

XI. OPÉRATEURS.

D'une manière générale, un *opérateur* sur un espace vectoriel est une transformation linéaire de cet espace dans lui-même. Si l'espace est de dimension finie, on retrouve la notion bien connue d'endomorphisme. Si on fixe une base, l'opérateur est entièrement défini par sa matrice dans cette base.

Dans un espace vectoriel de dimension infinie, les opérateurs sont évidemment plus difficiles à manipuler. De même que pour les fonctionnelles linéaires, un opérateur linéaire n'est pas forcément continu. Nous ferons une place à part aux opérateurs continus, mais, ayant en vue la Mécanique quantique, nous ne pourrons pas éviter de considérer aussi des opérateurs discontinus, car les plus importants (impulsion, moment cinétique, hamiltonien) sont malheureusement discontinus.

Parmi les espaces de dimension infinie, ceux qui se rapprochent le plus des espaces de dimension finie sont les espaces de Hilbert. En Mécanique quantique, les fonctions d'onde (ou vecteurs d'état) sont justement des vecteurs évoluant dans des espaces du type \mathbb{L}_2 , c'est-à-dire des espaces de Hilbert. Comme nous avons en vue la Mécanique quantique, nous éviterons d'envisager les opérateurs sur d'autres espaces que les espaces de Hilbert. La perte de généralité qui en résulte est sans conséquence pour des physiciens, et aura l'avantage d'éviter bon nombre de complications.

Dans tout le chapitre, on appellera \mathcal{H} l'espace de Hilbert, $\langle \mid \rangle$ le produit scalaire, et $\| \cdot \|$ la norme. On ne précisera la nature de \mathcal{H} que dans les exemples.

1. Opérateurs continus et opérateurs discontinus fermés.

La continuité d'un opérateur linéaire se réduit à la continuité en 0. En effet, si l'opérateur $A : f \mapsto Af$ est continu à l'origine de l'espace \mathcal{H} , on aura par définition de la continuité :

$$\forall \varepsilon > 0, \exists \eta > 0, \|f\| < \eta \implies \|Af\| < \varepsilon \quad (1.1.)$$

Si on remplace f par $f - f_0$ cela donne, puisque $A(f - f_0) = Af - Af_0$:

$$\forall \varepsilon > 0, \exists \eta > 0, \|f - f_0\| < \eta \implies \|Af - Af_0\| < \varepsilon \quad (1.2.)$$

Opérateurs

ce qui est l'expression de la continuité au “point” f_0 .

La continuité d'un opérateur linéaire se traduit aussi par l'existence d'une constante $M > 0$ telle que

$$\forall f \in \mathcal{H}, \quad \|Af\| \leq M\|f\| \quad (1.3.)$$

Prenons en effet $\varepsilon = 1$ dans (1.1); on obtient alors $\exists \eta > 0, \|f\| < \eta \implies \|Af\| < 1$, ce qui, puisque $A(\eta g) = \eta Ag$, équivaut à $\exists \eta > 0, \|f\| < 1 \implies \|Af\| < \frac{1}{\eta}$. Mais pour *tout* $f \in \mathcal{H}$, on peut dire en posant $h = (1/\|f\|)f$, que $\|h\| \leq 1$ d'où $\|Ah\| < \frac{1}{\eta}$ c'est-à-dire $\|Af\| < \frac{1}{\eta}\|f\|$, ce qui est bien (1.3) en prenant $M = \frac{1}{\eta}$. On exprime souvent l'inégalité (1.3) en disant que l'opérateur A est *borné*; c'est pourquoi on trouve souvent dans la littérature l'expression “opérateur borné” (“bounded operator”), pour dire “opérateur continu”.

La plus petite possible de ces constantes M s'appelle la *norme* de l'opérateur A .

Un exemple que nous avons déjà rencontré d'opérateur continu est la *transformation de Fourier* sur $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$. Cet opérateur est continu car on a la formule de Plancherel $\|\hat{f}\| = \sqrt{2\pi}\|f\|$; l'inégalité (1.3) est donc vérifiée avec $M = \sqrt{2\pi}$.

Un autre exemple d'usage très courant est l'opérateur de projection sur un sous-espace fermé. Un tel opérateur vérifie par nature l'inégalité $\|Pf\| \leq \|f\|$. En voici deux cas.

— dans $\mathcal{H} = \mathbb{L}_2([-1, +1])$, la projection P_N sur le sous-espace de dimension finie \mathcal{T}_N des polynômes trigonométriques de degré $\leq N$.

— dans $\mathcal{H} = \mathbb{L}_2([-1, +1])$ ou $\mathcal{H} = \mathbb{L}_2(-\infty, +\infty) = \mathbb{L}_2(\mathbb{R})$, la projection P_+ sur le sous-espace de dimension infinie (mais fermé) \mathcal{H}_+ des fonctions paires.

On rencontre très souvent aussi l'*opérateur de translation* T_a ; si $f \in \mathbb{L}_2(\mathbb{R})$, la fonction T_af est $f(x - a)$. On voit que $\|T_af\| = \|f\|$, puisque par un changement de variable évident on a

$$\int_{-\infty}^{+\infty} |f(x - a)|^2 dx = \int_{-\infty}^{+\infty} |f(x)|^2 dx$$

Toutefois les opérateurs les plus importants de la Mécanique quantique sont discontinus. Voici les principaux.

— Multiplication par x (opérateur de position): c'est l'opérateur qui à une fonction $f(x)$, associe la fonction $xf(x)$.

— Dérivation : c'est l'opérateur qui à une fonction f , associe sa dérivée df/dx . L'opérateur d'impulsion est $i\hbar \cdot d/dx$.

— Plus généralement, on peut considérer la multiplication par une fonction quelconque $V(x)$: c'est l'opérateur qui, à une fonction $f(x)$, associe la fonction $V(x)f(x)$; on a aussi l'opérateur d^n/dx^n ; ainsi l'opérateur *Hamiltonien* est

$$Hf = -\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + V \cdot f$$

Non seulement ces opérateurs sont discontinus, mais ils ne sont même pas partout définis : sur l'espace $\mathbb{L}_2(\mathbb{R})$ il y a des fonctions non dérivables, pour lesquelles on ne peut pas définir l'opérateur de dérivation ; il y a aussi des fonctions qu'on peut certes multiplier par x ou par $V(x)$, mais sans que la fonction produit soit elle aussi dans $\mathbb{L}_2(\mathbb{R})$, de sorte que l'opérateur de multiplication ne peut pas être défini en tant qu'opérateur de $\mathbb{L}_2(\mathbb{R})$ dans lui-même.

Pour de tels opérateurs, on prendra en compte leur domaine de définition : au lieu d'avoir un opérateur A qui transforme *toute* fonction de l'espace de Hilbert \mathcal{H} en une autre fonction de \mathcal{H} , on aura un opérateur A défini sur un sous-espace vectoriel \mathcal{D}_A de l'espace de Hilbert \mathcal{H} , qui transforme toute fonction de \mathcal{D}_A en une fonction de \mathcal{H} .

Remarques. 1. On pourrait aussi convenir que l'opérateur A serait défini sur l'espace de Hilbert \mathcal{H} tout entier, mais qu'en revanche l'espace image serait plus gros que \mathcal{H} : ainsi l'opérateur de multiplication par x , au lieu de n'être défini que sur le sous-espace \mathcal{D}_x des fonctions $f \in \mathbb{L}_2(\mathbb{R})$ telles que $xf \in \mathbb{L}_2(\mathbb{R})$, serait défini sur $\mathbb{L}_2(\mathbb{R})$, mais l'espace-image serait un espace plus grand que $\mathbb{L}_2(\mathbb{R})$, contenant aussi les fonctions $f(x)$ telles que $f(x)/x$ soit de carré intégrable. Mais la pratique montre que ce serait une mauvaise solution : les produits scalaires et autres commodités des espaces de Hilbert (bases, etc.) sont utilisables sur les sous-espaces \mathcal{D}_A , mais ne s'étendent pas forcément à des espaces plus gros.

2. Comme cela a déjà pu être observé pour les fonctionnelles linéaires, la continuité dépend de la métrique. Nous ne considérons ici que des opérateurs sur des espaces de Hilbert, où la métrique est fixée par le produit scalaire. Il est facile de vérifier que par exemple l'opérateur de multiplication par x , défini sur \mathcal{D}_x , est discontinu : les fonctions $f_n(x) = \frac{1}{\sqrt{n}} e^{-x^2/2n}$ ont pour norme $(\pi/n^3)^{1/4}$, donc la suite f_n tend vers zéro dans $\mathbb{L}_2(\mathbb{R})$; d'autre part les fonctions $xf_n(x)$ sont toutes de carré intégrable (ce qui prouve que les f_n sont toutes dans \mathcal{D}_x), et leurs normes sont $\frac{1}{\sqrt{2}}(n\pi)^{1/4}$, ce qui tend vers l'infini et prouve ainsi que l'opérateur est discontinu. Mais si on avait pris sur \mathcal{D}_x la métrique $\|f\| = \sqrt{\|f\|^2 + \|xf\|^2}$, tout en conservant la métrique hilbertienne sur l'espace image, le même opérateur serait alors continu : en effet, avec *n'importe quelle* suite f_n de \mathcal{D}_x , le fait que $\|f_n\|$ tend vers zéro implique automatiquement que $\|xf_n\|_2$ tend vers zéro. Toutefois la nouvelle norme $\|f\|$ dépend de l'opérateur, est fabriquée *ad hoc*, ne dérive d'aucun produit scalaire, et (pour ce qui concerne la Mécanique quantique) ne correspond à rien de physique. Il s'agit d'un exemple scolaire qui sert uniquement ici à rappeler que la discontinuité n'est pas intrinsèque et dépend de la métrique.

Opérateurs

Pour de tels opérateurs, non partout définis, et discontinus, on introduit une nouvelle notion : celle d'opérateur *fermé*

Définition. Soit A un opérateur sur l'espace de Hilbert \mathcal{H} , ayant pour domaine de définition le sous-espace vectoriel dense⁽¹⁾ \mathcal{D}_A . On dit que A est fermé si pour toute suite f_n dans \mathcal{D}_A , telle que f_n et Af_n soient toutes deux convergentes, on a $\mathcal{H}\text{-lim } f_n \in \mathcal{D}_A$ et $\mathcal{H}\text{-lim } Af_n = A \mathcal{H}\text{-lim } f_n$.

On dit que A est préfermé s'il est défini sur un sous-espace dense \mathcal{D} de \mathcal{H} et si pour toute suite f_n de \mathcal{D} qui tend (dans \mathcal{H}) vers zéro et telle que, la suite Af_n soit convergente dans \mathcal{H} , on ait nécessairement $\mathcal{H}\text{-lim } Af_n = 0$.

On vérifie en une ligne qu'un opérateur continu sur \mathcal{H} est forcément fermé, et qu'un opérateur continu sur un sous-espace dense \mathcal{D} de \mathcal{H} est forcément préfermé. L'intérêt de cette nouvelle notion est de donner une possibilité de prolongement pour les opérateurs discontinus, analogue au prolongement par continuité. Imaginons en effet un opérateur qui serait défini a priori sur un sous-espace dense de \mathcal{H} , et qui serait continu : un tel opérateur se prolonge alors par continuité à \mathcal{H} tout entier (tout comme une fonction continue sur \mathbb{Q} se prolonge par continuité à \mathbb{R}).

De la même façon, imaginons un opérateur A qui serait discontinu, et défini sur un sous-espace dense \mathcal{D} . Si l'opérateur A est préfermé, alors on peut le prolonger à un domaine plus grand \mathcal{D}_A , qui est défini comme étant l'ensemble des éléments de \mathcal{H} qui sont des limites de suites f_n de \mathcal{D} telles que f_n et Af_n soient toutes deux convergentes. Si $f \in \mathcal{D}_A$, il existe une suite f_n de \mathcal{D} qui tend vers f , et telle que Af_n tende elle aussi vers une limite g : on prolonge alors A en posant $Af = g$. La définition ci-dessus garantit l'unicité et donc la cohérence de ce prolongement : si on avait pris une autre suite $f_n^{(1)}$ qui tend vers f , telle que $Af_n^{(1)}$ tende vers $g^{(1)}$, on aurait $g^{(1)} = g$.

Remarques. Contrairement au prolongement par continuité, qui permet de prolonger l'opérateur du domaine dense \mathcal{D} à \mathcal{H} tout entier, le prolongement par fermeture permet de prolonger l'opérateur du domaine dense \mathcal{D} à un domaine plus grand (et donc dense aussi) \mathcal{D}_A , mais qui *n'est pas* \mathcal{H} tout entier. Ce domaine étendu \mathcal{D}_A est alors le domaine de définition maximal de l'opérateur ; on peut montrer que si ce domaine maximal est \mathcal{H} tout entier, l'opérateur est continu. Autrement dit :

Théorème (de Banach) : un opérateur fermé dont le domaine est \mathcal{H} tout entier est continu.

L'adhérence \overline{K} d'un sous-ensemble K de \mathcal{H} est définie comme l'ensemble des éléments de \mathcal{H} qui sont des limites de suites convergentes f_n incluses dans K . Un prolongement par continuité s'effectue de K à \overline{K} ; si $K = \mathcal{D}$ qui est dense, l'adhérence est \mathcal{H} , et si A est continu il se prolonge donc de \mathcal{D} à \mathcal{H}

⁽¹⁾ La plupart des auteurs omettent cette restriction ; toutefois, les opérateurs qu'on rencontre en Mécanique quantique ont toujours un domaine dense, de sorte que la considération de domaines non denses ne ferait que nous compliquer la vie, déjà assez difficile comme cela.

Mais \mathcal{D}_A n'est pas l'ensemble des limites de *n'importe quelle* suite convergente de \mathcal{D} ; on ne considère que les suites f_n pour lesquelles Af_n est elle aussi convergente: il y a donc moins de suites admises, et c'est pourquoi \mathcal{D}_A n'a en général aucune raison d'être \mathcal{H} tout entier (lorsque cela arrive, c'est que l'opérateur A est continu, d'après le théorème de Banach mentionné).

On comprendra encore mieux le rapprochement entre l'adhérence $\overline{\mathcal{D}}$ et l'extension \mathcal{D}_A considérant à nouveau la norme $\|f\| = \sqrt{\|f\|^2 + \|Af\|^2}$. Une suite f_n de \mathcal{D} qui converge dans \mathcal{H} ainsi que Af_n est une suite intrinsèquement convergente (ou suite de Cauchy) pour la norme $\| \cdot \|$. Ainsi l'adhérence $\overline{\mathcal{D}}$ de \mathcal{D} est le complété du sous-espace \mathcal{D} pour la norme $\| \cdot \|$, tandis que l'extension \mathcal{D}_A de \mathcal{D} est le complété du sous-espace \mathcal{D} pour la norme $\| \cdot \|$. $\overline{\mathcal{D}} = \mathcal{H}$ est complet pour la norme $\| \cdot \|$, tandis que \mathcal{D}_A est complet pour la norme $\| \cdot \|$.

