbol{i}S(t,q,p)/\hbar}\right) dp \\ &= \boldsymbol{i}\hbar \int \frac{\boldsymbol{i}}{\hbar}\frac{\partial S}{\partial t}\frac{A(p)}{\sqrt{2\pi\hbar}} e^{\boldsymbol{i}S(t,q,p)/\hbar} dp = \int H(q,p)\frac{A(p)}{\sqrt{2\pi\hbar}} e^{\boldsymbol{i}S(t,q,p)/\hbar} dp \end{aligned}$$

soit encore, en introduisant l'expression du hamiltonien en (31) :

$$\boldsymbol{i}\hbar\frac{\partial\psi}{\partial t} = \int \left(\frac{p^2}{2m} + V(q)\right)\frac{A(p)}{\sqrt{2\pi\hbar}} e^{\boldsymbol{i}S(t,q,p)/\hbar} dp \tag{33}$$

Mais d'autre part, en dérivant partiellement $\psi(t,q)$ deux fois par rapport à q, nous avons [23] :

$$\begin{aligned} \frac{\partial^2\psi}{\partial q^2} &= \frac{\partial}{\partial q}\int \frac{\partial}{\partial q}\left(\frac{A(p)}{\sqrt{2\pi\hbar}} e^{\boldsymbol{i}S(t,q,p)/\hbar}\right) dp \\ &= \frac{\partial}{\partial q}\int \frac{\boldsymbol{i}}{\hbar}\frac{\partial S}{\partial q}\frac{A(p)}{\sqrt{2\pi\hbar}} e^{\boldsymbol{i}S(t,q,p)/\hbar} dp \\ (\partial^2 S/\partial q^2 = 0) \quad &= \frac{\boldsymbol{i}}{\hbar}\int \frac{\partial}{\partial q}\left(\frac{\partial S}{\partial q}\frac{A(p)}{\sqrt{2\pi\hbar}} e^{\boldsymbol{i}S(t,q,p)/\hbar}\right) dp = -\frac{1}{\hbar^2}\int \left(\frac{\partial S}{\partial q}\right)^2 \frac{A(p)}{\sqrt{2\pi\hbar}} e^{\boldsymbol{i}S(t,q,p)/\hbar} dp \end{aligned}$$

20. A ce propos il est instructif de citer une réflexion bien connue de Feynman à ce propos : « *D'où vient [l'équation de Schrödinger] ? He bien vous ne pouvez la déduire de quoi que ce soit que vous connaissiez. Elle est sortie de l'esprit de Schrödinger !* ».

21. Nous reviendrons sur ce point en détail dans [Oli17].

22. Dans le cas du corpuscule libre, nous avons $H(p,q) = E(p)$ (le hamiltonien est réduit à l'énergie cinétique) et $S(t,q,p) = pq - E(p)t$ est une fonction qui satisfait bien les équations $(\mathbf{HJ1})$ et $(\mathbf{HJ2})$.

23. Nous utilisons ici le fait que d'après $(\mathbf{HJ2})$:

$$\frac{\partial^2 S}{\partial^2 q} = \frac{\partial}{\partial q}\frac{\partial S}{\partial q} = \frac{\partial p}{\partial q} = 0$$

Grâce à l'équation (**HJ2**), nous obtenons alors un analogue presque identique à (29) soit

$$\frac{\partial^2\psi}{\partial q^2} = -\frac{1}{\hbar^2}\int p^2 \frac{A(p)}{\sqrt{2\pi\hbar}} e^{\boldsymbol{i}S(t,q,p)/\hbar}dp \tag{34}$$

Finalement, la combinaison de (33) et (34) donne l'équation

$$\boldsymbol{i}\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial q^2} - V(q)\psi(t,q) \tag{35}$$

Nous venons de retrouver *« l'équation de Schrödinger »* d'une onde-corpuscule de masse m prise dans le champ de force dérivant du potentiel $V(q)$. L'heuristique présentée est basée sur l'hypothèse (implicite ici) que le système étudié possède une distribution continue en impulsion ; cependant nous aurions pu tout aussi bien considérer une distribution $\mu(dp)$ plus générale, à la place dp et en particulier prendre pour $\mu(dp)$ une distribution discrète. Cette distinction entre discret et continue est un point important de la mécanique quantique. L'onde-corpuscule libre de de Broglie est un exemple de système quantique continu ; nous reviendrons en détail sur le cas de l'oscillateur harmonique quantique (correspondant au potentiel $V(q) = 1/2\kappa q^2$ avec κ constante de l'oscillateur) et qui est le modèle du système quantique discret.

5.4. Grâce à l'équation de Schrödinger (35) les paquets d'ondes de de Broglie-Schrödinger, tels que définis en (32), s'intègrent naturellement dans le cadre de *« la mécanique ondulatoire »*, c'est-à-dire comme un cas particulier de ce qu'on appelle une *« fonction d'onde »*.

Définition 5.1. *Une « fonction d'onde » associée au corpuscule de masse m soumis au potentiel $V(q)$ est une fonction $\psi(t,q)$ dépendant de la variable temporelle t et de la variable spatiale q, telle que (i) : ψ_t appartienne à l'espace de Hilbert $\mathscr{H} = L^2(\mathbb{R})$ pour tout t ; (ii) : $\psi(t,q)$ vérifie l'équation de Schrödinger en (35) et (iii) : il existe t_0 tel que $\|\psi_{t_0}\| = 1$.*

5.5. Nous verrons au § 6, avec les *« Q-trajectoires »* (c.f. Définition 6.2), une généralisation encore un peu plus abstraite de la notion de fonction d'onde. Pour l'instant remarquons que si $\psi_t(q) = \psi(t,q)$ est une fonction d'onde de la Définition 5.1, alors d'après (35) et en utilisant le fait que le potentiel $V(q)$ est une fonction à valeurs réelles (et les propriétés du produit scalaire hermitien sur $\mathscr{H}$) :

$$\begin{aligned}
\frac{d}{dt}\langle\psi_t|\psi_t\rangle &= \left\langle\frac{\partial\psi}{\partial t}\middle|\psi(t,\cdot)\right\rangle + \left\langle\psi(t,\cdot)\middle|\frac{\partial\psi}{\partial t}\right\rangle\\
&= \left\langle -\frac{\boldsymbol{i}}{\hbar}\left(\frac{\partial^2\psi}{\partial q^2} + V(\cdot)\psi(t,\cdot)\right)\middle|\psi(t,\cdot)\right\rangle + \left\langle\psi(t,\cdot)\middle| -\frac{\boldsymbol{i}}{\hbar}\left(\frac{\partial^2\psi}{\partial q^2} + V(\cdot)\psi(t,\cdot)\right)\right\rangle\\
&= -\frac{\boldsymbol{i}}{\hbar}\left\langle\frac{\partial^2\psi}{\partial q^2}\middle|\psi(t,\cdot)\right\rangle + \frac{\boldsymbol{i}}{\hbar}\left\langle\psi(t,\cdot)\middle|\frac{\partial^2\psi}{\partial q^2}\right\rangle
\end{aligned}$$

Mais par une double intégration par parties (et en admettant que $\psi_t(\pm\infty) = 0$), il vient

$$\left\langle\frac{\partial^2\psi}{\partial q^2}\middle|\psi(t,\cdot)\right\rangle = \left\langle\psi(t,\cdot)\middle|\frac{\partial^2\psi}{\partial q^2}\right\rangle$$

de sorte que finalement :

$$\frac{d}{dt}\|\psi_t\|^2 = 0$$

La norme (dans $\mathscr{H} = L^2(\mathbb{R})$) de la fonction ψ_t est donc indépendante de t : le fait que l'identité $\|\psi_t\| = 1$ soit vraie pour un certain t_0 entraîne qu'elle est vraie pour tout t. Nous avons ainsi vérifié l'analogue de la Proposition 4.3 dans le cas général des fonctions d'ondes au sens de la Définition 5.1.

Proposition 5.2. *Si la fonction $\psi_t(q) = \psi(t,q)$ dépendant la variable temporelle t et de la variable spatiale q est une fonction d'onde associée à l'équation de Schrödinger en (35), alors*

$$(\forall t) \qquad \int |\psi(t,q)|^2 dq = \|\psi_t\|^2 = 1$$

La Proposition 5.2 permet d'associer la notion de fonction d'onde à *« l'interprétation probabiliste de Born »* introduite dans[24] [Bor26] (voir aussi la conférence Nobel de 1954 : [Bor54]) : la probabilité $|\psi(t,q)|^2 dq$ décrit *« la probabilité de présence »* – à instant t – du corpuscule associé à la fonction d'onde $\psi(t,q)$. Ainsi, pour tout $a < b$ réel la probabilité de présence du corpuscule dans l'intervalle $[a\,;b]$ à l'instant t vaut-elle

$$\int_a^b |\psi(t,q)|^2 dt$$

5.6. Nous notons toujours ψ_t la fonction partielle spatiale associée à la fonctions d'ondes $\psi(t,q)$ de l'onde-corpuscule de masse m et soumis au champ de force dérivant du potentiel $V(q)$ (c.f. Définitions 5.1). Cela nous permet de reformuler l'équation de Schrödinger (35) sans dérivation partielle en écrivant

$$i\hbar \frac{d}{dt}\psi_t = H\psi_t \tag{36}$$

où le *« Q-hamiltonien »* H s'écrit comme la somme $H_0 + H_1$, chacun des opérateurs H_k agissant sur les fonctions complexes $f(q)$, de sorte que (sous réserve que cela a un sens)

$$H_0 f(q) = -\frac{\hbar^2}{2m}\frac{\partial^2 f}{\partial q^2} \quad \text{et} \quad H_1 f(q) = V(q)f(q) \tag{37}$$

La définition de $Hf(q)$ est claire lorsque $f(q)$ est (par exemple) dans l'espace de Schwarz $\mathscr{S}(\mathbb{R})$ avec $V(q)$ polynomial en q ; cependant, même si $\mathscr{S}(\mathbb{R})$ est dense dans $\mathscr{H}$, l'opérateur H en (37) ne possède pas d'extension bornée sur tout $\mathscr{H}$: en toute rigueur nous devons déterminer le sous-espace dense $\mathscr{D}$ de $L^2(\mathbb{R})$, contenu dans l'espace de Sobolev $\mathscr{H}^2$ et rendant l'opérateur $(H,\mathscr{D})$ autoadjoint ($\mathscr{D}$ dépend essentiellement de la forme de $V(q)$). Nous reviendrons sur cette question dans le cas de l'oscillateur harmonique quantique (c.f. § 8) ; pour l'instant nous développons les idées au niveau de l'heuristique.

5.7. La définition même d'une fonctions d'onde $\psi_t(q) = \psi(t,q)$ de l'onde-corpuscule de masse m prise dans le champ de force dérivant de $V(q)$ assure qu'à tout instant t la mesure $\mu_t(dq) = |\psi_t(q)|^2 dq$ est une distribution de probabilité sur $\mathbb{R}$. Afin de localiser le corpuscule à partir de sa fonction d'onde ψ_t, nous utilisons l'opérateur position Q (déjà introduit : c.f. [Oli16b]) et tel que

$$Qf(q) = qf(q) \tag{38}$$

24. Born à participer à la formulation de la mécanique des matrices de Heisenberg avec Bohr et Jordan ; à partir de 1923, Heisenberg et Jordan ont été ses assistants, lorsque celui-ci était professeur à Göttingen. Notons aussi que durant l'hiver 25-26 Born a travaillé avec Wiener sur l'interprétation du formalisme de Heisenberg en termes d'opérateurs linéaires agissant sur un espace vectoriel (c.f. Blog d'Alain Bouquet - CNRS). L'équation $pq - qp = \frac{h}{2\pi i}$ est gravée sur la pierre de tombale de Born à Göttingen.

Ainsi, la *« position (quantique) »* du corpuscule à l'instant t est par définition la quantité

$$\langle \psi_t | Q | \psi_t \rangle = \langle \psi_t | Q\psi_t \rangle = \int q |\psi_t(q)|^2 dq$$

(où nous reconnaissons le point moyen de la probabilité $\mu_t(dq) = |\psi_t(q)|^2 dq$). Si cette position est définie à tout instant t alors – à l'instar de la notion de vitesse de groupe (c.f. § 4.6) – il est possible de définir *« une notion de vitesse »*. En utilisant la dérivation sous le signe intégral, et l'équation de Schrödinger sous la forme (36), il vient en effet :

$$\begin{aligned} \frac{d}{dt} \langle \psi_t | Q | \psi_t \rangle &= \left\langle \frac{d}{dt}\psi_t \middle| Q\psi_t \right\rangle + \left\langle \psi_t \middle| \frac{d}{dt}(Q\psi_t) \right\rangle \\ &= \left\langle -\frac{\boldsymbol{i}}{\hbar} H\psi_t \middle| Q\psi_t \right\rangle + \left\langle \psi_t \middle| Q \frac{\partial}{\partial t}\psi_t \right\rangle = \left\langle \psi_t \middle| \frac{\boldsymbol{i}}{\hbar} HQ\psi_t \right\rangle + \left\langle \psi_t \middle| -\frac{\boldsymbol{i}}{\hbar} QH\psi_t \right\rangle \end{aligned}$$

où dans la dernière ligne du calcul, nous avons utilisé le fait que H est *« moralement autoadjoint »*. Avec la notation en crochets de Lie[25] nous pouvons écrire :

$$\frac{d}{dt} \langle \psi_t | Q | \psi_t \rangle = \left\langle \psi_t \middle| -\frac{\boldsymbol{i}}{\hbar}[Q, H] \middle| \psi_t \right\rangle \tag{39}$$

Mais le fait que le potentiel $V(q)$ soit une fonction à valeurs réelles assure que le crochet de Lie $[Q, H_1]$ est nul et finalement (39) s'écrit

$$\frac{d}{dt} \langle \psi_t | Q | \psi_t \rangle = \left\langle \psi_t \middle| -\frac{\boldsymbol{i}}{\hbar}[Q, H_0] \middle| \psi_t \right\rangle \tag{40}$$

Cette dernière expression nous incite à définir l'observable d'impulsion

$$P = -\frac{m\boldsymbol{i}}{\hbar}[Q, H_0] \tag{41}$$

Dans le cas où H_0 est donné par (37) et où Q est l'observable de position définie en (38), nous pouvons donner une expression simple de P : en effet, pour toute fonction ψ dans un sous-espace « *convenable* » de $\mathscr{H}$, nous avons :

$$\begin{aligned} [Q, H_0] \cdot \psi = -\frac{\hbar^2}{2m}\left(q\frac{\partial^2 \psi}{\partial q^2} - \frac{\partial^2}{\partial q^2}(q\psi) \right) &= -\frac{\hbar^2}{2m}\left(q\frac{\partial^2 \psi}{\partial q^2} - \frac{\partial}{\partial q}\left(\psi + q\frac{\partial \psi}{\partial q} \right) \right) \\ &= -\frac{\hbar^2}{2m}\left(q\frac{\partial^2 \psi}{\partial q^2} - \left(\frac{\partial \psi}{\partial q} + \frac{\partial \psi}{\partial q} + q\frac{\partial^2 \psi}{\partial q^2} \right) \right) = -\frac{\hbar^2}{m}\frac{\partial \psi}{\partial q} \end{aligned}$$

D'après (41) cela nous donne l'expression de *« l'observable d'impulsion »* P soit :

$$P\psi(q) = -\boldsymbol{i}\hbar \frac{\partial \psi}{\partial q} \tag{42}$$

Dans la suite, nous noterons

$$\langle \psi_t | P | \psi_t \rangle = \langle P\psi_t | \psi_t \rangle = -\boldsymbol{i}\hbar \int \frac{\partial \psi_t}{\partial q} \overline{\psi_t(q)} dq \tag{43}$$

l'impulsion de l'onde-corpuscule à l'instant t (du moins si elle est définie). La proposition suivante fait un lien direct entre la mécanique classique et la mécanique quantique.