On peut encore dire que l'ensemble \mathcal{G} des "points" (f, Af) lorsque f parcourt \mathcal{D} est un sous-espace vectoriel de l'espace de Hilbert $\mathcal{H} \times \mathcal{H}^{(2)}$: c'est le graphe de l'opérateur A . Dire que A est fermé équivaut à dire que le graphe de A est un sous-espace vectoriel fermé de $\mathcal{H} \times \mathcal{H}$. Dire que A est préfermé équivaut à dire que l'adhérence $\overline{\mathcal{G}}$ du graphe \mathcal{G} dans $\mathcal{H} \times \mathcal{H}$ est un graphe d'opérateur: l'opérateur A est discontinu en tant que transformation linéaire de \mathcal{H} dans \mathcal{H} , mais continu en tant que transformation linéaire de \mathcal{G} dans \mathcal{H} ; ainsi le prolongement par fermeture est un prolongement par continuité déguisé.

Pour résumer, on peut dire que toute $f \in \mathcal{D}_A$ est la limite d'une suite f_n telle que

- $\forall n, f_n \in \mathcal{D}$;
- $\exists g \in \mathcal{H}, g = \mathcal{H}\text{-lim } Af_n$;

pour prolonger A à \mathcal{D}_A on pose $Af = g$; sur le domaine étendu \mathcal{D}_A , A sera un opérateur fermé.

La notion d'opérateur préfermé est utile pour la raison suivante: comme cela a été dit plus haut, on ne peut malheureusement pas se restreindre à l'étude des seuls opérateurs continus, car ceux de la Mécanique quantique sont discontinus. Par contre les opérateurs de la Mécanique quantique sont préfermés, et même fermés si on convient de toujours considérer leur prolongement maximal. C'est ce que nous ferons dans la suite: les opérateurs (dérivation, etc.) seront définis a priori sur des domaines non nécessairement maximaux, mais (après avoir dûment vérifié qu'ils sont préfermés) on les supposera toujours implicitement prolongés à leur domaine maximal.

Voici une série d'exercices ayant pour but de montrer que les opérateurs de dérivation et de multiplication par une fonction sont bien des opérateurs fermés ou du moins préfermés, ainsi que de préciser le domaine maximal sur lequel on peut les prolonger pour avoir un opérateur fermé.

1. On pose $f_n(x) = e^{-nx^2/2}$; ces fonctions forment une suite sur l'espace $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$.

⁽²⁾ L'espace de Hilbert $\mathcal{H} \times \mathcal{H}$ est muni du produit scalaire $\langle (f_1, g_1), (f_2, g_2) \rangle = \langle f_1, f_2 \rangle + \langle g_1, g_2 \rangle$.

Opérateurs

a) Calculer leurs normes et en déduire que la suite f_n tend vers zéro dans \mathcal{H} .

b) Calculer les normes des dérivées f'_n , constater que la suite f'_n ne tend pas vers zéro dans \mathcal{H} , et en conclure que l'opérateur $f \mapsto f'$ est discontinu.

2. Soit $V(x)$ une fonction donnée, qu'on supposera définie et continue sur $] -\infty, +\infty[$.

a) On définit le domaine $\mathcal{D}_V = \{f \in \mathcal{H} \mid V \cdot f \in \mathcal{H}\}$; vérifier que c'est un sous-espace vectoriel dense de \mathcal{H} (indication: pour $f \in \mathcal{H}$ considérer la suite f_n des fonctions définies par $f_n(x) = f(x)$ si $|x| \leq n$, et $f_n(x) = 0$ si $|x| > n$).

b) Sur le domaine \mathcal{D}_V introduit en a) on définit l'opérateur A_V (multiplication par V): $A_V f$ est la fonction $V(x)f(x)$. Montrer que A_V est fermé (indication: faire jouer l'inégalité de la moyenne dans les intégrales $\int_{|x| \leq M}$ en utilisant le fait que (pour tout M fini) V doit avoir un maximum sur l'intervalle $\{|x| \leq M\}$).

3. On étudie maintenant l'opérateur de dérivation.

a) On prend le domaine \mathcal{D} égal à l'ensemble des fonctions de \mathcal{H} qui sont continuellement dérivables et dont la dérivée est dans \mathcal{H} (ce domaine n'est pas maximal). Vérifier que \mathcal{D} est un sous-espace vectoriel dense dans \mathcal{H} (indication: utiliser des théorèmes du chapitre **X**).

b) Soit f_n une suite de fonctions continuellement dérivables, qui tend vers 0 dans \mathcal{H} , et telle que la suite des dérivées f'_n ait une limite g (il s'agira de montrer que $g = 0$). Montrer que la fonction

$$G(x) = \int_0^x g(t) dt$$

est bien définie et que pour tout x on a

$$|G(x) - f_n(x)| \leq |f_n(0)| + \sqrt{|x|} \|g - f'_n\|$$

(utiliser l'inégalité de Schwarz dans l'intégrale $\int g(t) - f'_n(t) dt$). Montrer aussi que $f_n(0)$ tend vers zéro (utiliser le fait que les f_n , qui tendent vers zéro dans \mathcal{H} , sont dérivables, et leur appliquer habilement la formule des accroissements finis).

c) Prouver (en utilisant b) que pour tout $M > 0$ on a l'inégalité

$$\int_{-M}^{+M} |G(x) - f_n(x)|^2 dx \leq M |f_n(0)| + M^2 \|g - f'_n\|^2$$

et en déduire que $G = 0$.

d) Conclure. Caractériser le domaine maximal de l'opérateur.

2. Valeurs propres et spectre d'un opérateur.

En dimension finie, on dit que λ est valeur propre de l'endomorphisme A s'il existe un vecteur non nul V tel que $AV = \lambda V$. Une manière un peu plus sophistiquée de dire cela est que *le noyau de $\lambda I - A$ n'est pas réduit à 0* (I représentant l'endomorphisme *identité*). Ce noyau est appelé le *sous-espace propre* associé à la valeur propre λ , et sa dimension est appelée la *multiplicité* de la valeur propre λ .

On démontre⁽³⁾ que l'espace vectoriel \mathcal{E} (de dimension finie) sur lequel A opère est la somme directe des sous-espace propres ; sa dimension est donc la somme des multiplicités des valeurs propres. Le nombre de valeurs propres ne peut donc pas excéder la dimension de l'espace.

Si λ n'est égal à aucune des valeurs propres, le noyau de $\lambda I - A$ est réduit à zéro, et l'endomorphisme est injectif, donc inversible. L'ensemble fini des valeurs propres, qui est un sous-ensemble du corps \mathbb{C} , est appelé *le spectre* de l'endomorphisme A . L'application $\lambda \mapsto (\lambda I - A)^{-1}$, appelée *résolvante* de A , est définie sur le complémentaire du spectre, et prend ses valeurs dans l'espace $End(\mathcal{E})$.

En dimension infinie, la situation est plus compliquée. Étant donné un opérateur A sur un espace de Hilbert \mathcal{H} , on peut toujours considérer $\lambda I - A$ qui est aussi un opérateur sur \mathcal{H} , et si son noyau est de dimension $\alpha > 0$, on dit que λ est une valeur propre de multiplicité α . Mais en dehors de cela il y a les phénomènes nouveaux que voici, qui ne peuvent pas se produire en dimension finie :

- le noyau peut être réduit à 0, sans que pour autant $\lambda I - A$ soit inversible (en dimension finie, un endomorphisme injectif est automatiquement surjectif, mais cela n'est pas vrai en dimension infinie), et le spectre peut donc contenir autre chose que des valeurs propres ;

- le noyau peut être réduit à 0, et $\lambda I - A$ être inversible, mais cet inverse peut être discontinu ;

- l'opérateur A n'est lui-même pas forcément continu, et peut n'être défini que sur un sous-espace dense \mathcal{D}_A : si tel est le cas, l'opérateur $(\lambda I - A)^{-1}$ ne peut pas être surjectif (c'est-à-dire avoir pour image \mathcal{H} tout entier), même s'il est bien défini et continu.

Le dernier cas fait bien comprendre pourquoi on a intérêt à considérer des opérateurs fermés, c'est-à-dire ayant un domaine maximal. Imaginons un opérateur A défini sur un sous-espace dense \mathcal{D} , mais préfermé et non

⁽³⁾ Voir par exemple J. Lelong-Ferrand, J. M. Arnaudiès *Algèbre* Dunod, Paris, 1971, chapitre XI.

Opérateurs

fermé, et soit A_0 son extension fermée, définie sur le domaine maximal \mathcal{D}_0 . Il est clair que si $\lambda I - A_0$ est inversible, il *ne pourra pas* en être de même de $\lambda I - A$, car les éléments de $\mathcal{D}_0 - \mathcal{D}$ n'auront pas d'image. Ainsi, pour une opération linéaire telle que la dérivation, le fait pour $\lambda I - A$ d'être inversible ou non dépendrait du choix du domaine et ne serait pas "intrinsèque".

On supposera donc *toujours* dans la suite que les opérateurs sont fermés. Pour faire la théorie spectrale de ces opérateurs, on fixe alors les définitions suivantes.

Définitions. 1. On appelle ensemble résolvant l'ensemble des valeurs complexes de λ pour lesquelles $\lambda I - A$ possède un inverse continu.

N.B. Si λ est résolvant, $\lambda I - A$ applique le domaine maximal \mathcal{D}_A bijectivement sur \mathcal{H} , et son inverse $(\lambda I - A)^{-1}$ applique bijectivement \mathcal{H} sur \mathcal{D}_A , et en outre de façon continue. Ce critère n'est correct que pour des opérateurs fermés.

2. On appelle spectre de A le complémentaire (dans \mathbb{C}) de l'ensemble résolvant.

N.B. Le spectre n'est donc pas (contrairement à ce qui se passe en dimension finie) formé que de valeurs propres; il contient aussi les valeurs λ pour lesquelles $\lambda I - A$ n'est pas surjectif, et même les valeurs de λ pour lesquelles $(\lambda I - A)^{-1}$ existe, mais n'est pas continu; ces cas ne se produisent pas quand la dimension est finie.

3. On appelle indice de nullité de λ , noté $\text{nul}_A(\lambda)$, la dimension du noyau de $\lambda I - A$, et indice de défaut de λ , noté $\text{def}_A(\lambda)$, la codimension de l'image de $\lambda I - A$.

N.B. Si λ est valeur propre, $\text{nul}_A(\lambda)$ est donc sa multiplicité. La codimension d'un sous-espace vectoriel est la dimension de son complémentaire orthogonal (ici, dans \mathcal{H}). En dimension finie, on aurait *toujours* $\text{nul}_A(\lambda) = \text{def}_A(\lambda)$, mais en dimension infinie ces deux paramètres peuvent prendre indépendamment l'un de l'autre n'importe quelle valeur entière ou infinie.

4. On appelle spectre essentiel de A l'ensemble des valeurs de λ pour lesquelles $\text{nul}_A(\lambda) = \text{def}_A(\lambda) = \infty$.

Remarques. Le spectre est une partie du plan complexe, qui est formée de différentes composantes: il y a les valeurs propres, pour lesquelles $\text{nul}_A(\lambda) > 0$ (y compris $\text{nul}_A(\lambda) = \infty$), le spectre essentiel, mais aussi des composantes intermédiaires, correspondant à différentes valeurs des indices. On remarque que les valeurs propres de multiplicité finie sont par nature toujours hors du spectre essentiel. Le spectre essentiel tire son nom du fait qu'il constitue la partie la plus stable du spectre lorsqu'on perturbe l'opérateur.

Théorème 1. *Le spectre est toujours un ensemble fermé.*

Démonstration. On utilise le fait qu'un opérateur proche de l'identité est forcément inversible et son inverse continu; soit A un opérateur borné, et ε un nombre complexe; alors, pour ε assez petit, $I + \varepsilon A$ est inversible et son inverse est continu. Prouvons d'abord ce lemme technique.

Puisque A est borné, il existe une constante M telle que pour toute $f \in \mathcal{H}$, $\|Af\| \leq M\|f\|$. Cela implique que $\|(I + \varepsilon A)f\| \geq \|f\| - |\varepsilon| \|Af\| \geq (1 - |\varepsilon|M)\|f\|$. Cette inégalité prouve déjà que si $|\varepsilon|M < 1$, $I + \varepsilon A$ est injective.

Voyons la surjectivité. Si g est dans \mathcal{H} , soit f la somme de la série $\sum_{n \geq 0} (-1)^n \varepsilon^n A^n g$, où A^n est l'opérateur A itéré n fois. On a donc $\|A^n g\| \leq M^n \|g\|$. Chaque terme de la série peut donc être majoré en norme comme ceci: $\|(-1)^n \varepsilon^n A^n g\| \leq (|\varepsilon|M)^n \|g\|$. Cela montre que si $|\varepsilon|M < 1$, la série est normalement convergente et donc convergente dans \mathcal{H} (puisque \mathcal{H} est complet). Sa somme définit donc un élément f de \mathcal{H} . En appliquant l'opérateur $I + \varepsilon A$ à f , on obtient g , donc g est l'image de f . On a ainsi prouvé que si $|\varepsilon|M < 1$, $I + \varepsilon A$ est surjective.

Le fait que l'inverse soit continu résulte immédiatement de l'inégalité $\|(I + \varepsilon A)f\| \geq (1 - |\varepsilon|M)\|f\|$ obtenue plus haut. En effet, en posant $g = (I + \varepsilon A)f$ ou, ce qui est équivalent, $f = (I + \varepsilon A)^{-1}g$, cette inégalité donne $\|f\| \leq [1/(1 - |\varepsilon|M)] \|g\|$.

En conclusion : si A est un opérateur borné de norme M , alors l'opérateur $I + \varepsilon A$ est inversible si $|\varepsilon| < 1/M$.

Pour montrer que le spectre est un ensemble fermé dans \mathbb{C} , on montre que son complémentaire, l'ensemble résolvant, est ouvert. L'argument consiste à montrer que si λ_0 est résolvant, tout λ assez proche de λ_0 est résolvant aussi. Ou encore : si $\lambda_0 I - A$ est inversible et son inverse continu il en sera de même de $\lambda I - A$ si λ est assez proche de λ_0 . Or on peut écrire

$$\lambda I - A = (\lambda - \lambda_0)I + \lambda_0 I - A = (\lambda_0 I - A)[I + (\lambda - \lambda_0)(\lambda_0 I - A)^{-1}]$$

Si $\lambda_0 I - A$ est inversible et d'inverse continu, appelons M la norme de cet inverse; d'après le lemme, l'opérateur $[I + (\lambda - \lambda_0)(\lambda_0 I - A)^{-1}]$ sera inversible et d'inverse continu si $|\lambda - \lambda_0| < 1/M$, donc aussi $\lambda I - A$. CQFD

On pourrait montrer par des arguments du même type que le spectre essentiel est aussi un ensemble fermé⁽⁴⁾.

On va immédiatement étudier quelques exemples.

⁽⁴⁾ T. Kato *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1966, pages 235 et 242–243.

Opérateurs

Exemple 1. multiplication par une fonction. Soit $V(x)$ une fonction à valeurs complexes, continuellement dérivable, et l'opérateur A qui, à toute fonction $f(x)$ de $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$, associe la fonction $V(x)f(x)$.

Cet opérateur est continu si V est une fonction continue bornée (cela résulte de l'inégalité de la moyenne, car $\|Af\| \leq \sup_{x \in \mathbb{R}} |V(x)| \cdot \|f\|$), mais discontinu dans le cas contraire. Il s'agit d'examiner l'inversion de $\lambda I - A$. Pour cela on considère l'équation $\lambda f - Af = g$, qui se traduit par $\forall x, \lambda f(x) - V(x)f(x) = g(x)$. On obtient l'inverse par division : $f(x) = g(x)/[\lambda - V(x)]$. λ est résolvant si et seulement si l'application $g \mapsto g/(\lambda - V)$ est définie pour tout $g \in \mathcal{H}$ et continue. On va voir que le spectre est l'adhérence de $\text{Im}(V)$, l'ensemble des valeurs prises par la fonction V . Supposons en effet que λ ne soit pas adhérente à $\text{Im}(V)$; il existe alors une distance minimum de λ aux points de $\overline{\text{Im}(V)}$, soit α . Le dénominateur de $f = g/(\lambda - V)$ ne s'annule jamais et en outre on a toujours $|f(x)| \leq \frac{1}{\alpha}|g(x)|$, donc d'après l'inégalité de la moyenne $\|f\| \leq \frac{1}{\alpha}\|g\|$, ce qui prouve la continuité. Inversement, il faut prouver que si λ est adhérente à $\text{Im}(V)$, la transformation $f \mapsto \lambda f - V \cdot f$ ne peut pas être inversée. Nous savons déjà que le spectre est un ensemble fermé, donc il suffit de voir le cas où $\lambda \in \text{Im}(V)$. Soit x_0 tel que $\lambda = V(x_0)$. Nous avons supposé que V est continuellement dérivable (sans cette hypothèse supplémentaire, l'argument deviendrait très compliqué); alors, dans un voisinage de x_0 on aura, d'après la formule des accroissements finis, $|V(x_0) - V(x)| \leq M|x - x_0|$, d'où $|\lambda f(x) - V(x)f(x)| \leq M|x - x_0||f(x)|$ (dans un voisinage de x_0). Cela suffit à prouver que la transformation $f \mapsto \lambda f - V \cdot f$ ne peut pas être inversée : si cela était, une fonction $g \in \mathbb{L}_2(\mathbb{R})$ qui reste par exemple égale à M dans un voisinage de x_0 serait l'image d'une fonction $f(x)$, minorée par $1/|x - x_0|$ dans ce voisinage; or une telle fonction ne peut pas être de carré intégrable.

Exemple 2. Dérivation. L'opérateur est maintenant la dérivation : $Bf = f'$. Mais on peut se ramener au cas précédent par transformation de Fourier. En effet, l'équation $g = \lambda f - Bf = \lambda f - f'$ est équivalente à l'équation $\widehat{g} = \lambda \widehat{f} + i\xi \widehat{f}$, ce qui est le problème précédent, avec $V(\xi) = i\xi$. Comme la transformée de Fourier est un isomorphisme de l'espace $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$, elle conserve les propriétés qui décident si λ est ou non une valeur résolvante : $f \mapsto \lambda f - f'$ est inversible si et seulement si $f \mapsto \lambda f + i\xi f$ l'est; et son inverse est continu si et seulement si $g \mapsto g/(\lambda + i\xi)$ l'est. On en déduit que le spectre de B est l'axe imaginaire.

Remarque. L'opérateur $f \mapsto \lambda f - f'$, tout comme l'opérateur $f \mapsto \lambda f - xf$ ou $f \mapsto \lambda f + i\xi f$, est *toujours* injectif : si $\forall x \in \mathbb{R}, (\lambda - x)f(x) = 0$ ou si $\forall \xi \in \mathbb{R}, (\lambda + i\xi)f(\xi) = 0$, alors $\forall x \in \mathbb{R}, f(x) = 0$. Le spectre

ne contient donc *aucune* valeur propre. Par contre, si la fonction $V(x)$ est constante sur tout un intervalle $[a, b]$, alors les valeurs de λ égales à cette constante sont, pour l'opérateur $f \mapsto Vf$, des valeurs propres de multiplicité infinie (la vérification de cette affirmation est un bon exercice; déterminer aussi les sous-espaces propres correspondants).

Exemple 3. Transformation de Fourier. L'opérateur est $\mathcal{F} : f \mapsto \widehat{f}$. Il s'agit de voir pour quelles valeurs de λ la transformation $\lambda I - \mathcal{F}$ possède un inverse continu.

On va commencer par montrer que les fonctions d'Hermite $\varphi_n(x) = H_n(x) e^{-x^2/2}$ (voir X.27) sont des fonctions propres de \mathcal{F} : $\widehat{\varphi_n} = \sqrt{2\pi} i^n \varphi_n$. Cela se prouve par récurrence. On sait que la transformée de Fourier de $e^{-x^2/2}$ est la fonction $\sqrt{2\pi} e^{-\xi^2/2}$ (chapitre VI, section 2, exemple 1), c'est-à-dire $\widehat{\varphi_0} = \sqrt{2\pi} \varphi_0$. Pour les suivants, on procède par récurrence: supposons que la relation $\widehat{\varphi_n} = \sqrt{2\pi} i^n \varphi_n$ soit vraie pour n , et prouvons qu'elle sera vraie aussi pour $n+1$. Pour cela, on va utiliser la relation de récurrence des polynômes d'Hermite, $H_{n+1} = H'_n - 2x H_n(x)$ (cf. X.23): celle-ci conduit à la relation $\varphi_{n+1} = \varphi'_n - x \varphi_n$ pour les φ_n . En lui appliquant la transformation de Fourier, on obtient $\widehat{\varphi_{n+1}} = -i\xi \widehat{\varphi_n} + i \frac{d}{d\xi} \widehat{\varphi_n}$. En utilisant maintenant l'hypothèse de récurrence, on parvient à $\widehat{\varphi_{n+1}}(\xi) = \sqrt{2\pi} i^{n+1} [-\xi \varphi_n(\xi) + \frac{d}{d\xi} \varphi_n(\xi)] = \sqrt{2\pi} i^{n+1} \varphi_{n+1}(\xi)$, ce qui prouve la récurrence.

Ceci montre que la transformation de Fourier \mathcal{F} sur l'espace $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$ a des valeurs propres; ce qui, rappelons-le, n'était pas le cas des opérateurs $f \mapsto Vf$ ou $f \mapsto f'$. Ces valeurs propres sont $\sqrt{2\pi} i^n$: comme i^n ne prend en réalité que les quatre valeurs 1, i , -1 , et $-i$, on peut dire que \mathcal{F} a pour valeurs propres les quatre nombres complexes $\sqrt{2\pi}$, $\sqrt{2\pi} i$, $-\sqrt{2\pi}$, et $-\sqrt{2\pi} i$. Les sous-espaces propres correspondants sont de dimension infinie; ils sont engendrés (respectivement) par les familles $\varphi_0, \varphi_4, \varphi_8, \varphi_{12}, \varphi_{16}, \dots$, $\varphi_1, \varphi_5, \varphi_9, \varphi_{13}, \varphi_{17}, \dots$, $\varphi_2, \varphi_6, \varphi_{10}, \varphi_{14}, \varphi_{18}, \dots$, et $\varphi_3, \varphi_7, \varphi_{11}, \varphi_{15}, \varphi_{19}, \dots$. Ainsi ces valeurs propres sont de multiplicité infinie.

Une autre question se pose: le spectre est-il composé uniquement de ces valeurs propres, ou contient-il encore autre chose que les valeurs propres? Le fait que λ soit valeur propre signifie que l'opérateur $\lambda I - \mathcal{F}$ est non injectif. Mais (contrairement à ce qui se passe en dimension finie), λ peut être dans le spectre même si $\lambda I - \mathcal{F}$ est injectif: il suffit pour cela que $\lambda I - \mathcal{F}$ soit non surjectif (c'est ce qui arrivait dans le spectre des opérateurs Vf ou f' , voir exemples 1 et 2), ou même, que $\lambda I - \mathcal{F}$, quoique injectif et surjectif, ait un inverse discontinu. Il faut donc contrôler ce qui arrive lorsque λ n'est aucune des quatre valeurs $\sqrt{2\pi}$, $\sqrt{2\pi} i$, $-\sqrt{2\pi}$, et $-\sqrt{2\pi} i$.

Pour cela, on va utiliser le fait que la famille φ_n est une base orthogonale

Opérateurs

de $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$, ou, mieux, que $\phi_n = \varphi_n / \|\varphi_n\| = \varphi_n / \pi^{1/4} 2^{n/2} \sqrt{n!}$ (cf. X.27) est une base orthonormée. Il s'agit d'étudier l'inversion de l'opérateur $\lambda I - \mathcal{F}$, autrement dit, étant donné g quelconque dans \mathcal{H} , trouver $f \in \mathcal{H}$ telle que $\lambda f - \hat{f} = g$. Or les fonctions de $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$ sont des sommes de "séries d'Hermite" de la forme $\sum c_n \phi_n$, telles que les séries numériques $\sum |c_n|^2$ soient convergentes (cf. chapitre **X**, théorème **1**). On peut donc traduire le problème ainsi : "étant donnée une série $g = \sum b_n \phi_n$ telle que $\sum |b_n|^2$ soit convergente, trouver une série $f = \sum a_n \phi_n$ telle que $\sum |a_n|^2$ soit convergente et que $\lambda f - \hat{f} = g$. Or $\lambda f - \hat{f} = \sum a_n [\lambda \phi_n - \widehat{\phi_n}] = \sum a_n [\lambda - \sqrt{2\pi}i^n] \phi_n$, et puisque φ_n est une base, l'équation $\lambda f - \hat{f} = g$ se traduit par

$$\forall n \geq 0, \quad b_n = [\lambda - \sqrt{2\pi}i^n] a_n \quad \Longleftrightarrow \quad \forall n \geq 0, \quad a_n = \frac{b_n}{\lambda - \sqrt{2\pi}i^n}$$

Il est facile de voir que la condition $\sum |b_n|^2 < \infty \implies \sum |a_n|^2 < \infty$ sera satisfaite pour tout λ autre que les quatre valeurs propres, car on peut minorer le dénominateur $\lambda - \sqrt{2\pi}i^n$: si λ est distinct des quatre valeurs propres, il existe un $\alpha > 0$ tel que $\forall n, |\lambda - \sqrt{2\pi}i^n| \geq \alpha$, d'où

$$\forall n \geq 0, \quad |a_n| = \frac{|b_n|}{|\lambda - \sqrt{2\pi}i^n|} \leq \frac{1}{\alpha} |b_n|$$

L'opérateur $\lambda I - \mathcal{F}$ est donc bien inversible. En outre

$$\|f\|^2 = \sum_{n \geq 0} |a_n|^2 \leq \frac{1}{\alpha^2} \sum_{n \geq 0} |b_n|^2 = \frac{1}{\alpha^2} \|g\|^2$$

ce qui prouve la continuité de l'opérateur $(\lambda I - \mathcal{F})^{-1}$.

Conclusion : le spectre de \mathcal{F} est formé uniquement des quatre valeurs propres $\sqrt{2\pi}$, $\sqrt{2\pi}i$, $-\sqrt{2\pi}$, et $-\sqrt{2\pi}i$, toutes de multiplicité infinie.

Exercice : vérifier que l'ensemble de ces quatre valeurs propres est le spectre essentiel de \mathcal{F} .

(conseil : calculer les indices de défaut et de nullité.)

3. L'oscillateur quantique.

Dans cette section, on va illustrer la théorie des opérateurs développée en **1** et **2**, en montrant comment elle s'applique à la Mécanique quantique. L'idée essentielle a été avancée pour la première fois par Erwin Schrödinger (*Quantisierung als Eigenwertproblem II*, Annalen der Physik, vol **79**, 1926, page 489⁽⁵⁾). Dans ce mémoire, la Mécanique quantique est construite

⁽⁵⁾ La traduction française de 1933 a été rééditée récemment : Erwin Schrödinger *Mémoires sur la Mécanique ondulatoire* Editions Jacques Gabay, Paris, 1988. La théorie de l'oscillateur quantique à une dimension, qui est présentée ici avec le formalisme des opérateurs, se trouve aux pages 49 à 53 de cette édition.

en partant de l'analogie entre la Mécanique newtonienne et l'Optique géométrique, selon le postulat (appelé “principe de correspondance”) que *la Mécanique quantique serait à la Mécanique classique ce que l'Optique ondulatoire est à l'Optique géométrique*. Voici un passage du début de ce mémoire :

L'existence d'une *relation intime* entre la théorie de HAMILTON et le phénomène de propagation des ondes est loin d'être quelque chose de nouveau. HAMILTON lui-même le connaissait très bien et l'avait même choisi comme point de départ de sa théorie de la dynamique, laquelle n'était d'ailleurs qu'un prolongement de son *Optique des milieux non homogènes*.

Le principe de variation de HAMILTON (principe de moindre action) peut être considéré comme un principe de FERMAT pour une propagation d'ondes dans l'espace de configuration et son équation aux dérivées partielles n'exprime autre chose que le principe d'Huyghens pour cette même propagation.

et plus loin :

L'hypothèse à laquelle je crois pouvoir attacher un haut degré de certitude est la suivante : la manière correcte de concevoir ou de représenter les phénomènes mécaniques consiste à les rattacher à une *propagation d'ondes* dans l'espace des points représentatifs et non à un mouvement de points représentatifs. L'étude du mouvement des points représentatifs qui forme l'objet de la Mécanique classique, n'est qu'un procédé d'approximation et son emploi est tout aussi peu justifié que l'emploi de l'optique géométrique, ou optique de rayons, dans le cas de phénomènes lumineux réels.