25. L'utilisation du crochet de Lie $[Q, H]$ au lieux de $[H, Q]$ vient d'une notation usuelle pour les *« crochets de Poisson »* dans la mécanique hamiltonienne classique (c.f. [Oli17]).

Proposition 5.3. *Si $\psi_t(q) = \psi(t,q)$ est la fonction d'onde associée à une onde-corpuscule de masse m soumise au champ de force dérivant du potentiel $V(q)$ alors (si cela a un sens)*

$$\langle \psi_t | P | \psi_t \rangle = m \frac{d}{dt} \langle \psi_t | Q | \psi_t \rangle$$

5.8. *« Le principe de correspondance de Schrödinger affirme que toute quantité classiquement mesurable d'un système correspond à une observable du système quantique, c'est-à-dire un opérateur densément défini et autoadjoint sur l'espace de Hilbert des états quantiques »* (c.f. § 6.2 infra). Une fois l'interprétation probabiliste de Born posée, la définition de l'observable de position Q est très raisonnable ; en se basant sur la relation classique entre la position et l'impulsion des variables hamiltonienne, la définition de l'observable d'impulsion P en (42) devient elle aussi raisonnable. Cette dernière remarque enlève un peu du mystère qui entoure le principe de correspondance. Notons aussi qu'avec les définitions des opérateurs P et Q ainsi obtenus nous pouvons écrire, pour toute fonction $f \in L^2(\mathbb{R})$ (pour laquelle cela a un sens) :

$$[P,Q]f(q) = PQf(q) - QPf(q) = -\boldsymbol{i}\hbar \left(\frac{\partial}{\partial q}\big(qf(q)\big) - q\frac{\partial f}{\partial q} \right) = -\boldsymbol{i}\hbar f(q)$$

soit encore

$$[P,Q]f(q) = \frac{h}{2\pi \boldsymbol{i}} f(q) \tag{44}$$

Cette relation – *« l'identité de Born »* – est gravée sur la tombe de Born (voir Fig. 1).

5.9. Le rôle de la Q-transformée de Fourier joué dans le formalisme des paquets d'ondes de de Broglie-Schrödinger – tels qu'introduits en (32) – semble avoir disparu dans le formalisme plus général des fonctions d'ondes au sens de la Définition 5.1. Il n'en est rien. En effet, soit ψ_t la fonction partielle spatiale associée à la fonction d'onde $\psi(t,q)$ et pour tout t, supposons que la Q-transformée de Fourier $\hat{\psi}_t(p) := \mathfrak{F}_-[\psi_t](p)$ soit dans $L^2(\mathbb{R}) \cap \mathscr{C}^1(\mathbb{R})$. Alors (en utilisant la dérivation sous le signe intégral) il vient :

$$\begin{aligned} \frac{\partial \psi}{\partial q} = \frac{\partial}{\partial q}\Big(\mathfrak{F}_+\big[\hat{\psi}_t\big](q)\Big) &= \frac{1}{\sqrt{2\pi\hbar}} \int \frac{\partial}{\partial q}\Big(\hat{\psi}_t(p) e^{\boldsymbol{i}pq/\hbar}\Big)\, dp \\ &= \frac{\boldsymbol{i}}{\hbar}\frac{1}{\sqrt{2\pi\hbar}} \int p\hat{\psi}_t(p) e^{\boldsymbol{i}pq/\hbar} dp = \frac{\boldsymbol{i}}{\hbar}\mathfrak{F}_+\big[Q\hat{\psi}_t\big](q) \end{aligned}$$

et par suite, en partant de la définition (43) :

$$\begin{aligned} \langle \psi_t | P | \psi_t \rangle &= -\boldsymbol{i}\hbar \int \frac{\partial \psi_t}{\partial q} \overline{\psi_t(q)} dq \\ &= \int \mathfrak{F}_+\big[Q\hat{\psi}_t\big](q) \overline{\mathfrak{F}_+\big[\hat{\psi}_t\big](q)} dq = \Big\langle \mathfrak{F}_+\big[Q\hat{\psi}_t\big] \Big| \mathfrak{F}_+\big[\hat{\psi}_t\big] \Big\rangle \end{aligned}$$

Finalement par le théorème de Plancherel [Oli16a, Théorème 6.2]), nous obtenons :

$$\langle \psi_t | P | \psi_t \rangle = \langle Q\hat{\psi}_t | \hat{\psi}_t \rangle = \langle \hat{\psi}_t | Q | \hat{\psi}_t \rangle \tag{45}$$

Plus générale (quand cela a un sens), on montre de même que pour tout $k \geq 1$

$$\langle \psi_t | P^k | \psi_t \rangle = \langle \hat{\psi}_t | Q^k | \hat{\psi}_t \rangle$$

La relation (45) est très liée au *« principe de correspondance de Schrödinger »* tel que décrit par de Broglie dans [dB38, p. 48 - 50] : la position et l'impulsion forment un couple (p,q) de variables conjuguées du formalisme hamiltonien qui *« correspond »* au couple (Q,P)

FIGURE 1. *Pierre tombale de Max Born à Göttingen (source Wikipédia). Si Q et P désignent (dans nos notations) l'observable position et l'observable impulsion d'un système quantique, alors « l'identité de Born »*

$$(*)\ :\ PQ - QP = \frac{h}{2\pi \boldsymbol{i}} I$$

(où I est l'opérateur identité) contient à elle seule deux points qui distinguent la mécanique classique de la mécanique quantique. Premièrement $(*)$ ne peut être valable en dimension finie car dans ce cas la trace de PQ est égale à la trace de QP. Deuxièmement, si l'espace est de dimension infini et que les deux opérateurs P et Q vérifient $(*)$, alors il est nécessaire que P où Q soit non borné : cette remarque est implicite dans un articled de Weyl de 1928 et se trouve formellement établie par Wintner en 1947 (c.f. [RS81, Example 2, p. 274]*). Notons enfin que l'identité de Born est à l'origine d'un théorème structurel (lié au principe de correspondance de Schrödinger) et établit par Stone et von Neumann en 1932 (c.f.* [Ros04] *pour une histoire sélective autour de ces questions).*

des observables position-impulsion dont la conjugaison *quantique* est assurée par (45). Cette dernière remarque, nous permet de formuler une version (en fait assez générale) du *« l'inégalité de Heisenberg »*. En supposant que la position $\langle \psi_t | Q | \psi_t \rangle$ et l'impulsion $\langle \psi_t | P | \psi_t \rangle$ soient définies à l'instant t, nous pouvons définir (et nous y reviendrons de manière plus systématique au paragraphe suivant) *« l'écart type en position »*

$$\sigma_t(Q) := \sqrt{\langle \psi_t | Q^2 | \psi_t \rangle - \langle \psi_t | Q | \psi_t \rangle^2}$$

ainsi que *« l'écart type en impulsion »*

$$\sigma_t(P) := \sqrt{\langle\psi_t|P^2|\psi_t\rangle - \langle\psi_t|P|\psi_t\rangle^2} = \sqrt{\langle\hat{\psi}_t|Q^2|\hat{\psi}_t\rangle - \langle\hat{\psi}_t|Q|\hat{\psi}_t\rangle^2}$$

Il est immédiat de vérifier que $\sigma_t(Q)$ et $\sigma_t(P)$ coïncident respectivement avec les écarts types des probabilités $\mu_t(dq) = |\psi_t(q)|^2dq$ et $\eta_t(dp) := |\hat{\psi}_t(p)|^2dp$: en d'autres termes, $\sigma_t(Q) = \sigma(\mu_t)$ et $\sigma_t(P) = \sigma(\eta_t)$. Si $\xi_t(dp) := |\varphi_t(p)|^2dp$ est la probabilité, où

$$\varphi(p) = \frac{1}{\sqrt{2\pi}}\int \psi_t(q)e^{-ipq}dq$$

alors *« l'inégalité de Heisenberg-Gabor »* (c.f. [Oli16a, Théorème 9.1]) affirme que

$$\sigma(\mu_t)\sigma(\xi_t) \geq \frac{1}{2} \tag{46}$$

La relation entre $\sigma(\eta_t)$ et $\sigma(\xi_t)$ s'obtient facilement par changement de variable : en effet,

$$\hat{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}}\int \psi_t(q)e^{-i(p/\hbar)q}dq = \frac{1}{\sqrt{\hbar}}\varphi(p/\hbar)$$

de sorte que

$$\begin{aligned}
\sigma^2(\eta_t) &= \int p^2|\hat{\psi}_t(p)|^2dp - \left(\int p|\hat{\psi}_t(p)|^2dp\right)^2 \\
&= \hbar\int (p/\hbar)^2|\varphi_t(p/\hbar)|^2dp - \left(\int (p/\hbar)|\varphi_t(p/\hbar)|^2dp\right)^2 \\
&= \hbar^2\int p^2|\varphi_t(p)|^2dp - \left(\hbar\int p|\varphi_t(p)|^2dp\right)^2 \\
&= \hbar^2\left(\int (p)^2|\varphi_t(p)|^2dp - \left(\int p|\varphi_t(p)|^2dp\right)^2\right) = \hbar^2\sigma^2(\xi_t)
\end{aligned}$$

Nous venons de vérifier que $\sigma(\eta_t) = \hbar\sigma(\xi_t)$, ce qui permet d'écrire l'inégalité de Heisenberg-Gabor en (46) sous sa forme originelle $\sigma(\mu_t)\sigma(\eta_t) = \hbar/2$ et connue en mécanique quantique comme *« l'inégalité de Heisenberg »*.

Théorème 5.4. *Soit $\psi_t = \psi(t,\cdot)$ la fonction partielle spatiale d'une fonction d'onde associée à un corpuscule de masse m pris dans un champ de force conservatif (c.f. Définition 5.1) ; alors, à tout instant t où la position $\langle\psi_t|Q|\psi_t\rangle$ et l'impulsion $\langle\psi_t|P|\psi_t\rangle$ sont définies, nous avons*

$$\sigma_t(Q)\sigma_t(P) \geq \frac{\hbar}{2} \tag{47}$$

6. Forme abstraite de l'équation de Schrödinger

6.1. Nous avons déjà vu (c.f. [Oli16b]) que le travail de réflexion sur la formulation abstraite de la mécanique quantique à été menée au tournant des années 30 par Stone et von Neumann. Pour donner un aperçu élémentaire de cette formulation, considérons que $\mathscr{H}$ est un espace de Hilbert (séparable) et que le système étudié est décrit par un *« Q-hamiltonien »*[26] H, c'est-à-dire un opérateur défini sur un sous-espace $\mathscr{D}$ dense dans $\mathscr{H}$ et tel que $(H,\mathscr{D})$ soit autoadjoint. Par analogie avec l'équation de Schrödinger associée

26. Nous supposerons ici que H ne dépend pas du temps.

en (35) au corpuscule de masse m pris dans un champ de force, *« l'équation de Schrödinger abstraite »* associée au Q-hamiltonien $(H, \mathscr{D})$ s'écrit

$$i\hbar \frac{d\Psi}{dt} = H\Psi(t) \tag{48}$$

L'existence et l'unicité des solutions de l'équation (48) sont assurées par le *« théorème de Stone »* [27] (c.f. [Oli16b, Théorème 7.19]) : en effect $(H, \mathscr{D})$ étant autoadjoint $(iH/\hbar, \mathscr{D})$ est le *« générateur infinitésimal »* d'un *« semi-groupe fortement continue de contraction »* (s.g.f.c.c.) habituellement noté $\{\exp(itH/\hbar)\ ;\ t \geq 0\}$: par suite, pour tout Q-état initial $\psi_0 \in \mathscr{D}$, le chemin $\Psi(t) = \exp(itH/\hbar)\psi_0$ est l'unique Q-trajectoire de $(H, \mathscr{D})$ t.q. $\Psi(0) = \psi$ (i.e. l'unique chemin de la sphère unité de $\mathscr{D}$ t.q. $i\hbar\partial\Psi/\partial t = H\Psi(t)$ avec $\Psi(0) = \psi_0$). Nous énonçons ce résultat sous la forme suivante.

Théorème 6.1. *Soit $(H, \mathscr{D})$ une Q-hamiltonien ; si ψ_0 est un Q-état dans $\mathscr{D}$ (i.e. $\psi_0 \in \mathscr{D}$ et $\|\psi_0\| = 1$), alors il existe un unique chemin $\Psi(t)$ dans $\mathscr{D}$ qui est solution de (48) et t.q. $\Psi(0) = \psi_0$.*

Si nous considérons que $t \mapsto \|\Psi(t)\|^2$ et la solution de (48) t.q. $\Psi(0) = \psi_0 \in \mathscr{D}$, alors (dérivation sous le signe intégral) :

$$\begin{aligned}
\frac{d}{dt}\|\Psi(t)\|^2 = \frac{d}{dt}\left\langle \Psi(t) \middle| \Psi(t) \right\rangle &= \left\langle \frac{d\Psi}{dt} \middle| \Psi(t) \right\rangle + \left\langle \Psi(t) \middle| \frac{d\Psi}{dt} \right\rangle \\
&= \left\langle -\frac{i}{\hbar} H\Psi(t) \middle| \Psi(t) \right\rangle + \left\langle \Psi(t) \middle| -\frac{i}{\hbar} H\Psi(t) \right\rangle = 0
\end{aligned}$$

où nous avons utilisé le fait que $(H, \mathscr{D})$ est autoadjoint : il en découle que la norme $\|\Psi(t)\|$ est constante,, i.e. $\|\Psi(t)\| = \|\psi_0\|$, pour tout $t \geq 0$. Ces dernières remarques nous permettent de définir de manière consistante les Q-trajectoires comme généralisation abstraite des fonctions d'onde telles que définies dans la Définition 5.1.