La méthode de HAMILTON mentionnée par Schrödinger est la suivante : dans un milieu non homogène (c'est-à-dire d'indice de réfraction variable $\mathcal{N}(x, y, z)$, la propagation d'ondes a pour équation

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} + k^2 \mathcal{N}(x, y, z)^2 f = 0 \quad (3.1.)$$

où $k = 2\pi/\lambda$ est le nombre d'onde. Lorsque l'onde f se propage, les “surfaces d'onde”, d'équation $W(x, y, z) = C^{\text{te}}$ vérifient l'équation dite *équation eikonale*

$$\left| \frac{\partial W}{\partial x} \right|^2 + \left| \frac{\partial W}{\partial y} \right|^2 + \left| \frac{\partial W}{\partial z} \right|^2 = |\vec{\text{grad}} W|^2 = k^2 \mathcal{N}(x, y, z)^2 \quad (3.2.)$$

qui décrit l'onde selon l'optique géométrique ; les “rayons lumineux” sont alors les lignes orthogonales aux surfaces d'onde.

Hamilton fait remarquer que dans un champ de force dérivant d'un potentiel $V(x, y, z)$, les trajectoires des points matériels d'énergie E sont également les lignes orthogonales à des surfaces d'équation $W(x, y, z) = C^{\text{te}}$, si on prend pour W une solution de l'équation

$$|\vec{\text{grad}} W|^2 = 2(E - V) \quad (3.3.)$$

Opérateurs

dite équation de Hamilton-Jacobi. Cette dernière est formellement semblable à l'équation eikonale 3.2. C'est pourquoi, dans l'ouvrage cité par Schrödinger, Hamilton fait remarquer que les trajectoires des points matériels d'énergie E soumis au champ de potentiel $V(x, y, z)$ sont identiques aux rayons lumineux dans un milieu non homogène d'indice de réfraction $\mathcal{N}(x, y, z)$, si on pose

$$\mathcal{N}(x, y, z) = \sqrt{2[E - V(x, y, z)]} \quad (3.4.)$$

Ainsi, les trajectoires seraient les “rayons lumineux” du phénomène ondulatoire correspondant à l'équation des ondes

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} + 2k^2 [E - V] f = 0 \quad (3.5.)$$

qu'on obtient à partir de 3.1 en remplaçant $\mathcal{N}(x, y, z)^2$ par $2[E - V(x, y, z)]$, conformément à 3.2. Il reste à deviner quelle valeur donner à k ; après une courte évaluation de ce que doit être l'ordre de grandeur de k d'après les données physiques connues, Schrödinger fait le rapprochement avec l'hypothèse (formulée trois ans avant) par de Broglie et conclut que k doit être \sqrt{m}/\hbar , m étant la masse de la particule. Voici encore un passage :

On retrouve ainsi un théorème que M. DE BROGLIE avait établi en faisant largement appel à la théorie de la Relativité, pour les “ondes de phase” d'un électron, dans ses belles recherches⁽⁶⁾ qui ont formé le point de départ du présent travail. On voit qu'il s'agit là au fond d'un théorème beaucoup plus général qui ne découle pas nécessairement de la théorie de la Relativité, mais qui est valable pour n'importe quel système conservatif de la Mécanique classique.

En remplaçant ainsi k par \sqrt{m}/\hbar , l'équation 3.5 devient

$$-\frac{\hbar^2}{2m} \left[\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \right] + V f = E f \quad (3.6)$$

qui est restée sous le nom d'*équation de Schrödinger*.

On remarque que si on interprète le premier membre de cette équation comme un opérateur

$$A : f \longmapsto -\frac{\hbar^2}{2m} \left[\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \right] + V f$$

l'équation 3.6 s'interprète comme l'équation aux valeurs propres de l'opérateur A : la solution de 3.6, si elle existe, est le vecteur propre associé à la valeur propre E .

⁽⁶⁾ L. DE BROGLIE, *Annales de Physique*, vol **3**, 1925, page 22 (thèse, Paris, 1924.

Schrödinger a proposé de poser comme un principe de la Mécanique quantique que

Les niveaux quantiques sont déterminés tous à la fois, comme valeurs propres de l'équation 3.6 qui contient en elle-même ses propres conditions aux limites.

Les conditions aux limites dont il est question implicitement sont les conditions à l'infini : pour que E puisse être une valeur propre de l'opérateur A , il faut que le vecteur propre associé soit une fonction de carré intégrable, c'est-à-dire un élément de l'espace de Hilbert $\mathbb{L}_2(\mathbb{R}^3)$. Ceci est lié à l'interprétation (effectuée au même moment par Max Born) de la fonction f , dont le carré doit représenter une densité de probabilité. Notons en passant que les valeurs et vecteurs propres concernent les états stationnaires ou états liés, mais les valeurs du spectre autres que les valeurs propres correspondent aux processus de collision.

L'idée initiale de Schrödinger, qui par l'intermédiaire des mathématiques conduisait à des conséquences expérimentalement justes, a été développée et systématisée. La synthèse a été faite par le mathématicien John von Neumann et présentée dans son livre paru en 1943 aux États-Unis *Mathematische Grundlagen der Quantenmechanik* (Dover Publ. N.Y.) Le formalisme adopté aujourd'hui encore universellement pour la Mécanique quantique est pour l'essentiel celui proposé par von Neumann, et est entièrement basé sur la théorie des opérateurs dans les espaces de Hilbert. Nous l'appellerons le *formalisme standard* de la Mécanique quantique. Pour l'Histoire, il est intéressant de signaler que la théorie des espaces de Hilbert et des opérateurs est l'œuvre quasi exclusive de l'école *mathématique de Göttingen*, dont Johann von Neumann est issu ; elle a été développée entre 1900 et 1920 (en gros), donc *avant la Mécanique quantique*. En 1933, cette école a déménagé à Princeton (New Jersey) ; seuls les membres les plus âgés, dont David Hilbert, sont restés en Allemagne.

Les principes de ce *formalisme standard* de la Mécanique quantique se résument comme suit.

1. L'état de préparation d'un système quantique est décrit par un élément d'un espace de Hilbert \mathcal{H} , qui est toujours un espace $\mathbb{L}_2(X)$, mais X ("l'espace de configuration") dépend de la nature du système et de son environnement : il ne peut être déterminé qu'à partir du problème classique correspondant.

2. Les grandeurs qu'on peut mesurer physiquement sont représentées par un opérateur autoadjoint A sur l'espace \mathcal{H} ; on ne peut mesurer simultanément, avec une précision arbitraire, que des grandeurs correspondant à des opérateurs qui commutent.

Opérateurs

3. Les valeurs observées sont nécessairement des valeurs propres de ces opérateurs, et pour un système “préparé” dans l’état $\psi \in \mathcal{H}$, la probabilité d’avoir la valeur a_n (valeur propre de A) est le carré du produit scalaire $\langle \psi | f_n \rangle$, où f_n est la fonction propre normée associée à la valeur propre a_n .

Le troisième principe sous-entend que, en vertu du principe des probabilités totales qui veut que la somme des probabilités de toutes les possibilités soit 1, on doit avoir pour toute $\psi \in \mathcal{H}$ de norme 1, la relation

$$\sum_n |\langle \psi | f_n \rangle|^2 = 1$$

la somme étant étendue à toutes les fonctions propres. Cela suppose, comme nous l’avons vu au chapitre **X**, que les fonctions propres de A forment une base. Si on a l’esprit de système (c’était le cas du mathématicien von Neumann), on peut ajouter cette exigence à la définition d’une grandeur physique, mais cela complique l’axiomatique sans bénéfices pratiques. Nous nous contenterons de la constater dans les cas “intéressants”.

Pour illustrer son approche, le premier exemple présenté par Schrödinger (toujours dans le même mémoire, déjà cité) est celui de l’oscillateur quantique à une dimension. Nous allons examiner de près ce problème.

L’équation classique de l’oscillateur est $m\ddot{x} = -Kx$, où K est le paramètre d’élasticité. Le potentiel est $V(x) = \frac{1}{2}Kx^2$, donc l’équation (dite de Schrödinger) qui correspond à 3.6 dans ce cas particulier est

$$-\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + \frac{1}{2}Kx^2 f = E f \quad (3.7.)$$

Nous devons en chercher des solutions de carré intégrable.

Si on interprète ce problème dans le cadre du formalisme *standard*, l’opérateur A correspond à la grandeur *énergie* (A est appelé l’*hamiltonien* du système). Nous allons calculer les valeurs propres E_n (appelées *niveaux d’énergie* du système) et les vecteurs propres normalisés f_n correspondants. D’après le principe 3 mentionné plus haut, les E_n sont les seules valeurs qui peuvent être observées dans un processus de mesure, et si le système a été préparé dans l’état ψ , elles seront observées avec la probabilité $p_n = |\langle \psi | f_n \rangle|^2$; cela signifie que si on reproduit mille fois le processus en préparant chaque fois le système de la même façon, on observera environ $1000 p_1$ fois l’énergie E_1 , $1000 p_2$ fois l’énergie E_2 , $1000 p_3$ fois l’énergie E_3 , etc.

On procède comme pour l’étude de l’opérateur \mathcal{F} à la section **2**, en introduisant les fonctions d’Hermite $\varphi_n(x) = H_n(x) e^{-x^2/2}$. Appliquons

d'abord l'opérateur $A_0 = -\frac{d^2}{dx^2} + x^2$ à ces fonctions. Nous avons déjà vu que les opérateurs $f \mapsto -f''$ et $f \mapsto x^2 f$, quoique discontinus, sont fermés si on les définit sur un domaine maximal. Il n'est pas difficile de vérifier qu'il en sera de même de leur somme A_0 (considérer d'abord A_0 sur le sous-espace dense $\mathcal{S}(\mathbb{R})$, et vérifier qu'il est préfermé). On obtient en dérivant :

$$\frac{d^2 \varphi_n}{dx^2} = \frac{d^2}{dx^2} [H_n(x) e^{-x^2/2}] = [H_n''(x) - 2x H_n'(x) + (x^2 - 1) H_n(x)] e^{-x^2/2}$$

ce qui montre que $A_0 \varphi_n(x) = [-H_n''(x) + 2x H_n'(x) + H_n(x)] e^{-x^2/2}$. Or on sait⁽⁷⁾ que les polynômes d'Hermite vérifient la relation $H_n''(x) - 2x H_n'(x) + 2n H_n(x) = 0$, de sorte que $A_0 \varphi_n(x) = (2n+1) H_n(x) e^{-x^2/2} = (2n+1) \varphi_n(x)$. Cette relation signifie que les fonctions φ_n sont des vecteurs propres de l'opérateur A_0 , les valeurs propres correspondantes étant $2n + 1$. On voit que (contrairement à ce qui se passait pour \mathcal{F}), ces valeurs propres sont toutes distinctes ; ainsi chacune a pour unique vecteur propre associé φ_n , et par conséquent a pour multiplicité 1.

On vérifie par le même procédé que pour \mathcal{F} , que le spectre de A_0 ne contient que ces valeurs propres : si $g \in \mathbb{L}_2(\mathbb{R})$, on a $g = \sum b_n \varphi_n$ avec $\sum |b_n|^2$ convergente, donc la solution de l'équation $\lambda f - A_0 f = g$ est $f = \sum a_n \varphi_n$, où $a_n = b_n / (\lambda - 2n - 1)$. On peut donc énoncer :

Théorème 2. *Le spectre de l'opérateur $A_0 = -\frac{d^2}{dx^2} + x^2$ sur l'espace $\mathbb{L}_2(\mathbb{R})$ est formé de la famille des valeurs propres simples $\lambda_n = 2n + 1$, les vecteurs propres correspondants étant les fonctions d'Hermite $\varphi_n(x) = H_n(x) e^{-x^2/2}$.*

On se ramène ensuite à l'opérateur hamiltonien

$$A = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} K x^2$$

de l'équation 3.7 par un changement de variable linéaire $x \mapsto y = \alpha x$. En effet, $\frac{d^2}{dy^2} = \alpha^2 \frac{d^2}{dx^2}$, donc l'équation 3.7 se transforme en

$$-\frac{\hbar^2}{2m} \frac{1}{\alpha^2} \frac{d^2 f}{dy^2} + \frac{1}{2} K \alpha^2 y^2 f = E f$$

ce qui, en divisant par $\frac{1}{2} K \alpha^2$, donne

$$-\frac{\hbar^2}{m K \alpha^4} \frac{d^2 f}{dy^2} + y^2 f = \frac{2E}{K \alpha^2} f$$

⁽⁷⁾ Si on ne le sait pas, **Exercice :** la démontrer par récurrence à partir de la relation $H_{n+1}(x) = H_n'(x) - 2x H_n(x)$ qui, elle, résulte directement de la définition des polynômes d'Hermite.

Opérateurs

On voit qu'en choisissant $\alpha = [\hbar^2/(mK)]^{1/4}$, cette équation devient $A_0 f = \lambda f$, avec $E = \frac{1}{2} K \alpha^2 \lambda$. Ainsi les valeurs propres de l'opérateur A sont $E_n = \frac{1}{2} K \alpha^2 (2n + 1) = \hbar \omega (n + \frac{1}{2})$, où $\omega = \sqrt{K/m}$, qui est la fréquence (ou plutôt la pulsation) de l'oscillateur classique.

En conclusion : l'équation 3.7, postulée par Schrödinger en vertu de son *principe de correspondance*, n'a de solutions dans $\mathbb{L}_2(\mathbb{R})$ que pour les valeurs de E égales à $\hbar \omega (n + \frac{1}{2})$, et ces solutions sont alors les fonctions $\varphi_n(x [Km]^{1/4} / \sqrt{\hbar})$. La signification physique de cette conclusion est que

a) l'oscillateur, en changeant d'état, ne peut émettre ou absorber de l'énergie que par multiples entiers de $\hbar \omega$;

b) lorsqu'il est dans l'état d'énergie $E_n = \hbar \omega (n + \frac{1}{2})$, la densité de présence de la particule dans l'espace est $H_n(x [Km]^{1/4} / \sqrt{\hbar})^2 e^{-x^2 \sqrt{Km}/\hbar}$.

4. Opérateurs autoadjoints et unitaires.

On va commencer par définir l'*adjoint* d'un opérateur. Cette notion généralise celle d'*endomorphisme adjoint* sur les espaces de dimension finie, ou (une base étant fixée) de *matrice adjointe*. Si \mathcal{E} est un espace hermitien de dimension finie n , et si $A \in \text{End}(\mathcal{E})$, l'endomorphisme adjoint A^* de A est défini par la relation $\forall X \in \mathcal{E}, \forall Y \in \mathcal{E}, \langle AX | Y \rangle = \langle X | A^* Y \rangle$. Cette relation définit entièrement A^* , car il est clair que pour tout $Y \in \mathcal{E}$, $f : X \mapsto f(X) = \langle AX | Y \rangle$ est une forme linéaire sur \mathcal{E} , à laquelle correspond par dualité hermitienne un élément unique Z tel que $\forall X \in \mathcal{E}, f(X) = \langle X | Z \rangle$. Ce vecteur Z dépend linéairement de X et l'endomorphisme adjoint est la transformation $X \mapsto Z$; on pose $Z = A^* X$.

Dans un espace de Hilbert \mathcal{H} de dimension infinie, la situation est plus subtile, car l'existence d'un vecteur Z tel que $f(X) = \langle X | Z \rangle$ n'est garantie que pour les formes linéaires *continues* (théorème de F. Riesz). Si l'opérateur A est lui-même continu, $f : X \mapsto f(X) = \langle AX | Y \rangle$ est continue également et tout se passe bien comme en dimension finie ; mais si A est discontinu et défini seulement sur un domaine dense, on ne peut plus procéder de cette façon. On utilise alors le biais suivant : l'opérateur adjoint ne sera pas défini partout, mais seulement sur un sous-espace \mathcal{D}_{A^*} (comme A lui-même). Plus précisément, on prend pour \mathcal{D}_{A^*} l'ensemble des $h \in \mathcal{H}$ pour lesquels la fonctionnelle linéaire $f \mapsto \langle Af | h \rangle$ est continue. Cette fonctionnelle n'est bien définie a priori que sur \mathcal{D}_A (toujours supposé dense dans \mathcal{H}), mais puisqu'elle est supposée continue et que \mathcal{D}_A est dense, elle se prolonge par continuité à \mathcal{H} tout entier.

On voit qu'il est essentiel de supposer \mathcal{D}_A dense ; sinon, l'opérateur adjoint A^* ne peut pas être défini univoquement sur la foi du théorème

de Riesz.

Le nouveau domaine introduit, \mathcal{D}_{A^*} , est un sous-espace vectoriel de \mathcal{H} ; cela se voit immédiatement : si $f \mapsto \langle Af \mid h_1 \rangle$ et $f \mapsto \langle Af \mid h_2 \rangle$ sont continues, il en sera de même de leur somme, qui est $f \mapsto \langle Af \mid h_1 + h_2 \rangle$, donc $h_1 + h_2 \in \mathcal{D}_{A^*}$, et de même pour λh .

Il faudrait pouvoir s'assurer que \mathcal{D}_{A^*} est dense dans \mathcal{H} . Pour cela, on peut prouver que

Théorème 3. *Si A est un opérateur fermé, le domaine \mathcal{D}_{A^*} est dense, et l'opérateur adjoint A^* est alors lui aussi un opérateur fermé. En outre, l'adjoint de l'adjoint, $(A^*)^*$ (qui est alors bien défini) est identique à A .*

Démonstration. Le meilleur moyen de prouver cela est de considérer les graphes, qui sont des sous-espaces fermés de $\mathcal{H} \times \mathcal{H}$. Soit en effet $G_A = \{(f, Af) \mid f \in \mathcal{D}_A\}$ le graphe de A ; par hypothèse il est fermé. Soit aussi le sous-espace $\widehat{G}_A = \{(-Af, f) \mid f \in \mathcal{D}_A\}$, qui est l'image de G_A par la transformation $(f, g) \mapsto (-g, f)$, et qui évidemment est fermé aussi. La relation $\forall f \in \mathcal{H}, \forall g \in \mathcal{H}, \langle Af \mid g \rangle = \langle f \mid A^*g \rangle$ signifie que A^* a pour graphe le supplémentaire orthogonal de \widehat{G}_A dans $\mathcal{H} \times \mathcal{H}$ (le fait que ce supplémentaire orthogonal est bien un graphe d'opérateur résulte de l'existence, déjà signalée, de l'opérateur adjoint). Pour montrer que \mathcal{D}_{A^*} est dense dans \mathcal{H} il suffit de montrer qu'il n'existe dans \mathcal{H} aucun vecteur non nul f qui soit orthogonal à \mathcal{D}_{A^*} ; or s'il en existait un, le couple $(f, 0)$ serait orthogonal dans $\mathcal{H} \times \mathcal{H}$ à tout couple (g, h) tel que $g \in \mathcal{D}_{A^*}$, donc en particulier à tout couple de G_{A^*} : on aurait $\forall g \in \mathcal{D}_{A^*}, (f, 0) \perp (g, A^*g)$. Puisque le graphe G_{A^*} de A^* est le supplémentaire orthogonal de \widehat{G}_A (qui est fermé), cela impliquerait que $(f, 0)$ est un élément de \widehat{G}_A , donc que $(0, -f)$ est un élément de G_A . Autrement dit, $-f$ serait l'image par A de 0, ce qui est absurde. Par conséquent \mathcal{D}_{A^*} est dense. CQFD

Remarques. Pour définir l'adjoint, il n'est même pas nécessaire que A soit fermé; *pré*fermé suffit : en effet, si A est *pré*fermé et si A_0 est son prolongement fermé, le graphe G_A a pour adhérence le graphe G_{A_0} ; le supplémentaire orthogonal de G_A est le même que celui de G_{A_0} (si $\langle f \mid g \rangle = 0 \forall f \in G_A$, on peut passer à la limite en faisant tendre f vers un élément quelconque de G_{A_0}). Dans les espaces de Hilbert, l'adhérence d'un sous-espace V est $(V^\perp)^\perp$, le supplémentaire du supplémentaire. De même, le supplémentaire orthogonal de \widehat{G}_A est le même que celui de \widehat{G}_{A_0} . On peut dire que si A est *pré*fermé, son prolongement fermé est $(A^*)^*$, l'adjoint de l'adjoint.

L'existence de l'adjoint A^* n'exige pas que A soit fermé ou *pré*fermé, mais seulement que le domaine \mathcal{D}_A sur lequel A est défini soit dense; par contre, si A n'est pas *pré*fermé, le domaine \mathcal{D}_{A^*} de l'adjoint ainsi obtenu ne sera pas dense, et par conséquent $(A^*)^*$ ne sera pas défini. En ajoutant quelques détails à la démonstration du théorème 3, on peut montrer que si \mathcal{D}_A est dense, alors A est *pré*fermé si et seulement si \mathcal{D}_{A^*} est dense.

Bien entendu, ces remarques deviennent triviales pour les opérateurs continus : dans ce cas, l'adjoint existe toujours et est continu, et les domaines sont \mathcal{H} tout entier. Rappelons

Opérateurs

qu'un opérateur fermé dont le domaine est \mathcal{H} tout entier est forcément continu (théorème de Banach).

Exercice. Sur un espace de dimension finie, on peut caractériser les endomorphismes par leur matrice: si l'endomorphisme A a pour matrice $\{a_{ij}\}$, l'endomorphisme adjoint a pour matrice $\{\overline{a_{ji}}\}$. Montrer par un procédé direct et calculatoire que cette relation entre les matrices est bien équivalente à la relation entre les graphes $G_{A^*} \perp \widehat{G}_A$. Examiner plus particulièrement le cas des dimensions 1 et 2.

Définition. On dit qu'un opérateur A (dont le domaine est toujours supposé dense) est autoadjoint, si $A = A^*$.

Si A est autoadjoint, le supplémentaire orthogonal de G_A dans $\mathcal{H} \times \mathcal{H}$ est égal à \widehat{G}_A .

Voici quelques exemples (la plupart déjà rencontrés).

— La multiplication par une fonction réelle (disons continuellement dérivable) $V(x)$, définie sur le domaine $\mathcal{D}_V = \{f \in \mathcal{H} \mid Vf \in \mathcal{H}\}$. Pour déterminer le domaine \mathcal{D}_{V^*} de l'adjoint, on cherche, conformément à la définition, les éléments g pour lesquels la fonctionnelle $f \mapsto \langle Vf \mid g \rangle = \int V(x) f(x) \overline{g(x)} dx$ est continue; or d'après le théorème de Riesz, il en est ainsi si et seulement si Vg est de carré intégrable, ce qui signifie bien que $\mathcal{D}_{V^*} = \mathcal{D}_V$.

— L'opérateur $\mathcal{P} = i\hbar \frac{d}{dx}$, sur le domaine $\mathcal{D}_{\mathcal{P}} = \{f \in \mathcal{H} \mid \widehat{\xi f} \in \mathcal{H}\}$. (même argument que précédemment, mais appliqué aux transformées de Fourier).

— L'opérateur $A = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}Kx^2$. Pour alléger l'écriture on prend $A_0 = -\frac{d^2}{dx^2} + x^2$. En utilisant la base hilbertienne ϕ_n formée par les fonctions d'Hermite normalisées, on peut caractériser le domaine de A_0 comme étant l'ensemble des éléments f de \mathcal{H} dont les "coefficients d'Hermite" $a_n = \langle f \mid \phi_n \rangle$ sont tels que la série numérique $\sum_{n \geq 0} (2n+1)^2 |a_n|^2$ converge. Le domaine de l'opérateur adjoint A_0^* est alors l'ensemble des $g = \sum b_n \phi_n \in \mathcal{H}$ tels que la fonctionnelle $f \mapsto \langle A_0 f \mid g \rangle = \sum (2n+1) a_n b_n$ est continue, ce qui est le cas si et seulement si la série numérique $\sum (2n+1)^2 |b_n|^2$ est convergente: cela montre bien que $\mathcal{D}_{A_0^*} = \mathcal{D}_{A_0}$.

— Plus généralement, l'opérateur $A = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x)$, qui est l'opérateur hamiltonien usuel en Mécanique quantique non relativiste. Dans le cas particulier précédent, on pouvait utiliser une base de vecteurs propres; cela n'est plus possible en général. Interprétons le premier terme de l'opérateur comme une dérivation au sens des distributions, qui existe toujours ($\forall \varphi \in \mathcal{S}(\mathbb{R}), \langle f'', \varphi \rangle = -\int f(x) \varphi''(x) dx$ est toujours bien défini); posons alors $\mathcal{D}_A = \{f \in \mathbb{L}_2(\mathbb{R}) \mid -\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + V f \in \mathbb{L}_2(\mathbb{R})\}$. La fonctionnelle $f \mapsto \int [-\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2}(x) + V(x) f(x)] \overline{g(x)} dx$ est bien définie pour $f \in \mathcal{D}_A$. Si V est une fonction infiniment dérivable à croissance polynomiale, et si f appartient à $\mathcal{S}(\mathbb{R})$, qui est un sous-espace de \mathcal{D}_A , on peut écrire $\langle \overline{g}, -\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + V f \rangle = \langle f, -\frac{\hbar^2}{2m} \frac{d^2 \overline{g}}{dx^2} + V \overline{g} \rangle$ en considérant \overline{g} comme une distribution. Mais cela revient à $\langle -\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + V f \mid g \rangle = \langle f \mid -\frac{\hbar^2}{2m} \frac{d^2 \overline{g}}{dx^2} + V \overline{g} \rangle$. D'après le théorème de Riesz, ceci est une fonctionnelle linéaire continue de f si et seulement si $-\frac{\hbar^2}{2m} \frac{d^2 \overline{g}}{dx^2} + V \overline{g}$ est un élément de $\mathbb{L}_2(\mathbb{R})$. Cette conclusion est obtenue pour $f \in \mathcal{S}(\mathbb{R})$, mais $\mathcal{S}(\mathbb{R})$ est dense dans $\mathbb{L}_2(\mathbb{R})$ donc la fonctionnelle se prolonge par continuité. Ainsi $\mathcal{D}_{A^*} = \mathcal{D}_A$. Si V n'est pas infiniment dérivable, cette argumentation reste

valable à condition d'interpréter correctement $\langle \bar{g}, -\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + V f \rangle = \langle f, -\frac{\hbar^2}{2m} \frac{d^2 \bar{g}}{dx^2} + V \bar{g} \rangle$: c'est la difficulté classique liée à la multiplication des distributions.

— Plus généralement, sur $\mathbb{L}_2(\mathbb{R}^3)$ au lieu de $\mathbb{L}_2(\mathbb{R})$, l'opérateur $A = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2} + V(x, y, z)$ (même argumentation).

Remarques. La relation $A = A^*$ inclut le fait que les domaines sont égaux : $\mathcal{D}_A = \mathcal{D}_{A^*}$. Cela implique que les domaines sont maximaux : un opérateur autoadjoint est *automatiquement* fermé. Si on restreint A à un domaine plus petit que \mathcal{D}_A (mais toujours dense), l'opérateur ainsi diminué sera préfermé et non plus fermé ; son adjoint sera cependant toujours le même. Beaucoup d'auteurs introduisent pour cela la notion d'opérateur *essentiellement* autoadjoint : c'est un opérateur préfermé, dont l'extension fermée est autoadjointe. Pour éviter l'inflation terminologique, on supposera (comme cela a déjà été convenu plus haut) que les opérateurs sont toujours considérés implicitement dans leur extension maximale.

On introduit aussi la notion d'opérateur *symétrique* : c'est un opérateur dont l'adjoint est une extension : A est symétrique si A^* prolonge A ; ainsi un opérateur symétrique est forcément préfermé. Un opérateur symétrique est simplement un opérateur préfermé A pour lequel on a $\forall f \in \mathcal{H}, \forall g \in \mathcal{H}, \langle Af | g \rangle = \langle f | Ag \rangle$. Un opérateur essentiellement autoadjoint est symétrique, mais malheureusement il y a des opérateurs symétriques non essentiellement autoadjoints. On est bien obligé de tenir compte de ces difficultés dans une théorie générale et abstraite des opérateurs, mais pour ce qui nous intéresse on peut s'en passer, car les opérateurs particuliers que nous considérerons seront, eux, autoadjoints (comme le sont les exemples mentionnés).