Définition 6.2. *Soit $(H, \mathscr{D})$ un Q-hamiltonien ; alors une « Q-trajectoire » est un chemin différentiable $t \mapsto \Psi(t)$ formé de Q-états dans $\mathscr{D}$ (i.e.$\Psi(t)$ appartient à la sphère unité de $\mathscr{D}$ pour tout $t \geq 0$) et qui est solution de l'équation de Schrödinger abstraite donnée en (48).*

6.2. L'opérateur $(A, \mathscr{K})$, supposé autoadjoint et densément défini sur $\mathscr{H}$, est appelé une *« observable compatible »* avec le Q-hamiltonien $(H, \mathscr{D})$, lorsque $\mathscr{K}$ contient $\mathscr{D}$. Une telle observable représente une quantité mesurable à partir d'expériences faites sur le système décrit par $(H, \mathscr{D})$: les résultats « *possibles* » des mesures sont alors *« identifiées par principe »* aux valeurs spectrales de $(A, \mathscr{K})$ (c'est-à-dire – d'après le critère de Weyl – aux valeurs propres ou aux valeurs propres approchées : c.f. [Oli16b, § 3.3]). La Q-moyenne de A évaluée le long d'une Q-trajectoire $\psi_t := \Psi(t)$ dans $\mathscr{D}$ est

$$\langle \psi_t | A | \psi_t \rangle = \int \overline{\psi_t(q)} A\psi_t(q) dq$$

D'autre part, sous réserve que A^2 soit une observable compatible avec $(H, \mathscr{D})$, alors nous pouvons définir la Q-dispersion de A évaluée en ψ_t en posant

$$\sigma_t(A) = \sqrt{\langle \psi_t | A^2 | \psi_t \rangle - \langle \psi_t | A | \psi_t \rangle^2}$$

27. C'est une forme partielle (pour les semi-groupes) qu'il est possible de voir (de manière anachronique) comme un corollaire du *« théorème de Hille-Yosida »* concernant les semi-groupes fortement continue de contraction (s.g.f.c.c.) dans le cas des espaces de Hilbert.

Le théorème suivant est une version du *« principe d'incertitude de Heisenberg »*.

Théorème 6.3. *Soit A et B deux observables compatibles avec $(H, \mathscr{D})$; si $[A, B]$ est aussi compatible avec $(H, \mathscr{D})$, alors pour toute Q-trajectoire $\psi_t = \Psi(t)$ dans $\mathscr{D}$,*

$$\sigma_t(A)\sigma_t(B) \geq \frac{1}{2}\left|\left\langle\psi_t|[A, B]|\psi_t\right\rangle\right|$$

Preuve. Pour $f, g \in \mathscr{H}$ l'inégalité triangulaire entraîne que

$$\|f\|^2 + \|g\|^2 + 2\|f\|\|g\| \geq \|f + g\|^2 = \|f\|^2 + \|g\|^2 + \langle f|g\rangle + \langle g|f\rangle$$

Après simplification (et en changeant éventuellement le signe de f) nous retrouvons (une forme de) l'inégalité de Cauchy-Schwarz, soit (c.f. [Oli16a, § 3.2, eq. (4)]) :

$$2\|f\|\|g\| \geq |\langle f|g\rangle + \langle g|f\rangle| \tag{49}$$

En posant $\underline{A}(t) = A - \langle\psi_t|A|\psi_t\rangle I$ et $\underline{B}(t) := B - \langle\psi_t|B|\psi_t\rangle I$ il vient $\sigma_t(A) = \|\underline{A}(t)\psi_t\|$ et $\sigma_t(B) = \|\underline{B}(t)\psi_t\|$) : d'après (49) avec $f = \underline{A}(t)\psi_t$ et $g = \underline{B}(t)\psi_t$, il vient alors

$$\begin{aligned} 2\sigma_t(A)\sigma_t(B) &\geq \left|\left\langle\underline{A}(t)\psi_t|\underline{B}(t)\psi_t\right\rangle + \left\langle\underline{B}(t)\psi_t|\underline{A}(t)\psi_t\right\rangle\right| \\ &\geq \left|\left\langle\psi_t|\underline{A}(t)\underline{B}(t)|\psi_t\right\rangle + \left\langle\psi_t|\underline{B}(t)\underline{A}(t)|\psi_t\right\rangle\right| = \left|\left\langle\psi_t|\underline{A}(t)\underline{B}(t) - \underline{B}(t)\underline{A}(t)|\psi_t\right\rangle\right| \end{aligned}$$

i.e. $\sigma_t(A)\sigma_t(B) \geq 1/2|\langle\psi_t|[\underline{A}(t), \underline{B}(t)]|\psi_t\rangle|$. La conclusion vient du fait que

$$\begin{aligned} [\underline{A}(t), \underline{B}(t)] &= [A, B] - \langle\psi_t|B|\psi_t\rangle\,[A, I] - \langle\psi_t|A|\psi_t\rangle\,[B, I] \\ &\qquad + \langle\psi_t|A|\psi_t\rangle\,\langle\psi_t|B|\psi_t\rangle\,[I, I] = [A, B] \end{aligned}$$

□

Nous avons vu au § 5.6 que dans le cas où $\Psi(t) := \psi_t$ est la Q-trajectoire associée à la fonction d'onde $\psi_t(q) = \psi(t, q)$ du corpuscule de masse m pris dans un champ de force, les observables position et impulsion sont respectivement définies pour tout f dans le domaine du hamiltonien en posant $Qf(q) = qf(q)$ et $Pf(q) = -i\hbar\partial f/\partial q$. Or d'après *« l'identité de Born »* (c.f. § 5.8 supra) nous avons $[P, Q]f = -i\hbar f$. En se plaçant le long d'une Q-trajectoire quelconque $\psi_t = \Psi(t)$, nous pouvons combiner l'identité de Born avec le *« principe d'incertitude »* tel qu'énoncé dans le Théorème 6.3 afin de retrouver *« l'inégalité de Heisenberg »* (c.f. Théorème 5.4) : en effet, pour tout $t \geq 0$,

$$\sigma_t(Q)\sigma_t(P) \geq \frac{1}{2}|\langle\psi_t|[P, Q]|\psi_t\rangle| = \frac{1}{2}|\langle\psi_t| - i\hbar I|\psi_t\rangle| = \frac{\hbar}{2}\|\psi_t\|^2 = \frac{\hbar}{2}$$

6.3. Nous allons démontrer *« le théorème d'Ehrenfest »* [28] portant sur l'évolution des valeurs moyennes des observables (indépendantes du temps) le long d'une Q-trajectoire.

Théorème 6.4 (Ehrenfest)**.** *Soient $(H, \mathscr{D})$ un hamiltonien densément défini sur $\mathscr{H}$ et $\Psi(t)$ une Q-trajectoire dans le domaine de H. Si A est une observable compatible avec $(H, \mathscr{D})$, alors*

$$\frac{d}{dt}\left\langle\Psi(t)|A|\Psi(t)\right\rangle = -\frac{i}{\hbar}\left\langle\Psi(t)|[A, H]|\Psi(t)\right\rangle$$

---

28. Ehrenfest (1880-1933), élève de Boltzmann et ami d'Einstein et de Bohr, était un brillant physicien prématurément et tragiquement disparu.

Preuve. Nous appliquons le théorème de dérivation sous le signe intégral, de sorte que :

$$\begin{aligned}\frac{d}{dt}\left\langle\Psi(t)\middle|A\middle|\Psi(t)\right\rangle &= \left\langle\frac{d\Psi}{dt}\middle|A\middle|\Psi(t)\right\rangle + \left\langle\Psi(t)\middle|A\middle|\frac{d\Psi}{dt}\right\rangle \\ &= \left\langle-\frac{\boldsymbol{i}}{\hbar}H\Psi(t)\middle|A\middle|\Psi(t)\right\rangle + \left\langle\Psi(t)\middle|A\middle|-\frac{\boldsymbol{i}}{\hbar}H\Psi(t)\right\rangle \\ &= -\frac{\boldsymbol{i}}{\hbar}\left\langle\Psi(t)\middle|HA\middle|\Psi(t)\right\rangle + \frac{\boldsymbol{i}}{\hbar}\left\langle\Psi(t)\middle|AH\middle|\Psi(t)\right\rangle\end{aligned}$$

(par convention nous utilisons un produit scalaire semi-linéaire à gauche) soit encore

$$\frac{d}{dt}\left\langle\Psi(t)\middle|A\middle|\Psi(t)\right\rangle = -\frac{\boldsymbol{i}}{\hbar}\left\langle\Psi(t)\middle|AH - HA\middle|\Psi(t)\right\rangle$$

□

Lorsque c'est possible, le théorème d'Ehrenfest donne une interprétation classique cohérente à partir de la description quantique du système étudié ; nous pouvons vérifier cela dans le cas particulier du corpuscule de masse m dans dans le champ de force dérivant du potentiel $V(q)$. Pour un tel système, nous savons que le Q-hamiltonnien H est défini sur l'espace de Sobolev $\mathscr{H}^2(\mathbb{R})$ et que pour tout $\psi \in \mathscr{H}^2(\mathbb{R})$

$$H\psi = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial q^2} + V(q)\psi(q)$$

En extrapolant les expressions des observables de position et d'impulsion Q et P obtenues en (38) et (42) au § 5.7 pour le corpuscule libre, il vient pour tout Q-état $\psi \in \mathscr{H}^2(\mathbb{R})$:

$$Q\psi(q) = q\psi(q) \quad \text{et} \quad P\psi(q) = \mathfrak{F}_+[Q\mathfrak{F}_-[\psi]](q) = -\boldsymbol{i}\hbar\frac{\partial\psi}{\partial q}$$

Nous affirmons alors que le long d'une Q-trajectoire $\psi_t = \Psi(t)$

$$\frac{d}{dt}\langle\psi_t|Q|\psi_t\rangle = \frac{1}{m}\langle\psi_t|P|\psi_t\rangle \tag{50}$$

(ce qui est cohérent avec l'interprétation classique de la vitesse et de l'impulsion). Nous allons vérifier (50), grâce au théorème d'Ehrenfest. Le calcul du crochet de Lie $[Q, H]$ évalué en un Q-état ψ s'effectue en écrivant d'une part,

$$QH\psi = -\frac{\hbar^2}{2m}q\frac{\partial^2\psi}{\partial q^2} + qV(q)\psi(q)$$

alors que d'autre part,

$$\begin{aligned}HQ\psi &= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2}(q\psi(q)) + qV(q)\psi(q) \\ &= -\frac{\hbar^2}{2m}\frac{\partial}{\partial q}\left(\psi(q) + q\frac{\partial\psi}{\partial q}\right) + qV(q)\psi(q) = -\frac{\hbar^2}{2m}\left(\frac{\partial\psi}{\partial q} + \frac{\partial\psi}{\partial q} + q\frac{\partial^2\psi}{\partial q^2}\right) + qV(q)\psi(q)\end{aligned}$$

de sorte que

$$HQ = -\frac{\hbar^2}{m}\frac{\partial\psi}{\partial q} + QH\psi$$

Compte-tenu de l'expression de l'observable d'impulsion P, nous obtenons $[Q,H] = -i\hbar/mP$ et finalement, par le Théorème d'Ehrenfest nous retrouvons (50) puisque

$$\frac{d}{dt}\langle\psi_t|Q|\psi_t\rangle = \frac{i}{h}\langle\psi_t|[Q,h]|\psi_t\rangle = \frac{i}{h}\left\langle\psi_t\middle|-\frac{i\hbar}{m}P\middle|\psi_t\right\rangle = \frac{1}{m}\langle\psi_t|P|\psi_t\rangle$$

Nous traiterons le cas général du théorème d'Ehrenfest (i.e. avec des observables dépendant du temps) dans [Oli17] à partir de la correspondance entre crochet de Poisson *d'observables classiques* et crochet de Lie *d'observables quantiques*, telle qu'elle est mise en oeuvre dans le *« principe de correspondance de Schrödinger »*.

7. Equation de Schrödinger abstraite : cas du spectre discret

7.1. Soit $\mathscr{H}$ un sous-espace fermé de l'espace de Hilbert (complexe) $L^2(\mathbb{R})$, muni de la structure d'espace de Hilbert induite (par exemple $\mathscr{H}$ pourra coïncider avec l'espace de Hilbert engendré par une famille orthonormée). Dans la suite $(H,\mathscr{D})$ désigne un Q-hamiltonien, i.e. un opérateur autoadjoint densément défini sur $\mathscr{H}$: en d'autres termes, $\mathscr{D}$ est un sous-espace dense de $\mathscr{H}$ et H est une application linéaire fermée de $\mathscr{D}$ dans $\mathscr{H}$ qui est symétrique en ce sens que $\langle Hf|g\rangle = \langle f|Hg\rangle$, quel que soient $f,g\in\mathscr{D}$ (c.f. [Oli16a, § 10]). Le *« théorème de Stone »* (c.f. [Oli16b, Théorème 7.20]) donne une solution abstraite de l'équation de Schrödinger dans le cas général i.e. sans faire d'hypothèse sur la nature spectrale du Q-hamiltonien $(H,\mathscr{D})$ (c.f. Théorème 6.1 supra). Nous allons voir que dans le cas où le Q-Hamiltonien est à spectre discret il est possible de retrouver ce résultat sans l'aide du théorème de Stone[29].

7.2. Pour aborder la question de l'intégration de l'équation de Schrödinger dans le cas d'un système à spectre discret (comme l'oscillateur harmonique traité au § 8), nous supposons pour simplifier qu'il existe une suite infinie strictement croissante de valeurs propres $E_0 < E_1 < \cdots$ associées à $(H,\mathscr{D})$, chaque valeurs propres E_n (*« niveau d'énergie »* du système) étant supposée non dégénérée ; il existe alors une suite orthonormée $\phi_0,\phi_1,\ldots$ de Q-états dans $\mathscr{D}$ tels que

(51) $$H\phi_n = E_n\phi_n$$

Nous nous plaçons dans le cas purement discret, ce qui revient à supposer que $\mathscr{H}$ est le sous-espace fermé de $L^2(\mathbb{R})$ engendré par les ϕ_n : en d'autres termes $\mathscr{H}$ est l'espace de Hilbert pour lequel les Q-états $\phi_0,\phi_1,\ldots$ forment une base hilbertienne. Par définition, *« l'état stationnaire »* associé au niveau d'énergie E_n est le Q-état

(52) $$\Psi_n(t) := e^{-iE_nt/\hbar}\phi_n$$

L'équation aux valeurs propres en (51) assure que $\Psi_n(t)$ est une Q-trajectoire de $(H,\mathscr{D})$, i.e. un chemin différentiable de la sphère unité de $\mathscr{D}$ t.q. $i\hbar d\Psi_n/dt = H\Psi_n(t)$ $(= E_n\Psi_n(t))$. Le terme *« stationnaire »* est simple à comprendre : sous réserve que $q|\phi_n(q)|^2$ soit intégrable, *« la position moyenne »* le long de la Q-trajectoire $\Psi_n(t)$ i.e.

$$\left\langle\Psi_n(t)\middle|Q\middle|\Psi_n(t)\right\rangle = \int q|\psi_n(t,q)|^2dq = \int q\left|e^{-iE_nt/\hbar}\phi_n(q)\right|^2dq = \int q\,|\phi_n(q)|^2\,dq$$

est une quantité indépendante du temps.

29. Beaucoup de présentation de la mécanique quantique commence d'ailleurs par cette présentation élémentaire et très intuitive.

7.3. Par linéarité, il est immédiat que les combinaisons linéaires finies d'états stationnaires sont des solutions (non nécessairement stationnaires) de l'équation de Schrödinger. Afin de trouver des solutions plus générales, il est naturel de considérer des combinaisons linéaires infinies d'états stationnaires. Pour cette généralisation, il est cependant nécessaire de prendre quelques précautions. Commençons d'abord par noter qu'en posant, pour tout $t \geq 0$ et tout $\psi \in \mathscr{H}$,

$$U_t\psi = \sum_n \langle\psi|\phi_n\rangle\, \Psi_n(t) = \sum_{n=0}^{+\infty} \langle\psi|\phi_n\rangle\, e^{-iE_n t/\hbar}\phi_n \tag{53}$$

nous définissons un opérateur unitaire de $\mathcal{L}(\mathscr{H})$ et que l'ensemble formé des U_t pour $t \geq 0$ est *« un semi-groupe d'opérateurs unitaires »* (i.e. U_0 est l'identité de $\mathscr{H}$ et $U_{t+t'} = U_t U_{t'}$, pour tout $t, t' \geq 0$). Etant donné ψ_0 un élément de la sphère unité de $\mathscr{H}$, nous allons nous intéresser aux conditions assurant que la chemin de Q-états

$$\Psi(t) := U_t\psi_0 = \sum_n \langle\psi_0|\phi_n\rangle\, e^{-iE_n t/\hbar}\phi_n \tag{54}$$

correspond bien à une Q-trajectoire du système quantique de Q-hamiltonien $(H, \mathscr{D})$.

Proposition 7.1. *Pour tout $\psi_0 \in \mathscr{H}$ et t.q. $\|\psi_0\| = 1$, l'application $t \mapsto \Psi(t) = U_t\psi_0$ définie en (54) est un chemin continu de la sphère unité de $\mathscr{H}$ (i.e. un chemin continu de Q-états).*

Preuve. Notons $c_n := \langle\psi_0|\phi_n\rangle$; si $t_k \to t$, alors

$$\Psi(t_k) - \Psi(t) = \sum_n c_n \left(e^{-iE_n t_k/\hbar} - e^{-iE_n t/\hbar}\right)\phi_n \tag{55}$$

Ainsi, dans la limite où $k \to +\infty$, la convergence ponctuelle $c_n e^{-iE_n t_k/\hbar} \to c_n e^{-iE_n t/\hbar}$ (pour $n \in \mathbb{N}$) combinée à l'inégalité $|e^{-iE_n(t_k-t)/\hbar} - 1| \leq 2$ entraînent que

$$\left|c_n\left(e^{-iE_n t_k/\hbar} - e^{-iE_n t/\hbar}\right)\right| = \left|c_n\left(e^{-iE_n(t_k-t)/\hbar} - 1\right)\right| \leq 2|c_n|$$

Par suite, en appliquant le *« théorème de la convergence dominée »*, il vient

$$\lim_{k\to+\infty} \sum_n \left|c_n\left(e^{-iE_n t_k/\hbar} - e^{-iE_n t/\hbar}\right)\right|^2 = 0$$

Compte tenu de (55) cette dernière convergence permet de conclure que (dans $\mathscr{H}$)

$$\lim_{k\to+\infty} \Psi(t_k) = \Psi(t)$$

d'où la continuité du chemin $\Psi(t)$ pour $t \geq 0$.

□

Cependant, pour identifier les chemins de Q-états $\Psi(t) = U_t\psi_0$ (avec $\psi_0 = \Psi(0) \in \mathscr{H}$ et $\|\Psi_0\| = 1$) en tant que Q-trajectoires, il reste d'une part à déterminer sous quelle(s) condition(s) $\Psi(t) \in \mathscr{D}$ pour tout $t \geq 0$ et d'autre part, de vérifier que $t \mapsto \Psi(t)$ est bien un chemin différentiable solution de l'équation de Schrödinger. Comme déjà mentionné, le *« théorème de Stone »* donne directement la réponse à ces questions. Dans le cas du spectre discret, il est cependant possible de vérifier directement que la seule condition $\psi_0 = \Psi(0) \in \mathscr{D}$ implique automatiquement que $\Psi(t)$ est une Q-trajectoire.

7.4. Pour tout Q-état $\psi_0 \in \mathscr{H}$ (i.e. $\|\psi_0\| = 1$) nous avons vu que $t \mapsto \Psi(t) = U_t\psi_0$ $(t \geq 0)$ définie en (53) est un chemin continue sur la sphère unité de $\mathscr{H}$: pour déterminer sous quelle(s) condition(s) c'est un chemin de la sphère unité de $\mathscr{D}$, nous utilisons un résultat général sur les opérateurs à spectre discret (c.f. [Oli16b, Proposition 2.3]) qui, lorsqu'il est combiné à l'équation aux valeurs propres (51), nous donne la proposition suivante.

Proposition 7.2. *Si $E_0 < E_1 < \cdots$ est la suite des valeurs propres (supposées non dégénérées) du Q-hamiltonien $(H, \mathscr{D})$, alors*

(56) $$\mathscr{D} := \left\{\psi \in \mathscr{H}\ ;\ \sum_n |E_n \langle\psi|\phi_n\rangle|^2 < +\infty\right\}$$

et pour tout $\psi \in \mathscr{D}$

(57) $$H\psi = \sum_n E_n \langle\psi|\phi_n\rangle\, \phi_n$$

Les coefficients de Fourier $c_n = \langle\psi_0|\phi_n\rangle$ associés au Q-état initial $\psi_0 = \Psi(0)$ forment une suite de nombres complexes de carrés sommables t.q. (normalisation des Q-états) :

(58) $$\sum_n |c_n|^2 = 1$$

Il découle directement de (56) que le support du chemin $\Psi(t) = U_t\psi_0$ défini en (53) appartient à la sphère unité de $\mathscr{D}$ ssi $\psi_0 \in \mathscr{D}$, c'est-à-dire lorsque la normalisation (58) est doublée par la condition

(59) $$\sum_n |E_n c_n|^2 < +\infty$$

La proposition suivante précise la Proposition 7.1.

Proposition 7.3. *Pour tout $\psi_0 \in \mathscr{D}$ et t.q. $\|\psi_0\| = 1$, l'application $t \mapsto \Psi(t) = U_t\psi_0$ définie en (54) est un chemin continu de la sphère unité de $\mathscr{D}$.*

7.5. Nous allons maintenant vérifier directement que $\Psi(t) = U_t\psi_0$ est bien une Q-trajectoire de $(H, \mathscr{D})$, lorsque le Q-état initial $\psi_0 = \Psi(0)$ est dans $\mathscr{D}$. Pour ce faire, nous adaptons l'argument donnant la continuité de ce chemin lorsque $\psi_0 \in \mathscr{H}$ (c.f. Proposition 7.1 supra) à partir du lemme suivant (cas spécial de *« dérivation sous le signe sommes »*).

Lemme 7.4. *Soit $(\phi_0, \phi_2, \dots)$ une base hilbertienne de l'espace de Hilbert (séparable) $\mathscr{H}$ et pour tout $n \in \mathbb{N}$ et tout $t \geq 0$, soit $f_n(t) \in \mathbb{C}$ tel que $t \mapsto f_n(t)$ soit dérivable de dérivée notée $f_n'(t)$; si (i) : il existe $t_0 \geq 0$ t.q. $\sum_n |f_n(t_0)|^2 < +\infty$ et si (ii) : il existe $g_n \geq 0$ tels que $|f_n'(t)| \leq g_n$ avec $\sum_n g_n^2 < +\infty$, alors $\Psi(t) := \sum_n f_n(t)\phi_n$ $(t \geq 0)$ est un chemin dérivable de $\mathscr{H}$ tel que*

$$\frac{d\Psi}{dt} = \sum_n f_n'(t)\phi_n$$

Preuve. Les fonctions $t \mapsto f_n(t)$ étant dérivables pour $t \geq 0$, d'après le théorème des accroissements finis et l'hypothèse de domination (ii), nous avons pour tout $t, t' \geq 0$:

(60) $$|f_n(t) - f_n(t')| \leq |t - t'| g_n$$

de sorte que pour tout $t \geq 0$, en introduisant t_0 il vient :

(61) $$|f_n(t)|^2 \leq \frac{1}{2}|f_n(t) - f_n(t_0)|^2 + \frac{1}{2}|f_n(t_0)|^2 \leq \frac{1}{2}|t - t_0|^2 g_n^2 + \frac{1}{2}|f_n(t_0)|^2$$

Or par hypothèse $\sum_n |f_n(t_0)|^2 < +\infty$ et $\sum_n g_n^2 < +\infty$, de sorte que par (61) nous avons $\sum_n |f_n(t)|^2 < +\infty$: cela assure que $\Psi(t) := \sum_n f_n(t)\phi_n$ est un élément bien défini de $\mathscr{H}$ et ce, quelque soit $t \geq 0$. Maintenant supposons que $t_k \to t$ (avec $t_k \neq 0$) : alors

$$\frac{\Psi(t_k) - \Psi(t)}{t_k - t} = \sum_n G_{n,k}(t)\phi_n \quad \text{où} \quad G_{n,k}(t) := \frac{f_n(t_k) - f_n(t)}{t_k - t} \tag{62}$$

Les fonctions $t \mapsto f_n(t)$ étant dérivables (pour tout n), il vient d'une part que $G_{n,k}(t) \to f_n'(t)$, quand $k \to +\infty$. Mais d'autre part, partant de l'expression de $G_{n,k}(t)$ en (62), l'inégalité (60) entraîne $|G_{n,k}(t)| \leq g_n$. Avec l'hypothèse (ii), le théorème de la convergence dominée assure que

$$\lim_{k\to+\infty} \sum_n \left|G_{n,k}(t) - f_n'(t)\right|^2 = 0 \tag{63}$$

Comme les ϕ_n forment une base hilbertienne de $\mathscr{H}$, en combinant (62) et (63) il vient :

$$\lim_{k\to+\infty} \left\| \frac{\Psi(t_k) - \Psi(t)}{t_k - t} - \sum_n f_n'(t)\phi_n \right\|^2 = \lim_{k\to+\infty} \sum_n \left|G_{n,k}(t) - f_n'(t)\right|^2 = 0$$

d'où nous tirons l'identité $d\Psi/dt = \sum_n f_n'(t)\phi_n$ (qui a lieu dans $\mathscr{H}$).

□

7.6. Considérons $\Psi(t) = U_t\psi_0$, où le Q-état initial $\psi_0 = \sum_n c_n\phi_n$ est dans $\mathscr{D}$ – c'est-à-dire que les coefficients de Fourier $c_n = \langle\psi_0|\phi_n\rangle$ vérifient les conditions (58) et (59). D'après la Propositions 7.3, nous savons que $\Psi(t)$ est un chemin continu sur la sphère unité de $\mathscr{D}$; pour la différentiabilité, nous appliquons directement le Lemme 7.4 en écrivant

$$\Psi(t) := U_t\psi_0 = \sum_n f_n(t)\phi_n \quad \text{où} \quad f_n(t) = \langle\Psi(t)|\phi_n\rangle = c_n e^{-iE_nt/\hbar} \tag{64}$$

et où la base hilbertienne de $\mathscr{H}$ est formée des Q-états propres $\phi_0, \phi_1, \ldots$ de $(H, \mathscr{D})$ t.q. $H\phi_n = E_n\phi_n$. Les conditions du lemme sont ici trivialement vérifiées et tout spécialement la condition de domination des dérivées, du fait que

$$\|f_n'(t)\| \leq g(n) := \left\| -\frac{i}{\hbar} c_n E_n e^{-iE_nt/\hbar}\phi_n \right\| = \frac{1}{\hbar}|E_n c_n|$$

Insistons : le fait que $\sum_n g_n^2 < +\infty$ découle de la condition (59) assurant que l'état initial $\Psi(0) = \psi_0 = \sum_n c_n\phi_n$ est dans $\mathscr{D}$. Finalement, par une application de la dérivation sous le signe somme, il vient

$$i\hbar\frac{d\Psi}{dt} = i\hbar\sum_n f_n'(t)\phi_n = i\hbar\sum_n \left(-\frac{i}{\hbar}E_n f_n(t)\right)\phi_n = \sum_n E_n f_n(t)\phi_n$$

soit encore, d'après l'expression de $f_n(t)$ en (64),

$$i\hbar\frac{d\Psi}{dt} = \sum_n E_n\langle\Psi(t)|\phi_n\rangle\phi_n$$

La conclusion que $i\hbar d\Psi/dt = H\Psi(t)$ vient de la définition de H sur $\mathscr{D}$ donnée dans l'équation (57) de la Proposition 7.2 : cela fait de $\Psi(t) = U_t\psi_0$ une Q-trajectoire de $(H, \mathscr{D})$.