Théorème 4. *Le spectre d'un opérateur autoadjoint est entièrement contenu dans l'axe réel.*

Démonstration. Si A est autoadjoint, $\langle Af | f \rangle$ est toujours réel (quel que soit $f \in \mathcal{H}$). En effet, le produit scalaire $\langle f | g \rangle$ sur un espace de Hilbert complexe a la propriété que $\overline{\langle g | f \rangle} = \langle f | g \rangle$. D'autre part, si A est autoadjoint, on a aussi $\langle Af | f \rangle = \langle f | Af \rangle = \overline{\langle Af | f \rangle}$, d'où la conclusion. Maintenant, λ étant un nombre complexe quelconque, on peut écrire

$$\langle \lambda f - Af | f \rangle = \lambda \|f\|^2 - \langle Af | f \rangle$$

Ce qui, en ne prenant que la partie imaginaire, donne

$$\Im\{\langle \lambda f - Af | f \rangle\} = \Im\{\lambda\} \|f\|^2 \quad (4.1.)$$

D'autre part, d'après l'inégalité de Schwarz

$$|\langle \lambda f - Af | f \rangle| \leq \|\lambda f - Af\| \cdot \|f\| \quad (4.2.)$$

En combinant 4.1 et 4.2, et en utilisant le fait que le module d'un nombre complexe (ici $\langle \lambda f - Af | f \rangle$) est plus grand que sa partie imaginaire, on obtient l'inégalité :

$$\|\lambda f - Af\| \cdot \|f\| \geq |\Im\{\lambda\}| \|f\|^2$$

Opérateurs

d'où on déduit en simplifiant par $\|f\|$

$$\|\lambda f - Af\| \geq |\Im\{\lambda\}| \|f\| \quad (4.3.)$$

La simplification n'est légitime que si $\|f\| \neq 0$, mais si $\|f\| = 0$, 4.3 sera vraie pour une autre raison. Cette inégalité prouve déjà que si $\Im\{\lambda\} \neq 0$, l'opérateur $\lambda I - A$ est injectif. Pour montrer qu'il est surjectif (toujours si $\Im\{\lambda\} \neq 0$), montrons que son image \mathcal{R} est fermée et que son supplémentaire orthogonal est nul (un sous-espace de \mathcal{H} qui est fermé et de supplémentaire orthogonal nul, est forcément \mathcal{H} tout entier). Les éléments de \mathcal{R} sont les $\lambda f - Af$ lorsque f parcourt \mathcal{D}_A ; soit donc g tel que $\langle \lambda f - Af | g \rangle = 0 \forall f \in \mathcal{D}_A$. Si $g \in \mathcal{D}_A$, on aura en particulier (en prenant $f = g$) $\langle \lambda g - Ag | g \rangle = 0$, ce qui d'après 4.3 implique que $g = 0$. On a ainsi prouvé que tout $g \in \mathcal{D}_A$, orthogonal à \mathcal{R} , est nul. Comme \mathcal{D}_A est dense, cela suffit. D'autre part, \mathcal{R} est fermé: si f_n est une suite de \mathcal{D}_A telle que $g_n = \lambda f_n - Af_n$ tend dans \mathcal{H} vers une limite g , alors g_n est une suite de Cauchy (puisque'elle converge). D'après l'inégalité 4.3 on aura $\|g_n - g_m\| \geq |\Im\{\lambda\}| \|f_n - f_m\|$, qui prouve que f_n est aussi une suite de Cauchy et a donc une limite f dans \mathcal{H} . Ainsi, $g_n = \lambda f_n - Af_n$ et f_n ont toutes deux une limite dans \mathcal{H} , donc aussi Af_n , et puisque l'opérateur A est fermé, on a nécessairement $g = \lambda f - Af$, ce qui prouve bien que \mathcal{R} est fermé, et par conséquent égal à \mathcal{H} (On aurait pu énoncer séparément que l'image d'un opérateur fermé est fermée).

On voit ainsi que pour $\Im\{\lambda\} \neq 0$, l'opérateur $\lambda I - A$ est injectif et surjectif. Enfin, le fait que l'inverse est continu résulte directement de l'inégalité 4.3: $\|(\lambda I - A)^{-1}g\| \leq \|g\|/|\Im\{\lambda\}|$. CQFD

Définition. on appelle unitaire un opérateur inversible qui laisse invariant le produit scalaire. L'opérateur U est unitaire s'il est surjectif et si pour tous f, g dans \mathcal{H} , on a $\langle Uf | Ug \rangle = \langle f | g \rangle$.

Un opérateur unitaire est donc forcément continu, puisqu'il conserve la norme. Pour la même raison, il est forcément injectif (si $\|Uf\| = \|f\|$, il est clair que $\|Uf\| = 0 \Rightarrow \|f\| = 0$). Mais contrairement à ce qui se passe en dimension finie, cela ne garantit pas la surjectivité. L'adjoint U^* d'un opérateur unitaire U vérifie les relations $U^*U = UU^* = I$, autrement dit un opérateur unitaire est inversible, et son inverse est unitaire. Les opérateurs unitaires sur les espaces de Hilbert généralisent les matrices unitaires.

Parmi les opérateurs déjà rencontrés, les transformations de Fourier symétriques, les translations, les multiplications par des fonctions de module 1, sont des opérateurs unitaires:

— \mathcal{F}_α , qui à la fonction $f \in \mathbb{L}_2(\mathbb{R})$ associe la fonction

$$g_\alpha(\xi) = \sqrt{\frac{\alpha}{2\pi}} \int_{-\infty}^{+\infty} e^{i\alpha x\xi} f(x) dx$$

est unitaire (pour tout α réel non nul) d'après la relation de Parseval; les cas courants sont $\alpha = 1/\hbar$ (transformation de Fourier de la Mécanique quantique) et $\alpha = 1$ (transformation de Fourier symétrique usuelle);

— La translation T_a , opérateur qui à toute fonction $f \in \mathbb{L}_2(\mathbb{R})$ associe la fonction $f_a(x) = f(x - a)$, est unitaire, comme le montre un changement de variable évident dans l'intégrale

$$\langle T_a f | T_a g \rangle = \int_{-\infty}^{+\infty} f(x - a) \overline{g(x - a)} dx ;$$

— plus généralement, étant donné une fonction $y = \rho(x)$, croissante sur \mathbb{R} , et différentiable ainsi que son inverse, la transformation qui à toute fonction $f \in \mathbb{L}_2(\mathbb{R})$ associe la fonction $f_\rho(x) = f(\rho(x)) \cdot \sqrt{\rho'(x)}$, est unitaire;

— la multiplication par une fonction $V(x) = e^{i\varphi(x)}$ (où la fonction φ est réelle) est aussi un opérateur unitaire, puisque

$$\begin{aligned} \langle Vf | Vg \rangle &= \int_{-\infty}^{+\infty} e^{i\varphi(x)} f(x) \overline{e^{i\varphi(x)} g(x)} dx \\ &= \int_{-\infty}^{+\infty} e^{i\varphi(x)} e^{-i\varphi(x)} f(x) \overline{g(x)} dx . \end{aligned}$$

Théorème 5. *Le spectre d'un opérateur unitaire est entièrement contenu dans le cercle $\{|z| = 1\}$.*

Démonstration. En tous points semblable à celle du théorème 4. On la propose en exercice :

1. Montrer que $\langle \lambda f - Uf | \lambda f - Uf \rangle = (|\lambda|^2 + 1)\|f\|^2 + 2\Re\{\lambda \langle f | Uf \rangle\}$.
2. Majorer $|\Re\{\lambda \langle f | Uf \rangle\}|$ à l'aide de l'inégalité de Schwarz, et en déduire que

$$\|\lambda f - Uf\|^2 \geq (|\lambda| - 1)^2 \|f\|^2$$

3. Montrer que si $|\lambda| \neq 1$, $\lambda I - U$ et $\overline{\lambda}I - U^{-1}$ sont injectifs.

4. Pour montrer que si $|\lambda| \neq 1$, $\lambda I - U$ est surjectif, montrer d'abord que $(\lambda I - U)\mathcal{H}$ (image de \mathcal{H} par $\lambda I - U$) est fermé, en considérant des suites f_n telles que $(\lambda I - U)f_n$ converge et en utilisant 2; puis montrer en utilisant 3 que si g est orthogonal à tous les $(\lambda I - U)f$, il est forcément nul (en cas de difficulté, regarder la démonstration du théorème 4).

Opérateurs

5. Utiliser à nouveau 2 pour montrer que l'inverse de $\lambda I - U$ est continu.
6. Peut-on étendre ce procédé à d'autres cas ?

5. Fonctions d'opérateurs.

Dans les espaces de dimension finie, on définit aisément les puissances d'un endomorphisme A : la n -ième puissance de A , soit A^n , est l'endomorphisme obtenu en itérant n fois A . On peut aussi considérer des polynômes de A : si $P(x) = \sum a_n x^n$, l'endomorphisme $P(A)$ sera $\sum a_n A^n$.

Il est également possible de définir des séries de puissances d'un endomorphisme, mais il faut préciser dans quel sens la série converge. Ainsi, on peut définir l'exponentielle d'un endomorphisme :

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$$

La convergence s'interprète comme suit. Si A est un endomorphisme sur l'espace euclidien de dimension finie \mathcal{E} , il existe une constante M_A telle que $\forall X \in \mathcal{E}, \|AX\| \leq M_A \|X\|$ (rappel : il suffit de prendre pour M_A le maximum de $\|AX\|$ sur la boule $B = \{X \in \mathcal{E} \mid \|X\| \leq 1\}$). L'existence de cette constante M_A correspond à la continuité des opérateurs linéaires sur un espace de dimension finie. Il est clair que si on itère l'endomorphisme A , on aura $\|A^2 X\| \leq M_A \|AX\| \leq M_A^2 \|X\|$, puis $\|A^3 X\| \leq M_A \|A^2 X\| \leq M_A^2 \|AX\| \leq M_A^3 \|X\|$, etc. de sorte que $\|A^n X\| \leq M_A^n \|X\|$. Si on définit $e^A X$ comme la somme de la série $\sum_{n \geq 0} \frac{1}{n!} A^n X$, on voit que cette série converge car le terme général a pour norme $\frac{1}{n!} \|A^n X\|$, ce qui est majoré par $\frac{1}{n!} M_A^n \|X\|$.

Tout ce qui vient d'être fait se transpose tel quel aux opérateurs *continus* sur les espaces de Hilbert de dimension infinie ; en effet, l'inégalité $\|Af\| \leq M_A \|f\|$ est satisfaite par les opérateurs continus, donc la série $\sum \frac{1}{n!} A^n f$ est normalement convergente. Ce qui implique, puisqu'un espace de Hilbert est complet, que cette série est convergente et définit le vecteur $e^A f$. L'opérateur e^A est alors celui qui transforme f en $e^A f$.

La série $\sum \frac{1}{n!} x^n$ a un rayon de convergence infini, ce qui a évité de se poser des questions sur la valeur de la constante M_A . Mais on voit que si on considère une série $\sum a_n x^n$ ayant un rayon de convergence fini R , il faudra que $M_A < R$ pour définir la somme $\sum a_n A^n$. Voici un exemple :

Théorème 6 *Si A est un opérateur continu sur un espace de Hilbert \mathcal{H} , tel que $\forall f \in \mathcal{H}, \|Af\| \leq M_A \|f\|$, avec $M_A < 1$, alors l'opérateur $I - A$ est inversible.*

Démonstration. On a déjà démontré et utilisé ce résultat (voir la démonstration du théorème 1, section 2). Tout l'argument consiste simplement à remarquer que l'inverse de $I - A$ est la série $\sum_{n \geq 0} A^n$.

Exercice. En utilisant ce même argument, montrer que le spectre d'un opérateur continu est borné.

On peut généraliser ce résultat :

Théorème 7 Soit $\Phi(z)$ une fonction analytique dans le disque $\{|z| < R\}$, et soit A un opérateur continu sur un espace de Hilbert \mathcal{H} , tel que $\forall f \in \mathcal{H}, \|Af\| \leq M_A \|f\|$, avec $M_A < R$. Alors on peut définir l'opérateur $B = \Phi(A)$. Cet opérateur commute avec A : $AB = BA$.

Démonstration. La fonction $\Phi(z)$ étant analytique dans le disque $\{|z| < R\}$, est la somme d'une série $\sum a_n z^n$ de rayon de convergence $\geq R$, donc B sera défini par $Bf = \sum a_n A^n f$. Comme toutes les puissances A^n commutent avec A , il en est de même de B . CQFD

Le fait que les opérateurs commutent a une signification bien précise en Mécanique quantique : les opérateurs représentent des grandeurs physiques, et s'ils commutent ces grandeurs sont compatibles (observables simultanément). Il est logique que si une grandeur b est le carré ou l'exponentielle d'une autre grandeur a , c'est que a et b sont simultanément observables. D'autre part, si on les observe effectivement, on sait (d'après les principes de la Mécanique quantique) que les valeurs observées sont nécessairement des valeurs propres de l'opérateur ; la consistance logique de la Mécanique quantique exige donc aussi que si $B = \Phi(A)$, on ait la même relation pour les valeurs propres $b = \Phi(a)$. C'est bien le cas :

Théorème 8 Soit A un opérateur continu sur un espace de Hilbert \mathcal{H} , et $B = \Phi(A)$, Φ étant comme dans le théorème 7. Alors les valeurs propres de B sont les images par Φ des valeurs propres de A , et les vecteurs propres sont communs. Plus généralement, le spectre de B est l'image par Φ du spectre de A .

Démonstration. Si λ est une valeur propre, soit φ un vecteur propre associé. On a donc $A\varphi = \lambda\varphi$, d'où $A^n\varphi = \lambda^n\varphi$; en reportant dans la série on aura $\Phi(A)\varphi = \Phi(\lambda)\varphi$. Cela montre bien que $\Phi(\lambda)$ est valeur propre de $\Phi(A)$ avec le même vecteur propre associé.

Pour les autres valeurs $\lambda \in \mathbb{C}$, introduisons la fonction

$$\chi(z) = \frac{\Phi(\lambda) - \Phi(z)}{\lambda - z}.$$

cette fonction est analytique dans le même domaine que $\Phi(z)$; on peut donc considérer l'opérateur $\chi(A)$, et on a la relation $\Phi(\lambda) - \Phi(z) = (\lambda - z)\chi(z)$,

Opérateurs

qui sera vraie aussi pour les opérateurs :

$$\Phi(\lambda) I - \Phi(A) = (\lambda I - A) \chi(A) = \chi(A) (\lambda I - A) . \quad (5.1.)$$

Puisque $\chi(z)$ est analytique dans un disque de rayon $> M_A$, l'opérateur $\chi(A)$ est continu. La relation 5.1 montre que si l'opérateur $\Phi(\lambda) I - \Phi(A)$ a un inverse continu, il en sera de même de $(\lambda I - A)$, puisqu'il suffira de poser

$$(\lambda I - A)^{-1} = [\Phi(\lambda) I - \Phi(A)]^{-1} \chi(A) = .$$

Donc le spectre de $\Phi(A)$ contient l'image par Φ du spectre de A . Pour vérifier que non seulement il la contient, mais qu'il lui est identique, supposons un instant qu'il existe un λ dans le spectre de $\Phi(A)$ qui ne soit pas dans l'image ; cela signifie que λ n'est égal à aucun $\Phi(z)$ tel que z soit dans le spectre de A . Donc la fonction $\psi(z) = 1/[\lambda - \Phi(z)]$ est analytique pour tout z appartenant au spectre de A , et donc aussi dans un voisinage Ω de ce spectre. Quoique Ω ne soit pas forcément le disque $|z| < R$, on peut quand même définir l'opérateur $\psi(A)$, qui sera forcément l'inverse continu de $\lambda I - \Phi(A)$.