Proposition 7.5. *Pour tout $\psi_0 \in \mathscr{D}$ et t.q. $\|\psi_0\| = 1$, l'application $t \mapsto \Psi(t) = U_t\psi_0$ définie en (54) est une Q-trajectoire de $(H, \mathscr{D})$ t.q. $\Psi(0) = \psi_0$.*

8. L'oscillateur harmonique

8.1. Un des premiers exemple complètement résolu dans le cadre de la mécanique ondulatoire est exposé dans la deuxième de communication de Schrödinger dans la suite des articles fondateurs qu'il publie en 1926 (c.f. [Sch33, p. 49-55]) : Schrödinger l'appelle *« l'oscillateur de Planck »*. Ce modèle de la mécanique ondulatoire/quantique est maintenant connu sous le nom *« l'oscillateur harmonique quantique »*.

8.2. Considérons un mobile de masse m se déplaçant sur la droite réelle $\mathbb{R}$ et de position $\boldsymbol{q}(t)$: le mobile en question est un *« oscillateur harmonique (classique) »* de centre 0 (origine des coordonnées de position) lorsque l'accélération $\ddot{\boldsymbol{q}}(t)$ est proportionnelle à la *« position »* $\boldsymbol{q}(t)$ et dirigée vers 0 : par suite, (principe fondamental de la dynamique de Newton), $\boldsymbol{q}(t)$ vérifie l'équation de Newton $m\ddot{\boldsymbol{q}}(t) = -k\boldsymbol{q}(t)$ où k est la constante (positive) caractéristique de l'oscillateur (homogène à une masse divisée par le carré d'une durée). La force agissant sur notre mobile dérive du potentiel[30] $V(q) = 1/2kq^2$, de sorte que l'équation de Newton s'écrit $m\ddot{\boldsymbol{q}}(t) = -\nabla V(\boldsymbol{q}(t))$ (où $\nabla V(q) = dV/dq = kq$ désigne le gradient de V en q). Par définition, *« l'énergie mécanique »* $E(t)$ de l'oscillateur harmonique est la somme de *« l'énergie cinétique »* $E_c(t) = 1/2m\dot{\boldsymbol{q}}^2(t)$ et de *« l'énergie potentielle »* $E_p(t) = V(\boldsymbol{q}(t)) = 1/2k\boldsymbol{q}^2(t)$. Plus précisément, nous avons

$$E(t) = \frac{1}{2}m\Big(\dot{\boldsymbol{q}}^2(t) + \omega^2\boldsymbol{q}^2(t)\Big) \quad \text{où} \quad \omega = \sqrt{\frac{k}{m}} \tag{65}$$

Le paramètre ω désigne la pulsation (vitesse angulaire) caractéristique de l'oscillateur : c'est une quantité homogène à une fréquence : plus précisément $\omega = 2\pi\nu$ où ν est la fréquence propre de l'oscillateur. La pulsation ω, nous permet alors d'écrire l'équation du mouvement sans référence explicite à la masse de l'oscillateur, i.e.

$$\ddot{\boldsymbol{q}}(t) = -\omega^2\boldsymbol{q}(t) \tag{66}$$

Sous cette forme, les conditions initiales $(\boldsymbol{q}(0), \dot{\boldsymbol{q}}(0))$ nous donnent :

$$\boldsymbol{q}(t) = \boldsymbol{q}(0)\cos(\omega t) + \dot{\boldsymbol{q}}(0)\sin(\omega t) = \rho\sin(\omega t + \varphi)$$

(où $\rho = (\boldsymbol{q}(0)^2 + \dot{\boldsymbol{q}}(0)^2)^{1/2}$ et φ est un angle de déphasage défini modulo 2π). Il découle alors directement de *« l'identité trigonométrique de Pythagore »* que l'énergie de l'oscillateur harmonique telle que définie en (65) ne dépend pas de t, puisque

$$E(t) = \frac{1}{2}m\Big(\omega^2\cos^2(\omega t + \varphi) + \omega^2\sin^2(\omega t + \varphi)\Big) = \frac{1}{2}m\omega^2$$

On dit que l'oscillateur harmonique est un système conservatif, en ce sens que l'énergie mécanique est une constante du mouvement. Ce résultat n'est cependant pas une simple coïncidence trigonométrique, mais plutôt un cas particulier du formalisme hamiltonien (le formalisme hamiltonien est essentiellement une généralisation de ce calcul). Nous pouvons voir cela en introduisant *« l'impulsion »*[31], i.e. $\boldsymbol{p}(t) := m\dot{\boldsymbol{q}}(t)$ et le hamiltonien

30. Comme d'habitude, nous notons ici q la variable position.

31. De manière plus abstraite (c.f. [Oli17]) l'impulsion p est la variable conjuguée de la vitesse $\dot{q}$, la conjugaison ici étant obtenue par *« la transformation de Legendre »* ; le lagrangien $L(q, \dot{q})$ du système étant défini comme la différence entre l'énergie cinétique et l'énergie potentielle, il vient

$$L(q, \dot{q}) = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}m\omega^2 q^2$$

défini pour $(q,p) \in \mathbb{R} \times \mathbb{R}$ par

(67) $$H(q,p) := \frac{p^2}{2m} + V(q) = \frac{p^2}{2m} + \frac{m\omega^2 q^2}{2}$$

de sorte que pour tout t,

(68) $$E(t) = H(\boldsymbol{q}(t), \boldsymbol{p}(t))$$

Maintenant, si nous dérivons (68) par rapport au temps, nous obtenons

$$\begin{aligned} \dot{E}(t) &= \frac{d}{dt} H(\boldsymbol{q}(t), \boldsymbol{p}(t)) \\ &= \dot{\boldsymbol{q}}(t) \frac{\partial H}{\partial q}\big(\boldsymbol{p}(t), \boldsymbol{q}(t)\big) + \dot{\boldsymbol{p}}(t) \frac{\partial H}{\partial p}\big(\boldsymbol{p}(t), \boldsymbol{q}(t)\big) = \dot{\boldsymbol{q}}(t) m\omega^2 \boldsymbol{q}(t) + \dot{\boldsymbol{p}}(t)\boldsymbol{p}(t)/m \end{aligned}$$

soit encore du fait que $\boldsymbol{p}(t) = m\dot{\boldsymbol{q}}(t)$ et en utilisant l'équation (66)

$$\dot{E}(t) = \dot{\boldsymbol{q}}(t) m\omega^2 \boldsymbol{q}(t) + m\ddot{\boldsymbol{q}}(t)\dot{\boldsymbol{q}}(t) = m\dot{\boldsymbol{q}}(t)\big(\omega^2 \boldsymbol{q}(t) + \ddot{\boldsymbol{q}}(t)\big) = 0$$

Nous retrouvons ainsi que l'énergie mécanique est une constante du mouvement.

8.3. Le *« principe de correspondance »* de Schrödinger donne une procédure générale permettant d'obtenir la version quantique des systèmes de la mécanique classique. Nous allons illustrer cette procédure dans le cas de *« l'oscillateur harmonique quantique »* (o.h.q.) unidimensionnel (de masse m et de pulsation ω). Considérons les Q-états sur la sphère unité de l'espace de Hilbert $\mathscr{H} = L^2(\mathbb{R})$; alors, la position classique q correspond à *« l'opérateur position »*, soit :

$$f(q) \mapsto Qf(q) = qf(q)$$

l'impulsion classique p correspondant elle à *« l'opérateur impulsion »* soit

$$f(q) \mapsto Pf(q) = \boldsymbol{i}\hbar \frac{df}{dq}$$

L'expression classique du hamiltonien $H(q,p)$ en (67) nous donne la forme du Q-hamiltonien de l'o.h.q. noté (abusivement) H soit :

(69) $$H = \frac{1}{2m} P^2 + \frac{m\omega^2}{2} Q^2$$

de sorte que pour tout Q-état $\psi(q)$:

(70) $$H\psi(q) = -\frac{\hbar^2}{2m} \frac{d^2\psi}{d^2 q} + \frac{m\omega^2}{2} q^2 \psi(q)$$

Les opérateurs P, Q et H ne sont pas des opérateurs bornés sur $\mathscr{H}$. C'est le Q-hamiltonien H qui détermine le domaine des Q-états dans lesquels le système peut (théoriquement) évoluer. Nous avons vu (c.f. [Oli16b]) que la version auto-adjointe de H en (70) est $(H, \mathscr{H}^2)$, où $\mathscr{H}^2$ est l'espace de Sobolev des fonctions de $\mathscr{H}$ qui sont deux fois différentiables au sens faible et de dérivée dans $\mathscr{H}$ (= $L^2(\mathbb{R})$).

Par sa définition abstraite, nous retrouvons l'expression de l'impulsion :

$$p = \frac{\partial L}{\partial \dot{q}} = m\dot{q}$$

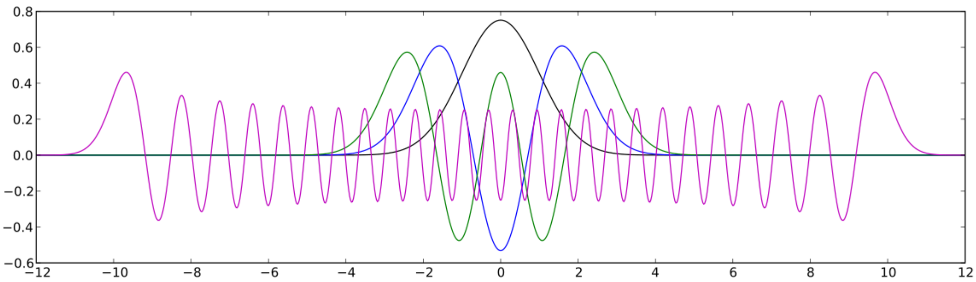


FIGURE 2. *Représentation des états propres de l'o.h.q.*

$$\phi_n(q) = \left(\frac{1}{\sqrt{\pi}\,2^n\,n!}\right)^{1/2} \exp\left(-\frac{q^2}{2}\right) C_n\,(q)$$

pour $n = 0, 2, 4$ *et* 50 *avec* $m\omega/\hbar = 1$ *(voir aussi* [Sch33, Fig. 2 p. 66]*).*

Théorème 8.1. (i) *: Les niveaux d'énergie accessibles à l'o.h.q. de masse-pulsation* (m, ω)*, c'est-à-dire les* $E \in \mathbb{R}$ *pour lesquels il existe un Q-état* ψ *dans* $\mathscr{H}^2$ *t.q.* $H\psi = E\psi$ *(on dit que* ψ *est un Q-état propre de* H*), forment une suite strictement croissante* $0 < E_0 < E_1 < \cdots$ *avec*

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \tag{71}$$

(ii) *: Chaque niveau d'énergie* E_n *est non-dégénéré (i.e. l'espace propre associé à* E_n *est de dimension* 1*) et associé à des Q-états propres proportionnels au Q-état*

$$\phi_n(q) = \sqrt{\frac{1}{2^n\,n!\,\sqrt{\pi}}} \exp\left(-\frac{mq^2\omega}{2\hbar}\right) C_n\left(q\sqrt{\frac{m\omega}{\hbar}}\right) \tag{72}$$

où nous avons introduit « le n*-ème polynôme d'Hermite (physique) »* [32] *(c.f.* [Oli16a, § 7.3]*)*

$$C_n(x) = (-1)^n e^{-x^2}\left(\frac{d}{dx}\right)^n\left(e^{-x^2}\right) \tag{73}$$

(iii) *: Les états stationnaires d'énergie* E_n *sont proportionnels à la fonction d'onde*

$$\psi_n(t, q) = \sqrt{\frac{1}{2^n n!\sqrt{\pi}}} \exp\left(-\frac{m\omega q^2}{2\hbar} - i\omega\left(n + \frac{1}{2}\right)t\right) C_n\left(q\sqrt{\frac{m\omega}{\hbar}}\right)$$

Le reste de la section est dédié à la démonstration de ce théorème. L'inégalité de Heisenberg, nous permet d'établir le résultat important suivant.

Lemme 8.2. *Si* E *est un niveau d'énergie de l'o.h.q. de masse-pulsation* (m, ω)*, alors* $E \geq \hbar\omega/2$*.*

Preuve. Nous nous plaçons le long de la Q-trajectoire stationnaire $\psi_t := e^{-iEt/\hbar}\phi$, où ϕ et un Q-état propre associé au niveau d'énergie E. Quitte à effectuer une translation spatiale adéquate (i.e. remplacer $\phi(q)$ par $\phi(q - \alpha)$ pour $\alpha = \int q|\phi(q)|^2 dq$), nous pouvons supposer

32. Ne pouvant utiliser la lettre H pour désigner les polynômes d'Hermite, nous avons utilisé la lettre C en référence au prénom du grand mathématicien. Dans [Sch33, eq. (27) p. 50], Schrödinger considère les polynômes d'Hermite tels que définis en (73), i.e. les polynômes d'Hermite *physiques* et fait référence à [CH24, V § 9 p. 261 eq. 43 & II § 10, 4, p. 76] pour les propriétés mathématiques correspondantes.

que $\langle\psi_t|Q|\psi_t\rangle = 0$. D'après le théorème d'Ehrenfest (c.f. Théorème 6.4) nous avons aussi $\langle\psi_t|P|\psi_t\rangle = 0$: ainsi, $\langle\psi_t|Q^2|\psi_t\rangle = \sigma_t^2(Q)$ et $\langle\psi_t|P^2|\psi_t\rangle = \sigma_t^2(P)$ et donc

$$E = \langle\psi_t|H|\psi_t\rangle = \frac{1}{2m}\langle\psi_t|P^2|\psi_t\rangle + \frac{m\omega^2}{2}\langle\psi_t|Q^2|\psi_t\rangle = \frac{1}{2m}\sigma_t^2(P) + \frac{m\omega^2}{2}\sigma_t^2(Q)$$

L'inégalité de Heisenberg (c.f. Théorème 5.4) assurant que $\sigma_t(Q)\sigma_t(P) \geq \hbar/2$ il vient $\sigma_t(P) \geq \hbar/(2\sigma_t(Q))$ de sorte que[33]

$$E \geq \frac{\hbar\omega}{2}\left(\frac{1}{4}\frac{\hbar}{m\omega}\frac{1}{\sigma_t^2(Q)} + \frac{m\omega}{\hbar}\sigma_t^2(Q)\right) = \frac{\hbar\omega}{2}\left(\frac{1}{4a^2\sigma_t^2(Q)} + a^2\sigma_t^2(Q)\right) \geq \frac{\hbar\omega}{2}$$

□

8.4. Il est utile de rendre l'équation de Schrödinger adimensionnelle : partant du fait que le quantum d'énergie associé à la pulsation ω vaut $\hbar\omega$, l'idée est d'effectuer un changement de la variable spatiale afin de *factoriser* $\hbar\omega$ dans l'expression du Q-hamiltonien H tel que défini en (69). Si nous posons $x = aq$ avec

$$a = \sqrt{\frac{m\omega}{\hbar}}$$

alors, pour toute fonction $\psi(x)$ de $\mathscr{H}^2$:

$$H\,|\psi(aq)\rangle = -\frac{\hbar^2}{2m}a^2\frac{d^2\psi}{dq^2}(aq) + \frac{\hbar\omega}{2}(aq)^2\psi(aq) = \hbar\omega\left(-\frac{1}{2}\frac{d^2\psi}{dq^2}(aq) + \frac{1}{2}(aq)^2\psi(aq)\right)$$

Cela nous mène à la proposition suivante.

Proposition 8.3. *Si nous définissons (dans la variable $x = aq$) les observables de position et d'impulsion soient respectivement*

$$X : \psi \mapsto x\psi(x) \quad \textit{et} \quad Y : \psi(x) \mapsto -\boldsymbol{i}\frac{d\psi}{dx} \tag{74}$$

ainsi que le hamiltonien adimensionnel

$$\tilde{H} := \frac{1}{2}X^2 + \frac{1}{2}Y^2 \tag{75}$$

alors nous avons l'équivalence

$$H\psi(aq) = E\psi(aq) \iff \tilde{H}\psi(x) = \frac{E}{\hbar\omega}\psi(x) \tag{76}$$

Si E est un niveau d'énergie, alors $\xi = E/(\hbar\omega)$ est une quantité sans dimension appelée *« niveau normalisé »* : grâce à (76) nous sommes amenés à résoudre l'équation aux valeurs propres de *« l'équation de Schrödinger adimensionnelle »*, c'est-à-dire à chercher les niveaux normalisés ξ et les Q-états propres $\psi \in \mathscr{H}^2$ t.q.

$$\tilde{H}\psi = \xi\psi \tag{77}$$

Nous allons suivre la méthode utilisée par Dirac en introduisant les *« opérateurs création et annihilaton »*. Cela consiste à remarquer que dans *« l'algèbre non commutative »* des endomorphismes densément définis sur $\mathscr{H}$, nous avons (sur $\mathscr{H}^2$) :

$$(X - \boldsymbol{i}Y)(X + \boldsymbol{i}Y) = X^2 + \boldsymbol{i}XY - \boldsymbol{i}YX + Y^2 = X^2 + Y^2 + \boldsymbol{i}[X, Y]$$

33. On vérifie aisément que $1/(2u)^2 + u^2 \geq 1$, pour tout u.

Or, pour toute fonction $f \in \mathscr{H}^2$, nous avons

$$XYf(x) - YXf(x) = -\boldsymbol{i}x\frac{df}{dx} + \boldsymbol{i}\frac{d}{dx}(xf(x)) = \boldsymbol{i}f(x)$$

En d'autres termes, $[X, Y] = \boldsymbol{i}I$ (où I est l'identité de $\mathscr{H}^2$) et par suite le hamiltonien adimensionnel introduit en (75) s'écrit :

(78) $$\tilde{H} = A^\dagger A + I/2$$

où les *« opérateurs création et annihilation »* sont définis en posant respectivement :

$$A^\dagger := \frac{X - \boldsymbol{i}Y}{\sqrt{2}} \quad \text{et} \quad A := \frac{X + \boldsymbol{i}Y}{\sqrt{2}}$$

Les opérateurs $(X, \mathscr{H}^2)$ et $(Y, \mathscr{H}^2)$ étant autoadjoints, l'opérateur création $(A^\dagger, \mathscr{H}^2)$ est l'adjoint de l'opérateur annihilation $(A, \mathscr{H}^2)$. Remarquons aussi que A n'est pas normal : en effet, en utilisant de nouveau l'identité $[X, Y] = \boldsymbol{i}I$, il vient :

$$[A, A^\dagger] = \frac{1}{2}[X + \boldsymbol{i}Y, X - \boldsymbol{i}Y] = \frac{1}{2}\Big([X, -\boldsymbol{i}Y] + [\boldsymbol{i}Y, X]\Big) = -\frac{\boldsymbol{i}}{2}\Big([X, Y] + [X, Y]\Big) = I$$

(et il n'existe donc pas de base hilbertienne de Q-états propres communes à A et $A^\dagger$). Cependant, de cette dernière identité nous tirons que

(79) $$\tilde{H} = A^\dagger A + I/2 = AA^\dagger - I/2$$

Il vient alors d'une part :

(80) $$[\tilde{H}, A] = [AA^\dagger - I/2, A] = [AA^\dagger, A] = AA^\dagger A - AAA^\dagger = A[A^\dagger, A] = -A$$

et de même :

(81) $$[\tilde{H}, A^\dagger] = A^\dagger$$

Les relations (80) et (81) peuvent être utilisées pour déterminer les niveaux normalisés de l'o.h.q. de masse-pulsation (m, ω).

Proposition 8.4. *Si ξ est une valeur propre du hamiltonien adimensionnel $\tilde{H}$ et si ϕ est un Q-état propre associé à ξ – i.e. $\tilde{H}\phi = 1/2(X^2 + Y^2)\phi = \xi\phi$ – alors*

$$(i)\ :\ \xi \geq 1/2 \qquad (ii)\ :\ \tilde{H}A\phi = (\xi - 1)A\phi \qquad (iii)\ :\ \tilde{H}A^\dagger\phi = (\xi + 1)A\phi$$

(Ces relations justifient le nom attribué aux opérateurs A et $A^\dagger$.)

Preuve. Supposons que ξ soit une valeur propre du hamiltonien adimensionnel $\tilde{H}$. Le fait que $\xi \geq 1/2$ est une simple transcription du Lemme 8.2 pour l'équation de Schrödinger adimensionnelle. Maintenant si nous considérons que ϕ un Q-état propre associé à ξ, alors d'une part nous pouvons utiliser (80) pour écrire

$$\tilde{H}A\phi = \big(A\tilde{H} - A\big)\phi = A\tilde{H}\phi - A\phi = (\xi - 1)A\phi$$

et de même avec (81) nous obtenons : $\tilde{H}A^\dagger\phi = (\xi + 1)A^\dagger\phi$.

□

De la Proposition 8.4, nous déduisons une remarque simple mais essentielle.

Proposition 8.5. *Soient* $1/2 \leq \xi < 3/2$ *et* ϕ *un Q-état de* $\mathscr{H}^2$*, alors*

$$\begin{aligned}(i) \quad &: \quad \tilde{H}\phi = \xi\phi \implies A\phi = 0\\ (ii) \quad &: \quad \tilde{H}\phi = \xi\phi \implies \xi = \frac{1}{2}\\ (iii) \quad &: \quad A\phi = 0 \implies \exists\theta \in \mathbb{R}, \quad \phi = e^{i\theta}|0\rangle \quad \textit{avec} \qquad |0\rangle := \frac{1}{\sqrt[4]{\pi}}\exp\left(-\frac{q^2}{2}\right)\end{aligned}$$

Preuve. (i) : Soit ϕ un Q-état t.q. $\tilde{H}\phi = \xi\phi$ pour $1/2 \leq \xi < 3/2$. Alors d'après la partie (ii) de la Proposition 8.4 nous avons $\tilde{H}A\phi = (\xi - 1)\phi$: mais comme $\xi - 1 < 1/2$ et que les valeurs propres de $\tilde{H}$ sont toutes minorées par $1/2$ (c.f. partie (i) de la Proposition 8.4), il est nécessaire $A\phi = 0$.

(ii) : Si $A\phi = 0$, alors nous pouvons utiliser la forme réduite du hamiltonien $\tilde{H}$ donnée en (79), de sorte que $0 = A^\dagger A\phi = \big(\tilde{H} - I/2\big)\phi$ i.e. $\tilde{H}\phi = 1/2\phi$.

(iii) : Soit $f \in \mathscr{H}^2$ dans le noyau de A ; cela signifie que $df/dq + qf(q) = 0$ et par suite :

$$\frac{d}{dq}\Big(\exp(q^2/2)f(q)\Big) = \exp(q^2/2)\left(\frac{df}{dq} + qf(q)\right) = 0$$

Formellement, nous obtenons $f(q) = f(0)\exp\left(-q^2/2\right)$: le fait que la fonction $\exp\left(-q^2/2\right)$ soit dans $\mathscr{H}^2$ assure qu'elle engendre le noyau de A. La conclusion vient en notant $\exp\left(-q^2/2\right)$ peut être normalisée en un Q-état i.e. $|0\rangle = 1/\sqrt[4]{\pi}\exp\left(-q^2/2\right)$.

□

Des Propositions 8.4 & 8.5 nous déduisons que [34] :

Corollaire 8.6. (i) *: La valeur propre minimale (fondamentale) du hamiltonien adimensionnel* $\tilde{H}$ *est* $\xi = 1/2$ *et cette valeur propre est non dégénérée et associée au Q-état (de* $\mathscr{H}^2$*)*

$$|0\rangle = \frac{1}{\sqrt[4]{\pi}}\exp\left(-q^2/2\right) \tag{82}$$

(ii) *: Les valeurs propres de* $\tilde{H}$ *forment une suite croissante* $1/2 = \xi_0 < \xi_1 < \cdots$ *avec*

$$\xi_n = n + 1/2 \tag{83}$$

(iii) *: Chaque valeur propre* ξ_n *est non dégénérée et associée au Q-état propre*

$$|n\rangle = \frac{1}{\sqrt{n!}}A^{\dagger n}|0\rangle \tag{84}$$

où $|0\rangle$ *est le Q-état défini dans l'équation (82) de la Proposition 8.5.*

Preuve. (i) : Nous savons d'après la Proposition 8.4 que les valeurs propres du hamiltonien adimensionnel sont toutes minorées par $1/2$: or il découle de la Proposition 8.5 que $\xi_0 = 1/2$ est effectivement une valeur propre ; de plus la Proposition 8.5 entraîne aussi que ξ_0 est non dégénérée et de Q-état propre $|0\rangle$ de $\tilde{H}$, car le noyau de A est de dimension 1 et engendré par Q-état $|0\rangle$.

(ii) : Pour $n \geq 1$ entier supposons que $n+1/2 \leq \xi < n+3/2$ soit une valeur propre de $\tilde{H}$ et que ϕ_n soit un Q-état propre associé ; alors, d'après la partie (iii) de la Proposition 8.4 (propriété de l'annihilateur), nous avons $HA\phi_n = (\xi - 1)A\phi_n$: or $A^\dagger A = \tilde{H} - I/2$ et donc :

$$\langle A\phi_n|A\phi_n\rangle = \big\langle\phi_n|A^\dagger A|\phi_n\big\rangle = \langle\phi_n|H - 1/2I|\phi_n\rangle = \xi\,\langle\phi_n|\phi_n\rangle - 1/2\,\langle\phi_n|\phi_n\rangle = \xi - 1/2$$

34. Voir aussi [Bas71, Ex. N° 2 p. 319].

Comme nous avons supposé que $\xi > 1/2$, cela signifie en particulier que $A\phi_n$ est non nul : par suite, $\xi - 1$ est une valeur propre de $\tilde{H}$ dans $[n-1/2\,; n+1/2[$ et pour laquelle

(85) $$|\phi_{n-1}\rangle := \frac{1}{\sqrt{\xi - 1/2}} A\,|\phi_n\rangle$$

est un Q-état propre. Par une récurrence descendante (finie) chacun des $\xi - k$ $(k = 1, \ldots, n)$ est une valeur propre de $\tilde{H}$. Pour $k = n$ nous savons donc que $\xi - n$ est une valeur propre t.q. $1/2 \le \xi - n < 3/2$, de sorte que par la partie (iii) de la Proposition 8.5, il est nécessaire que $\xi - n = 1/2$: en conclusion, $\xi_n := n + 1/2$ est l'unique valeur propre de $\tilde{H}$ dans l'intervalle $[n+1/2\,; n+3/2[$, ce qui conclut la preuve de (ii).

(iii) : Nous pouvons déterminer un Q-état propre pour chacune des valeurs propres $\xi_n = n + 1/2$, en l'exprimant à partir du Q-état $|0\rangle$ associé à la valeur propre fondamentale $\xi_0 = 1/2$ (c.f. Proposition 8.5). Le cas $n = 0$ est trivial par définition du Q-état $|0\rangle$; soit alors n un entier strictement positif et soit ϕ est un Q-état propre de la valeur propre $\xi_{n-1} = n - 1/2$. Alors d'après la partie (iii) de la Proposition 8.4 (propriété de l'opérateur création), nous savons que $\tilde{H}A^\dagger\phi = (\xi_{n-1} + 1)A^\dagger\phi = \xi_n A^\dagger\phi$: or

$$\begin{aligned}\left\langle A^\dagger\phi | A^\dagger\phi\right\rangle = \left\langle \phi | AA^\dagger | \phi\right\rangle &= \left\langle \phi | \tilde{H} + 1/2 I | \phi\right\rangle \\ &= \left\langle \phi | \tilde{H}\phi + \phi/2\right\rangle = \left\langle \phi \middle| (n-1/2)\phi + \phi/2\right\rangle = n\,\langle\phi|\phi\rangle = n\end{aligned}$$

Par suite $A^\dagger\phi$ est un élément non nul de $\mathscr{H}^2$ qui se normalise en un Q-état propre, soit $1/\sqrt{n}A^\dagger\phi$, associé à la valeur propre ξ_n. Comme par définition le Q-état $|0\rangle$ engendre le sous-espace propre de la valeur propre fondamentale (non dégénérée) $\xi_0 = 1/2$, une récurrence immédiate sur l'entier $n \ge 0$ montre que

$$|n\rangle := \frac{1}{\sqrt{n!}} A^{\dagger n}\,|0\rangle$$

est bien un Q-état propre de ξ_n.