CQFD

Jusqu'ici, on n'a considéré que des opérateurs continus. Toutefois, comme cela a déjà été dit, les opérateurs qui interviennent en Mécanique quantique ne sont pas forcément continus. On ne peut pas définir les puissances A^n d'un opérateur fermé discontinu aussi facilement que celles d'un opérateur continu, car l'itération exige que l'image du domaine \mathcal{D}_A soit incluse dans \mathcal{D}_A . Si tel n'est pas le cas, on peut poser $\mathcal{D}_{A^2} = \{f \in \mathcal{H} \mid Af \in \mathcal{D}_A\}$, puis $\mathcal{D}_{A^3} = \{f \in \mathcal{H} \mid Af \in \mathcal{D}_{A^2}\}$, et ainsi de suite. Mais le fait que \mathcal{D}_A soit dense n'entraîne pas automatiquement que \mathcal{D}_{A^2} , \mathcal{D}_{A^3} , etc. le soient aussi. Même si c'était vrai, cela ne permettrait encore de définir que des polynômes ; pour définir des séries infinies (par exemple l'exponentielle), il faudrait en outre que l'intersection de tous les \mathcal{D}_{A^n} soit dense.

Pour surmonter cette difficulté, il faut d'abord se restreindre aux opérateurs pour lesquels ça marche. Il se trouve que les opérateurs auto-adjoints font partie de ces privilégiés.

Une approche possible consiste à se ramener aux opérateurs continus par une transformation adéquate. Par exemple J. von Neumann a proposé d'utiliser la *transformation de Cayley*. Si A est un opérateur autoadjoint non nécessairement continu, son spectre est réel donc $(iI - A)^{-1}$ est continu. Le *transformé de Cayley* de A est alors l'opérateur $U = (iI + A)(iI - A)^{-1}$, qui est unitaire :

Théorème 9. *Si A est autoadjoint, $U = (iI + A)(iI - A)^{-1}$ est unitaire (transformation de Cayley). Inversement, si U est un opérateur unitaire*

tel que $U + I$ soit injectif et $\mathcal{D} = (U + I)\mathcal{H}$ dense, alors l'opérateur $A = -i(U - I)(U + I)^{-1}$, défini sur \mathcal{D} , est autoadjoint et a pour transformé de Cayley U .

Démonstration. Remarquons d'abord que l'opérateur U est bien défini : puisque A est autoadjoint, son spectre est inclus dans l'axe réel (théorème 4), donc $iI - A$ est inversible et d'inverse continu ; cet inverse transforme n'importe quel $f \in \mathcal{H}$ en un élément de \mathcal{D}_A , auquel on peut donc appliquer $iI + A$. Comme $iI + A$ est lui aussi inversible, il appliquera à son tour \mathcal{D}_A surjectivement sur \mathcal{H} . Il reste encore à vérifier la conservation du produit scalaire.

Pour cela, soient $f, g \in \mathcal{H}$, et posons $p = Uf$ et $q = Ug$; il s'agit de prouver que $\langle p | q \rangle = \langle f | g \rangle$. D'après la définition de U , on voit que $p = Uf \Leftrightarrow (iI + A)^{-1}p = (iI - A)^{-1}f$ et $q = Ug \Leftrightarrow (iI + A)^{-1}q = (iI - A)^{-1}g$. Posons donc $X = (iI + A)^{-1}p = (iI - A)^{-1}f$ et $Y = (iI + A)^{-1}q = (iI - A)^{-1}g$; X et Y sont ainsi des éléments du domaine \mathcal{D}_A , et on a $f = (iI - A)X$, $g = (iI - A)Y$, $p = (iI + A)X$, $q = (iI + A)Y$. Calculons les produits scalaires $\langle p | q \rangle$ et $\langle f | g \rangle$:

$$\begin{aligned}\langle p | q \rangle &= \langle iX + AX | iY + AY \rangle \\ &= \langle iX | iY \rangle + \langle iX | AY \rangle + \langle AX | iY \rangle + \langle AX | AY \rangle \\ &= \langle X | Y \rangle + i\langle X | AY \rangle - i\langle AX | Y \rangle + \langle AX | AY \rangle \\ \langle f | g \rangle &= \langle iX - AX | iY - AY \rangle \\ &= \langle iX | iY \rangle - \langle iX | AY \rangle - \langle AX | iY \rangle + \langle AX | AY \rangle \\ &= \langle X | Y \rangle - i\langle X | AY \rangle + i\langle AX | Y \rangle + \langle AX | AY \rangle\end{aligned}$$

Comme A est autoadjoint, le second terme est dans les deux cas annulé par le troisième, d'où

$$\begin{aligned}\langle p | q \rangle &= \langle X | Y \rangle + \langle AX | AY \rangle \\ \langle f | g \rangle &= \langle X | Y \rangle + \langle AX | AY \rangle\end{aligned}$$

ce qui montre bien que $\langle p | q \rangle = \langle f | g \rangle$.

Pour exprimer A en fonction de U , supposons d'abord que si A est donné et $U = (iI + A)(iI - A)^{-1}$. Revenons aux deux relations $f = (iI - A)X$ et $Uf = p = (iI + A)X$. On peut les écrire sous la forme $f = iX - AX$, $Uf = iX + AX$, donc en les ajoutant et soustrayant l'une de l'autre $Uf + f = 2iX$ et $Uf - f = 2AX$, d'où $2X = -i(U + I)f$ et $2AX = -i(U - I)(U + I)^{-1} \cdot 2X$; ceci étant vrai quel que soit $X \in \mathcal{D}_A$, on voit que $U + I$ est un opérateur qui envoie bijectivement \mathcal{H} sur \mathcal{D}_A , et on en déduit $A = -i(U - I)(U + I)^{-1}$,

Opérateurs

ce qui montre aussi que $(U + I)^{-1}$ (qui envoie bijectivement \mathcal{D}_A sur \mathcal{H}) est discontinu si A l'est).

Cette formule d'inversion vient d'être obtenue en supposant A donné et $U = (iI + A)(iI - A)^{-1}$, ce qui veut dire que si U est obtenu comme $(iI + A)(iI - A)^{-1}$, A sera égal à $-i(U - I)(U + I)^{-1}$. Si maintenant U est un opérateur unitaire quelconque donné a priori, on voit que l'opérateur autoadjoint A tel que $U = (iI + A)(iI - A)^{-1}$ n'existe pas forcément : il faut pour cela que $U + I$ soit injectif et envoie bijectivement \mathcal{H} sur un sous-espace dense de \mathcal{H} (par exemple si $U = -I$ cela ne marche pas !). Si tel est le cas, soient X et Y deux éléments quelconques de \mathcal{H} . Puisque U est unitaire, on aura

$$\begin{aligned}\langle UX - X \mid UY + Y \rangle &= \langle UX \mid UY \rangle + \langle UX \mid Y \rangle - \langle X \mid UY \rangle - \langle X \mid Y \rangle \\ &= \langle UX \mid Y \rangle - \langle X \mid UY \rangle \\ \langle UX + X \mid UY - Y \rangle &= \langle UX \mid UY \rangle - \langle UX \mid Y \rangle + \langle X \mid UY \rangle - \langle X \mid Y \rangle \\ &= -\langle UX \mid Y \rangle + \langle X \mid UY \rangle\end{aligned}$$

Autrement dit $\langle UX - X \mid UY + Y \rangle = -\langle UX + X \mid UY - Y \rangle$, ou encore $\langle -i(UX - X) \mid UY + Y \rangle = \langle UX + X \mid i(UY - Y) \rangle$. Si on pose $f = UX + X$, $f \in \mathcal{D}$ d'après l'hypothèse, et on aura $Af = -i(UX - X)$, de sorte que pour tout $Y \in \mathcal{H}$,

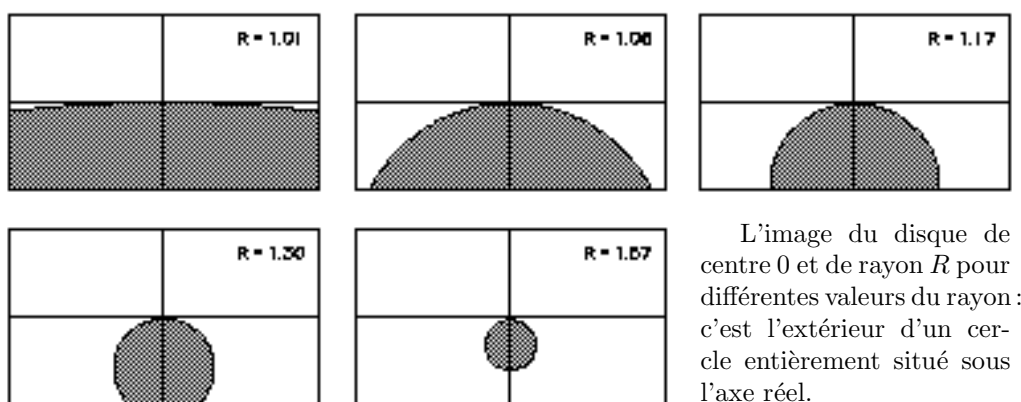
$$\langle Af \mid UY + Y \rangle = \langle f \mid i(UY - Y) \rangle$$

Cela prouve que la fonctionnelle $f \mapsto \langle Af \mid UY + Y \rangle$ est continue pour tout $Y \in \mathcal{H}$, et donc que la fonctionnelle $f \mapsto \langle Af \mid g \rangle$ est continue pour tout $g \in \mathcal{D}$; ce qui revient à dire que \mathcal{D} est inclus dans \mathcal{D}_{A^*} . Mais en échangeant les rôles de X et Y on arriverait à la conclusion que \mathcal{D}_{A^*} est inclus dans \mathcal{D} . CQFD

La transformation de Cayley est aussi une fonction d'opérateur : on peut dire que $U = f(A)$, avec $f(z) = (i + z)/(i - z)$. Mais, contrairement aux transformations envisagées avant, celle-ci s'applique à un opérateur non nécessairement continu. Pour la définir, on a procédé directement, sans passer par l'intermédiaire de séries entières qui doivent converger. Il serait possible de faire la même chose pour d'autres transformations. Par exemple, si on prend $f(z) = 1/(\lambda - z)$, ou $f(z) = (\lambda + z)/(\lambda - z)$, ou encore $f(z) = (\lambda + z)^2/(\lambda - z)^2$, etc. (avec λ non réel) on obtient des opérateurs $f(A)$ qui sont continus, bien que A ne le soit pas. On pourrait prendre aussi $f(z) = z/(\lambda - z)$, qui donnerait un opérateur $f(A)$ non continu, mais cependant bien défini.

On peut vérifier directement pour ces transformations particulières que les valeurs propres et les spectres sont transformés comme le dit le théorème ; de même, les opérateurs A et $f(A)$ commutent.

Si maintenant on prend un opérateur unitaire U et une fonction f analytique dans un disque de centre 0 et de rayon $R > 1$, la fonction $f(U)$ est bien définie par sa série entière. Soit aussi la fonction $g(z) = f(\frac{i+z}{i-z})$. Il est équivalent de dire que f est analytique dans un disque de rayon R , ou que g est analytique dans le domaine D_R des points $z = x + iy$ du plan tels que $2(R^2 + 1)y < (R^2 - 1)(x^2 + y^2 + 1)$, qui est l'extérieur d'un cercle (voir figure).



Si $R = 1$, ce cercle dégénère en la droite réelle et D_R est le demi-plan $\{y > 0\}$; si $R > 1$, ce cercle est situé sous la droite réelle, et donc le domaine D_R contient le demi-plan $\{y > 0\}$. La transformation homographique $z \mapsto w = (i + z)/(i - z)$ est une transformation conforme du domaine D_R sur le disque de centre 0 et de rayon R , privé du point $w = -1$. Ce point $w = -1$ correspond dans D_R au point à l'infini).

Si $w = (i + z)/(i - z)$, $g(z) = f(w)$. Par conséquent, si A est un opérateur autoadjoint non nécessairement continu, et si g est une fonction analytique dans D_R et sans singularité à l'infini (c'est-à-dire que $g(z)$ doit avoir une limite finie quand z tend dans D_R vers l'infini), on peut définir $g(A)$ sans difficulté en disant que c'est $f(U)$. Si g a une singularité à l'infini, on peut aussi poser $g(A) = f(U)$, si on parvient à résoudre la difficulté par un moyen direct (mais on ne peut se prononcer a priori). Par exemple, on peut définir e^{iA} en disant que c'est $f(U)$, où $f(w) = e^{(1-w)/(w+1)}$; mais cette fonction a un point singulier essentiel en $w = -1$, donc il faut prendre des précautions. On verra plus loin ce qui se passe sur des exemples.

6. Groupes unitaires.

Le problème mathématique de la Mécanique quantique non relativiste est de résoudre les équations de Schrödinger

$$i\hbar \frac{\partial \psi}{\partial t} = A\psi, \quad (6.1.)$$

de même que le problème mathématique de la Mécanique classique est de résoudre les équations de Newton $F = m\gamma$. L'opérateur A est appelé l'hamiltonien et représente l'énergie (ses valeurs propres sont les états d'énergie du système); c'est *par principe* un opérateur autoadjoint. Il opère sur l'espace $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$, et est en général discontinu.