Nous voulons maintenant montrer que chacune des valeurs propres de $\tilde{H}$ est non dégénérée. Comme nous savons déjà que la valeur propre fondamentale $\xi_0 = 1/2$ est non dégénérée, nous considérons (pour une récurrence) que $n \ge 0$ est un entier pour lequel la valeur propre ξ_n est non dégénérée : dans ce cas, nous savons que le Q-état $|n\rangle$ défini dans la partie (ii), engendre le sous-espace de ξ_n. Maintenant, si ϕ est un Q-état propre de ξ_{n+1}, alors d'après (85) (appliqué avec $\xi = \xi_{n+1} = n + 3/2$) le Q-état $1/\sqrt{n+1}A\phi$ est un Q-état propre de ξ_n $(= \xi_{n+1} - 1)$: par suite, il existe $\theta \in \mathbb{R}$ t.q.

(86) $$\frac{1}{\sqrt{n+1}} A\phi = e^{i\theta}\,|n\rangle$$

Pour conclure, notons que

$$A^\dagger A\phi = (H - I/2)\phi = \left(n + 1 + \frac{1}{2}\right)\phi - \frac{1}{2}\phi = (n+1)\phi$$

de sorte qu'avec (86), nous pouvons écrire

$$\phi = \frac{1}{n+1}A^\dagger A\phi = e^{i\theta}\frac{\sqrt{n+1}}{n+1}A^\dagger\left(\frac{1}{\sqrt{n!}}A^{\dagger n}\,|0\rangle\right) = e^{i\theta}\left(\frac{1}{\sqrt{(n+1)!}}A^{\dagger(n+1)}\,|0\rangle\right)$$

soit encore $\phi = e^{i\theta}\,|n+1\rangle$.

□

Avant de prouver le Théorème 8.1 rappelons (c.f. [Oli16a, § 7.3]) que *« les polynômes d'Hermite physiques »* $C_n(x)$ sont définis pour tout entier $n \geq 0$ par l'identité

$$C_n(x) := (-1)^n e^{x^2} \frac{d^n}{dx^n}\left(e^{-x^2}\right)$$

Lemme 8.7. $(\mathcal{P}_n) \ : \ C_n(x) = e^{x^2/2}\left(x - \frac{d}{dx}\right)^n \left(e^{-x^2/2}\right).$

Preuve. La propriété $(\mathcal{P}_0)$ étant vraie, nous pouvons considérer un entier $n \geq 0$ t.q $(\mathcal{P}_n)$ soit vraie ; alors, nous pouvons écrire que

$$\begin{aligned}
\left(x - \frac{d}{dx}\right)^{n+1}\left(e^{-x^2/2}\right) &= \left(x - \frac{d}{dx}\right)\left(e^{-x^2/2} C_n(x)\right) \\
&= (-1)^n \left(x - \frac{d}{dx}\right)\left(e^{x^2/2}\frac{d^n}{dx^n}\left(e^{-x^2}\right)\right) \\
&= (-1)^n \left(x e^{x^2/2}\frac{d^n}{dx^n}\left(e^{-x^2}\right) - x e^{x^2/2}\frac{d^n}{dx^n}\left(e^{-x^2}\right) - e^{x^2/2}\frac{d^{n+1}}{dx^{n+1}}\left(e^{-x^2}\right)\right) \\
&= (-1)^{n+1} e^{x^2/2}\frac{d^{n+1}}{dx^{n+1}}\left(e^{-x^2}\right) \\
&= e^{-x^2/2}\left((-1)^{n+1} e^{x^2}\frac{d^{n+1}}{dx^{n+1}}\left(e^{-x^2}\right)\right)
\end{aligned}$$

ce qui (après une petite transformation) démontre que $(\mathcal{P}_{n+1})$ est vraie.

□

Preuve du Théorème 8.1. En explicitant l'opérateur $A^\dagger$, nous avons

$$|n\rangle(x) = \frac{1}{\sqrt{n!}} A^{\dagger n} |0\rangle(x) = \frac{1}{\sqrt{2^n n!}}\left(x - \frac{d}{dx}\right)^n |0\rangle(x) = \frac{1}{\sqrt{2^n n!}}\left(x - \frac{d}{dx}\right)^n \left(\frac{1}{\sqrt[4]{\pi}} e^{x^2/2}\right)$$

soit encore d'après le Lemme 8.7

$$|n\rangle(x) = \sqrt{\frac{1}{2^n n! \sqrt{\pi}}}\, e^{-x^2/2} C_n(x),$$

Maintenant nous pouvons exprimer la solutions élémentaire de l'équation de Schrödinger de l'oscillateur harmonique quantique au niveau d'énergie $E_n := \hbar\omega(n + 1/2)$, soit :

$$\psi_n(t,q) = e^{-iE_n t/\hbar} |n\rangle\left(q\sqrt{\frac{m\omega}{\hbar}}\right) = \sqrt{\frac{1}{2^n n! \sqrt{\pi}}} \exp\left(-\frac{m\omega q^2}{2\hbar} - i\omega\left(n + \frac{1}{2}\right)t\right) C_n\left(q\sqrt{\frac{m\omega}{\hbar}}\right)$$

□

9. Postscriptum : Textes choisis

M. Planck (1882)
(Cité par J. L. Heilbron in [Hei88]*)*

« En dépit des grands succès remportés par la théorie atomique jusqu'à présent, il faudra en dernier ressort l'abandonner au profit de l'hypothèse d'une matière continue. »

M. Planck (citations)

« Une nouvelle vérité scientifique ne s'impose pas par la conversion de ses adversaires, mais plutôt parce que la disparition naturelle de ces mêmes adversaires fait place à une nouvelle génération de scientifiques familiarisée avec cette vérité dès le départ. »

« La science ne peut pas résoudre le mystère ultime de la nature et cela, parce qu'en dernière analyse, nous faisons nous-mêmes partie du mystère que nous essayons de résoudre. »

A. Einstein in [Ein05]

« L'énergie, durant la propagation d'un rayon de lumière, n'est pas distribuée de façon continue sur les espaces en constante augmentation, mais se compose d'un nombre fini de quanta d'énergie localisée en des points de l'espace, se déplaçant sans se diviser et étant capable d'être absorbés ou générés comme des entités. [...] Selon le point de vue d'après lequel la lumière incidente est composée de quanta d'énergie [...], la production de rayons cathodiques par la lumière peut être comprise de la manière suivante : la couche de surface du corps est pénétrée par des quanta d'énergie dont l'énergie est partiellement convertie en énergie cinétique des électrons. La conception la plus simple est qu'un quantum de lumière transfère toute son énergie à un électron unique [...] »

L. de Broglie (1994) in [dBL25]

« La relation du quantum n'aurait sans doute pas beaucoup de sens si l'énergie pouvait être distribuée d'une façon continue dans l'espace, mais nous venons de voir qu'il n'en est sans doute pas ainsi. On peut donc concevoir que par suite d'une grande loi de la Nature, à chaque morceau d'énergie de masse propre m_0, soit lié un phénomène périodique de fréquence ν_0 telle que l'on ait :

$$h\nu_0 = m_0c^2$$

ν_0 étant mesurée, bien entendu, dans le système lié au. morceau d'énergie. Cette hypothèse est la base de notre système : elle vaut, comme toutes les hypothèses, ce que valent les conséquences qu'on en peut déduire. Devons-nous supposer le phénomène périodique localisé à l'intérieur du morceau d'énergie ? Cela n'est nullement nécessaire et il [...] est sans doute répandu dans une portion étendue de l'espace. D'ailleurs que faudrait-il entendre par intérieur d'un morceau d'énergie ? L'électron est pour nous le type du morceau isolé celui que nous croyons, peut-être à tort, le mieux connaître ; or d'après les conceptions reçues, l'énergie de l'électron est répandue dans tout l'espace avec une très forte condensation dans une région de très petites dimensions dont les propriétés nous sont d'ailleurs fort mal connues. Ce qui caractérise l'électron comme atome d'énergie, ce n'est pas la petite place qu'il occupe dans l'espace, je répète qu'il l'occupe tout entier, c'est le fait qu'il est insécable, non sudivisible, qu'il forme une unité. »

E. Schrödinger (1926) in the abstract of [Sch26]

« The paper gives an account of the author's work on a new form of quantum theory. § 1. The Hamiltonian analogy between mechanics and optics. § 2. The analogy is to be extended to include real "physical" or "undulatory" mechanics instead of mere geometrical mechanics. § 3. The significance of wave-length ; macro-mechanical and micro-mechanical problems. § 4. The wave-equation and its application to the hydrogen atom. § 5. The intrinsic reason for the appearance of discrete characteristic frequencies. § 6. Other problems ; intensity of emitted light. § 7. The wave-equation derived from a Hamiltonian variation-principle ; generalization to an arbitrary conservative system. § 8. The wave-function physically means and determines a continuous distribution of electricity in space, the fluctuations of which determine the radiation by the laws of ordinary electrodynamics. § 9. Non-conservative systems. Theory of dispersion and scattering and of the "transitions" between the "stationary states." § 10. The question of relativity and the action of a magnetic field. Incompleteness of that part of the theory. »

G. Lochak (1994) in [Loc94, p. 125-126]
(Elève et biographe de de Broglie)

« Schrödinger devait accomplir deux progrès essentiels par rapport à de Broglie. Celui-ci restait encore trop près de la mécanique classique et de l'optique. Il était trop près de la mécanique classique parce qu'il faisait se propager l'onde le long de la trajectoire de la particule, donc le long du rayon de l'onde qu'il considérait comme donné *a priori*, se restreignant ainsi à ce qu'on appelle l'*approximation de l'optique géométrique*. En outre, pénétré de sa géniale analogie entre la mécanique et l'optique, il restait également trop près de l'optique et identifiât un peu trop l'onde qu'il avait découverte à une onde lumineuse, ce qui l'empêcha de trouver l'équation d'ondes de la mécanique ondulatoire [...] Au contraire, Schrödinger, venant en second et se tenant, pour cela, un peu plus loin de la sources, accorda à l'onde de matière son indépendance, tant par rapport à l'optique que par rapport à la mécanique. C'est ainsi qu'il découvrit la fameuse *équation de Schrödinger*. »

B. Bensaude-Vincent (1985) in [BV85]

« [...] Schrödinger définit en 1926, la fameuse équation et poursuit, à travers une série d'articles publiés aux Annalen der Physik, l'élaboration de la mécanique ondulatoire. Son approche est tout à fait différente de celle de Heisenberg et pourtant, avant même d'avoir achevé sa construction, Schrödinger établit l'équivalence formelle entre les deux mécaniques.

Equivalence ne signifie pas réconciliation. Les deux formalismes rivalisent en 1926. Dans l'ensemble c'est Schrödinger qui emporte la préférence. Son formalisme paraît plus intuitif, plus familier, moins déroutant. A l'été de 1926, Schrödinger, invité par Sommerfeld à donner un séminaire à Munich, rallie toute l'assistance en dépit des objections que lui adresse Heisenberg. W. Wien, alors directeur de l'Institut de Munich, aurait même déclaré que Schrödinger venait de démontrer l'absurdité de l'idée de "sauts quantiques" et que tous les problèmes en suspens allaient bientôt se résoudre dans la mécanique ondulatoire.

Heisenberg, désappointé, consulte Bohr qui aussitôt demande à Schrödinger de venir à Copenhague. Une semaine de discussions en septembre 1926, mais aucune réconciliation. Cet échec incite Bohr et Heisenberg à travailler sur les relations entre la mécanique quantique telle qu'ils la conçoivent et les données expérimentales. En février 1927, Heisenberg s'attaque à un problème en apparence très simple : décrire la trajectoire d'un électron dans une chambre de Wilson. De cette réflexion sortiront peu de mois après les fameuses relations d'indétermination, $\Delta p \cdot \Delta q \simeq h$. Sans retracer ici toute la démarche de Heisenberg, nous mentionnerons seulement un point important pour l'étude de la complémentarité : l'insistance d'Heisenberg sur l'influence d'Einstein. Quand il reconstruit sa démarche, Heisenberg prête à Einstein un rôle décisif. La "clé du problème", en 1925, lui aurait été livrée par le souvenir d'une phrase d'Einstein lors d'une conversation : "C'est toujours la théorie qui décide de ce qui peut être observé". »

M. Born (1926)
(Citation tirée du blog D'Alain Bouquet - CNRS)

« On n'obtient pas dans la théorie quantique de réponse à la question : *Quel est l'état après la collision ?* mais seulement une réponse à la question : *Quelle est la probabilité d'un résultat donné après la collision ?* [...] Du point de vue de notre mécanique quantique, il n'y a aucune quantité (Grösze) qui détermine causalement l'effet de la collision pour un événement individuel. Devons-nous espérer découvrir plus tard une telle propriété [. . .] et la déterminer dans les événements individuels ? [...] Pour ma part, je suis incliné à renoncer au déterminisme dans le monde atomique, mais ceci est une question philosophique à laquelle les arguments physiques seuls n'apportent pas de réponse. »

A. Einstein (lettre à M. Born : 12 décembre 1926)

« La mécanique quantique force le respect. Mais une voix intérieure me dit que nous sommes loin du compte. La théorie nous apporte beaucoup de choses, mais elle ne nous rapproche guère des secrets du Vieux. De toute façon, je suis convaincu que Lui, au moins, ne joue pas aux dès ! »

N. Bohr (Conférence de Côme 1927)
(citation tirée de [BV85, p. 238]*)*

« D'après l'essence de la théorie des quanta, nous devons nous contenter de considérer la représentation dans l'espace-temps et le principe de causalité, dont la combinaison est caractéristique des théories classiques, comme des traits complémentaires mais s'excluant mutuellement, de la description de l'expérience, qui symbolise l'idéalisation des possibilités d'observation et de définition. [...]

En réalité, il ne s'agit pas ici de conceptions contradictoires de phénomènes, mais de conceptions complémentaires, qui ne fournissent que par leur combinaison une généralisation naturelle du mode de description classique. »

W. Heisenberg (1958) in [Hei71, p. 34]

« Supposons que nous nous intéressions au mouvement d'un électron dans une chambre de Wilson et que l'on puisse, par une observation d'un genre quelconque, déterminer la position et la vitesse initiale de cet électron. Or cette détermination ne sera pas précise ; elle comportera au minimum les imprécisions provenant des relations d'incertitudes et comportera probablement des erreurs plus fortes, vu les difficultés de l'expérience. La première de ces imprécisions nous permet de traduire le résultat de notre observation dans le schéma mathématique de la théorie quantique. On écrit une fonction de probabilité qui représente la situation expérimentale au moment de la mesure et qui comprend les erreurs de mesure possibles. [...] La fonction de probabilité représente notre connaissance du fait, dans la mesure où un autre observateur pourrait connaitre la position de l'électron avec davantage de précision. »

W. Heisenberg (1958) in [Hei71, p. 183]

« Il est bien connu que la « réduction du paquet d'onde » apparaît toujours dans l'interprétation de Copenhague au moment où l'on passe du possible au réel. La fonction de probabilité, laquelle couvrait une vaste gamme de possibilités, se réduit soudain à une gamme bien plus étroite par le fait que l'expérience a conduit à un résultat spécifique et qu'un certain phénomène s'est effectivement produit. Dans le formalisme, cette réduction demande que ce qu'on appelle l'interférence des probabilités, phénomène le plus caractéristique de la théorie quantique, soit détruite par les interactions (partiellement indéfinissables et irréversibles) du système avec l'appareil de mesure et le reste du Monde. »

W. Heisenberg (1958) in [Hei71, p. 185]

« Parmi les autres adversaires de ce qu'on peut appeler parfois l'interprétation « orthodoxe », Schrödinger a adopté une position spéciale : il attribuerait la « réalité objective » non aux particules, mais aux ondes et il n'est pas prêt à interpréter les ondes comme n'étant « que des ondes de probabilité ». Dans son article intitulé « Existe-t-il des sauts quantiques ? », il tente de nier complètement l'existence de ces sauts (on pourrait mettre en doute le caractère approprié du termes « sauts quantiques » à cet endroit et le replacer par le terme moins provocateur de « discontinuité »). Or le travail de Schrödinger contient tout d'abord certaines erreurs sur l'interprétation usuelle. Il néglige le fait que seules les ondes dans l'espace de configuration (ou « matrices de transformation ») sont des ondes de probabilité dans l'interprétation habituelle, alors que les ondes corpusculaires à trois dimensions, ou ondes de rayonnement, ne le sont pas. Ces dernières possèdent tout autant ou tout aussi peut de « réalité » que les particules ; elles n'ont aucun lien direct avec les ondes de probabilité, mais ont une densité d'énergie et un quantité de mouvement continue, comme le champ électromagnétique dans la théorie de Maxwell. Schrödinger souligne donc avec raison que, sur ce point, les processus ne peuvent être conçus comme étant plus continues qu'ils ne le sont en général. Mais cette interprétation ne peut éliminer l'élément de discontinuité qu'on trouve partout en physique atomique ; tout écran de scintillations ou tout compteur Geiger démontre cela immédiatement. Dans l'interprétation usuelle de la mécanique quantique, cet élément est contenu dans la transition du possible au réel. Schrödinger ne fait lui même aucune contreproposition pour dire comment il a l'intention d'introduire l'élément de discontinuité – observable partout – et cela d'une manière qui diffère de l'interprétation habituelle. »

A. Einstein (Lettre à Schrödinger – 1950)
(citation tirée de [Ein89, p. 251]*)*

« La plupart des physiciens ne voient pas qu'ils sont en train de jouer de façon dangereuse avec la réalité - la réalité vue comme quelque chose d'indépendant de toute observation. Ils croient, de façon arbitraire, que la théorie quantique fournit une explication de la réalité, et même une description complète [...] Il est assez pénible de constater que nous ne sommes encore dans les langes. »

A. Einstein in [Ein34]

« Je crois toujours dans la possibilité d'un modèle de la réalité en une théorie qui représente les choses elles-mêmes et non la probabilité de leur avénement. »

E. Borel (1943) in [Bor43, p. 207 - 208]

« En fait il semble bien que, pour la plupart des savants du XIX^{e} siècle, l'onde lumineuse était réellement l'image d'une sinusoïde plutôt que la vibration d'un milieu hypothétique et doué de propriétés invraisemblables. Les mathématiciens ont, depuis longtemps, l'habitude de raisonner sur des être géométriques imaginaires [...], sans que l'absence de réalité les gêne le moins du monde et les empêche d'obtenir des résultats absolument réels et vérifiables. Il faut souhaiter que le développement des nouvelles théorie permette de les traduire un jour en langage dépouillé de tout appareil mathématique, langage dont les termes seraient liés entre eux par des relations simples et en apparence intuitives, comme c'est le cas des droites, plan et autre figures géométriques imaginaires.

Ce que l'on peut retenir d'essentiel, c'est que, dans ce langage nouveau, la notion de probabilité jouera un rôle fondamental et prendra la place qu'occupe dans la mécanique classique le déterminisme qui résulte de l'existence de solutions uniques à certaines équations différentielles, lorsque l'on donne, sous une forme convenable, les conditions initiales. »

R. Feynman (1963) in [Fey14]

« Dans quelques cas particuliers, l'amplitude peut varier de manière sinusoïdale dans l'espace et dans le temps comme $e^{i(\omega t - k\cdot r)}$, où r est le vecteur position pris depuis une certaine origine. [...] Une telle amplitude varie suivant une fréquence définie ω et un nombre d'onde k. Il se trouve alors que cela correspond à la situation classique limite où nous croirions avoir affaire à une particule dont l'énergie E serait connue et serait reliée à la fréquence par $E = \hbar\omega$ et dont la quantité de mouvement p serait aussi connue et serait reliée au nombre d'onde par $p = \hbar k$. [...] Cela signifie que l'idée de particule est limitée. L'idée de particule – avec une position, une impulsion, etc. – que nous utilisons si souvent est d'un certain point de vue peu satisfaisante. Par exemple, si l'amplitude pour trouver une particule à différentes places est donnée par $e^{i(\omega t - k\cdot r)}$, dont le carré du module est constant, cela veut dire que la probabilité de trouver la particule est la même partout. Cela signifie que nous ne savons pas où elle est – elle peut-être n'importe où – il y a une grande incertitude sur sa position. »

R. Feynman (citations)

« La nature ne nous autorise à calculer que des probabilités. Mais la science n'a pas pour autant trouvé ses limites. »

« Si vous croyez comprendre la mécanique quantique, c'est que vous ne la comprenez pas. »

RÉFÉRENCES

[Bas71] J. Bass. *Cours de mathématiques : Topologie. Intégration. Distributions. Equations intégrales. Analyse Harmonique.* Cours de mathématiques. Masson et Cie, 1971.

[BHJ26] M. Born, W. Heisenberg, and P. Jordan. Zur Quantenmechanik [Eng. trans. in van der Waerden 1968 : « On Quantum Mechanics II »]. *Zeitschrift f. Phys.*, 35 :557–615, 1926.

[BJ25] M. Born and P. Jordan. Zur Quantenmechanik [Eng. trans. in van der Waerden 1968 : « On Quantum Mechanics »]. *Zeitschrift f. Phys.*, 33 :858–888, 1925.

[Bor26] M. Born. Zur Quantenmechanik der stoßvorgänge (Sur la mécanique quantique des processus d'impact). *Zeitschrift für Physik*, 37(12) :863–867, 1926.

[Bor43] É. Borel. *L'évolution de la mécanique*. Bibliothèque de Philosophie scientifique. Flammarion, 1943.

[Bor54] M. Born. The statistical interpretation of quantum mechanics – Nobel Lecture, 1954.

[BV85] B. Bensaude-Vincent. L'évolution de la complémentarité dans les textes de Bohr (1927-1939). *Revue d'histoire des sciences*, 38(3) :231–250, 1985.

[CH24] R. Courant and D. Hilbert. *Methoden der mathematischen Physik I*. Spinger - Berlin, 1924.

[dB29] L. de Broglie. The wave nature of the electron. *Nobel Lectures, Physics 1922-1941 (Elsevier 1965)*, pages 244–256, 1929.

[dB38] L. de Broglie. *Le principe de correspondance et les interactions entre la matière et le rayonnement*. Number n° 704 in Actualités scientifiques et industrielles. Hermann – Paris, 1938.

[dBL25] de Broglie L. Recherche sur la théorie des quanta. *Annales de Physique*, III :3–128, 1925.

[Ein05] A. Einstein. über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt (Sur un point de vue heuristique concernant la production et la transformation de la lumière). *Annalen der Physik*, 17 :132–148, 1905.

[Ein34] A Einstein. *On the method of theoretical physics*. Mein Weltbild, Amsterdam, Querido Verlag, 1934.

[Ein89] A. Einstein. *Oeuvres choisies*, volume I. CNRS - Seuil, 1989.

[Fey14] R. Feynman. *Le cours de physique de Feynman - Mécanique quantique - 2ed*. Dunod, 2014.

[Hei25] W. Heisenberg. Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen [Eng. trans. in van der Waerden 1968 : « Quantum-Theoretical Re-interpretation of Kinematic and Mechanical Relations »]. *Zeitschrift f. Phys.*, 33 :879–893, 1925.

[Hei71] W. Heisenberg. *Physique et philosophie : La science moderne en révolution (Trad. de J. Hadamard)*. Poche, 1971.

[Hei88] J.L. Heilbron. *Planck 1858-1947. Une conscience déchirée*. Belin, 1988.

[Loc94] G. Lochak. *La Géométrisation de la physique*. Flammarion, Paris, 1994.

[Oli16a] E. Olivier. Mécanique quantique I : bases mathématiques. *Bull. d'Inf. App. & App.*, 103 :11–72, 2016.

[Oli16b] E. Olivier. Mécanique quantique II : opérateurs non bornés. *Bull. d'Inf. App. & App.*, 104 :115–153, 2016.

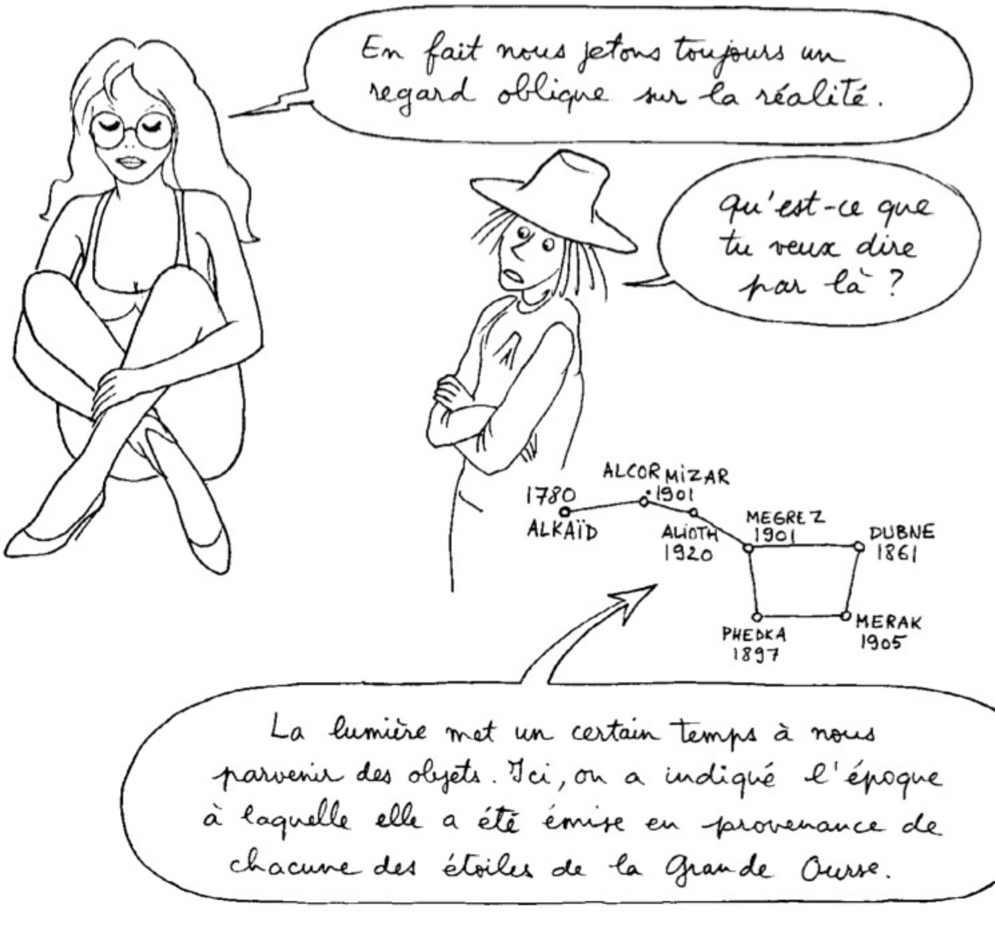
[Oli17] E. Olivier. Mécanique quantique IV : Du principe de Fermat à l'intégrale de chemin de Feynman. *Bull. d'Inf. App. & App.*, 106 :–, 2017.

[Ros04] J. Rosenberg. A Selective History of the Stone-von Neumann Theorem. *Contemporary Mathematics (AMS)*, 365 :331–353, 2004.

[RS81] M. Reed and B. Simon. *I : Functional Analysis*. Methods of Modern Mathematical Physics. Elsevier Science, 1981.

[Sch26] E. Schrödinger. An Undulatory Theory of the Mechanics of Atoms and Molecules. *Phys. Rev.*, 28 :1049–1070, 1926.

[Sch33] E. Schrödinger. *Mémoires sur la mécanique ondulatoire*. Librairie F. Alcan – Reédité par J. Gabay en 1988 – Trad A. Proca, 1933.



VOUZZAVEDIBISAR : Regard oblique (par Jean-Pierre PETIT)

B.I.A.A.

BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
Revue fondée par Edmond Bianco
Publication trimestrielle de l'Université d'Aix-Marseille
ISSN 0291-5413

Le bulletin d'informatique approfondie et applications est une revue pluridisciplinaire destinée à éclairer les connaissances fondamentales informatiques. Les fondements sont un domaine vaste allant de la structure intérieure de l'ordinateur, où se matérialise la machine universelle, à l'algorithme qui devient programme, pour aboutir à la notion de système. Nous contribuons ainsi à ce que les autres disciplines plus anciennes (sciences humaines et de la société, sciences de la matière et de l'énergie, sciences mathématiques, sciences de la nature, sciences de la terre, sciences de l'univers, sciences de la vie, etc.) n'aient pas tendance à considérer l'informatique comme un simple outil définitivement figé. Il importe de continuer à maîtriser les développements fondamentaux de l'informatique pour que nos disciplines puissent en tirer un meilleur parti.
Notre publication est ouverte à l'ensemble de la communauté scientifique. Le périodique est diffusé vers les bibliothèques universitaires de France et vers quelques bibliothèques des cinq continents.