Si on considérait 6.1 sur un espace de dimension finie, c'est-à-dire si A était un endomorphisme sur $\mathcal{E} = \mathbb{C}^n$, les fonctions ψ seraient des vecteurs à n composantes dépendant du temps et l'équation 6.1 serait un système de n équations différentielles linéaires du premier ordre. Si n était égal à 1, A serait simplement la multiplication par un nombre réel a et 6.1 serait une équation différentielle linéaire du premier ordre; ψ serait une fonction à valeurs complexes du temps t . La solution serait alors $\psi(t) = \psi(0) e^{-\frac{i}{\hbar}at}$. Pour $n > 1$, 6.1 serait un système d'équations, mais on pourrait toujours écrire sa solution sous la forme

$$\psi(t) = e^{-\frac{i}{\hbar}tA} \psi(0)$$

Pour calculer effectivement l'opérateur $e^{-\frac{i}{\hbar}tA}$, il faut diagonaliser la matrice A (on peut toujours diagonaliser une matrice autoadjointe), mais si on se contente d'une réponse de principe, l'exponentielle $U_t = e^{-\frac{i}{\hbar}tA}$ de la matrice exprime univoquement, quoique abstraitement, la solution du système d'équations.

On remarque que les matrices U_t forment un groupe: $U_{s+t} = U_s U_t$, $U_0 = I$, $U_t^{-1} = U_{-t}$. La solution d'un système d'équations différentielles linéaires du premier ordre est donc donnée par un groupe d'endomorphismes. Dans le cas du système 6.1, du fait que A est une matrice autoadjointe et du coefficient $i = \sqrt{-1}$ de $\frac{\partial \psi}{\partial t}$, les endomorphismes U_t sont unitaires (ce ne serait pas le cas pour le système $\frac{d\psi}{dt} = A\psi$).

Le but de la théorie des opérateurs sur les espaces de Hilbert est d'étendre ces considérations aux espaces de dimension infinie tels que les espaces $\mathbb{L}_2(\mathbb{R}^n)$.

Dans le cas de l'équation 6.1, l'opérateur est en général

$$A = -\frac{\hbar^2}{2m} \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right] + V(x, y, z) \quad (6.2.)$$

et il faut pouvoir préciser le sens de $U_t = e^{-\frac{i}{\hbar}tA}$. Si A était continu, il suffirait de le définir par la série $\sum \frac{1}{n!}(-\frac{i}{\hbar}tA)^n$, mais justement A n'est pas continu. Il n'existe pas de théorème du type "tout opérateur autoadjoint possède une exponentielle" (cf. les difficultés mentionnées à propos de la transformation de Cayley), mais on peut construire le groupe U_t dans la plupart des cas particuliers utiles. Pour fixer le langage, voici quelques définitions.

Définition. On appelle *groupe unitaire* une famille d'opérateurs U_t , paramétrée par $t \in \mathbb{R}$, telle que

- les U_t sont tous unitaires;
- $\forall s, t \in \mathbb{R}, U_t U_s = U_{s+t}$;
- pour $t = 0$, $U_t = I$;
- $\forall f \in \mathcal{H}$ la fonction $t \mapsto U_t f$ est continue sur \mathbb{R} , c'est-à-dire que pour tout $t \in \mathbb{R}$, $\|U_{t+h} f - U_t f\|$ tend vers zéro quand h tend vers zéro.

N.B. Pour la quatrième condition (continuité) il suffit de l'exiger en $t = 0$; la propriété de groupe la garantit alors sur tout \mathbb{R} : en effet, $\|U_{t+h} f - U_t f\| = \|(U_{t+h} - U_t) f\| = \|U_t (U_h f - f)\| = \|U_h f - f\|$.

Définition. Le *générateur infinitésimal* du groupe est l'opérateur (non nécessairement continu) G défini comme suit :

- le domaine \mathcal{D}_G est l'ensemble des éléments $f \in \mathcal{H}$ pour lesquels $\langle (U_t f - f)/it \mid g \rangle$ a pour tout $g \in \mathcal{H}$ une limite (dans \mathbb{C}) lorsque t tend vers zéro;
- Gf est l'élément de \mathcal{H} tel que, $\forall g \in \mathcal{H}$, $\langle Gf \mid g \rangle$ est cette limite.

On écrit souvent $U_t = e^{itG}$ pour un tel groupe.

Théorème 10. Le *générateur infinitésimal* d'un groupe unitaire est toujours autoadjoint.

Démonstration. La principale difficulté est de prouver que \mathcal{D}_G est dense. Voyons d'abord le reste.

La propriété de groupe implique que $U_t^{-1} = U_{-t}$, donc $\forall f, g \in \mathcal{H}$, $\langle U_t f \mid g \rangle = \langle f \mid U_{-t} g \rangle$, et donc aussi

$$\left\langle \frac{U_t f - f}{it} \mid g \right\rangle = \left\langle f \mid \frac{U_{-t} g - g}{-it} \right\rangle$$

Cette égalité, étant vraie $\forall t$, reste vraie quand on passe à la limite. Si $f \in \mathcal{D}_G$, la limite du premier membre est par définition $\langle Gf \mid g \rangle$, et le second membre tend alors aussi vers cette limite, ce qui d'après la définition signifie que g appartient à \mathcal{D}_G et que $\langle Gf \mid g \rangle = \langle f \mid Gg \rangle$; ce raisonnement restant

Opérateurs

identique à lui-même lorsqu'on échange f et g , ou lorsqu'on remplace t par $-t$, cela montre que $\mathcal{D}_G = \mathcal{D}_{G^*}$ et $G = G^*$.

Pour montrer que \mathcal{D}_G est dense, on va construire une suite d'opérateurs L_n qui converge vers I et tels que leurs images soient contenues dans \mathcal{D}_G . La densité de \mathcal{D}_G résulte alors simplement du raisonnement suivant : pour tout $f \in \mathcal{H}$, $f_n = L_n f \in \mathcal{D}_G$; d'autre part f_n tend vers f puisque $L_n f$ tend vers f . Ce procédé astucieux est emprunté à K. Yosida ⁽⁸⁾ avec quelques modifications.

L'opérateur L_n est défini ainsi : pour tous $f, g \in \mathcal{H}$, on pose

$$\langle L_n f \mid g \rangle = n \int_0^\infty e^{-ns} \langle U_s f \mid g \rangle ds \quad (6.3.)$$

La convergence de l'intégrale est garantie par l'inégalité de Schwarz $|\langle U_s f \mid g \rangle| \leq \|f\| \|g\|$, et l'opérateur L_n est bien défini par la donnée de tous les produits scalaires $\langle L_n f \mid g \rangle$. L'idée implicite est la suivante : il faut imaginer que $U_t = e^{itG}$, et l'opérateur L_n est la fonction d'opérateur $\Phi(G)$ avec $\Phi(x) = n \int_0^\infty e^{-ns} e^{isx} dx = n/(n - ix)$. Cette fonction tend bien vers 1 quand n tend vers l'infini, donc "on peut s'attendre" à ce que L_n tende vers I . Bien entendu il faut démontrer ce qui n'est ainsi que deviné. Mais prouvons d'abord que pour tout $h \in \mathcal{H}$, $f = L_n h \in \mathcal{D}_G$. Il faut montrer que pour tout $g \in \mathcal{H}$, $\langle \frac{1}{i\varepsilon} (U_\varepsilon f - f) \mid g \rangle$ tend vers une limite quand ε tend vers zéro. Or

$$\begin{aligned} \langle \frac{1}{i\varepsilon} (U_\varepsilon f - f) \mid g \rangle &= \frac{n}{i\varepsilon} \int_0^\infty e^{-ns} \langle U_{\varepsilon+s} h - U_s h \mid g \rangle ds \\ &= \frac{e^{n\varepsilon} - 1}{i\varepsilon} n \int_\varepsilon^\infty e^{-ns} \langle U_s h \mid g \rangle ds - \frac{n}{i\varepsilon} \int_0^\varepsilon e^{-ns} \langle U_s h \mid g \rangle ds \\ &= \frac{e^{n\varepsilon} - 1}{i\varepsilon} \left\{ \langle L_n h \mid g \rangle - n \int_0^\varepsilon e^{-ns} \langle U_s h \mid g \rangle ds \right\} - \frac{n}{i\varepsilon} \int_0^\varepsilon e^{-ns} \langle U_s h \mid g \rangle ds \end{aligned}$$

La continuité de la fonction $t \mapsto \langle U_t h \mid g \rangle$ (qui est exigée par la définition) implique que l'intégrale entre accolades tend vers zéro quand ε tend vers zéro, et le terme à l'extérieur des accolades tend vers $-n \langle U_0 h \mid g \rangle$; enfin, le facteur $(e^{n\varepsilon} - 1)/i\varepsilon$ tend vers $-in$ (quand ε tend vers zéro). Donc la limite du tout est $-in [\langle L_n h - h \mid g \rangle]$. On a ainsi prouvé que cette limite existe, donc que $L_n h$ est dans \mathcal{D}_G .

Il reste à vérifier que, quand n tend vers l'infini, $\|L_n f - f\|$ tend vers zéro. L'opérateur L_n a été défini par 6.3, qui doit se vérifier pour tous f et

⁽⁸⁾ Kôsaku Yosida *Functional Analysis* Springer-Verlag, Berlin, 1968, pp. 237 – 238.

g ; d'autre part, $\int_0^\infty n e^{-ns} ds = 1$, donc on a aussi

$$\langle f | g \rangle = \int_0^\infty n e^{-ns} \langle f | g \rangle ds$$

En soustrayant cela de 6.3 on obtient

$$\langle L_n f - f | g \rangle = \int_0^\infty n e^{-ns} \langle U_s f - f | g \rangle ds$$

et cela est vrai pour tous f et g dans \mathcal{H} ; prenons donc $g = L_n f - f$, ce qui donne

$$\|L_n f - f\|^2 = \int_0^\infty n e^{-ns} \langle U_s f - f | L_n f - f \rangle ds$$

Pour majorer cela, on utilise l'inégalité de Schwarz $|\langle U_s f - f | L_n f - f \rangle| \leq \|U_s f - f\| \cdot \|L_n f - f\|$, puis l'inégalité de la moyenne :

$$\left| \int_0^\infty n e^{-ns} \langle U_s f - f | L_n f - f \rangle ds \right| \leq \int_0^\infty n e^{-ns} \|U_s f - f\| ds \times \|L_n f - f\|$$

Par conséquent

$$\|L_n f - f\| \leq \int_0^\infty n e^{-ns} \|U_s f - f\| ds = \int_0^\infty n e^{-ns} \varphi(s) ds$$

où $\varphi(s) = \|U_s f - f\|$. Or, d'après la définition des groupes unitaires, la fonction $\varphi(s)$ est continue, majorée uniformément par $2\|f\|$, et s'annule pour $s = 0$. La conclusion résulte donc du fait que pour une telle fonction, l'intégrale $\int_0^\infty n e^{-ns} \varphi(s) ds$ tend vers zéro quand n tend vers l'infini, ce qui est laissé en exercice (faire le changement de variable $t = ns$). CQFD

Remarques. Ce théorème ne se prononce pas sur une réciproque du type "tout opérateur autoadjoint est le générateur infinitésimal d'un groupe unitaire", car ce serait faux. Mais c'est vrai pour les opérateurs qui ont déjà été présentés.

— Si G est l'opérateur $i \frac{d}{dx}$ (ou $i\hbar \frac{d}{dx}$), le groupe est celui des translations : $U_t f(x) = f(x+t)$; en effet on a bien $\lim \langle (U_t f - f)/it | g \rangle = \lim \int \overline{g(x)} [f(x+t) - f(x)]/it dx = \int \overline{g(x)} f'(x) dx$. La solution de l'équation différentielle $\frac{d\psi}{dt} = \frac{d\psi}{dx}$ est bien $\psi(t, x) = f(x+t)$.

— Si G est l'opérateur de multiplication par une fonction réelle $V(x)$, le groupe unitaire est celui des multiplications par $e^{itV(x)}$.

— Si G est l'opérateur $-\frac{d^2}{dx^2} + x^2$, le groupe est celui des transformations U_t qui, si on représente les éléments de \mathcal{H} en séries de fonctions d'Hermite, s'expriment par $U_t : f = \sum a_n \phi_n \mapsto \sum a_n e^{it(2n+1)} \phi_n$. On résoud donc ici l'équation de Schrödinger $i \frac{d\psi}{dt} = -\frac{d^2\psi}{dx^2} + x^2 \psi$ exactement de la même

Opérateurs

façon que Fourier a résolu l'équation de la chaleur (cf. chap. VI, section 1), sauf qu'on utilise des séries de fonctions d'Hermite à la place des séries de fonctions trigonométriques.

— Cela se généralise en principe aux opérateurs $-\frac{d^2}{dx^2} + V(x)$, avec une fonction potentiel V vérifiant certaines conditions (on peut pas garantir l'existence d'un groupe unitaire pour n'importe quelle fonction V).

On peut donc résumer tout cela comme suit.

Le groupe unitaire associé à un opérateur autoadjoint est la solution d'une équation de Schrödinger :

$$-i \frac{d\psi}{dt} = A\psi$$

qui se résoud formellement par $\psi(t) = e^{itA} \psi(0)$. Pour une condition initiale $\psi(0) = f$, la solution est en effet $\psi(t) = U_t f$. Par exemple l'équation :

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \left[\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right] + V(x, y, z) \psi$$

est de ce type ; on cherche une solution $\psi(t, x, y, z)$ qui pour chaque t fixé est une fonction de carré intégrable. Le groupe unitaire U_t qui la résoud donne la fonction à l'instant t , $\psi_t(x, y, z) = \psi(t, x, y, z)$, à partir de la fonction initiale $f(x, y, z) = \psi(0, x, y, z)$. Le fait que le groupe soit unitaire se traduit par la conservation de la norme : $\|\psi_t\|_2 = \|f\|_2$. Cette propriété de l'équation de Schrödinger est essentielle, car elle correspond à la *conservation de la probabilité* ; le carré de la fonction d'onde, $|\psi_t(x, y, z)|^2$, représente la densité de présence de la particule dans l'espace, dont l'intégrale sur tout l'espace doit être égale à 1. Ainsi, si au départ $\int |f(x, y, z)|^2 dx dy dz = \|f\|_2^2 = 1$, on aura aussi à tout instant ultérieur $\int |\psi(t, x, y, z)|^2 dx dy dz = \|f\|_2^2 = 1$. On voit que, du point de vue mathématique, cette conservation de la probabilité se traduit par le fait que l'opérateur hamiltonien est auto-adjoint.

7. L'espace $\mathbb{L}_2(S)$ et les harmoniques sphériques.

L'espace de Hilbert $\mathbb{L}_2(S)$ est celui des fonctions de carré intégrable sur la sphère $S = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$. Si on représente les points de l'espace en coordonnées sphériques

$$\begin{aligned} x &= r \sin \theta \cos \varphi \\ y &= r \sin \theta \sin \varphi \\ z &= r \cos \theta \end{aligned}$$

ceux de la sphère S correspondent à $r = 1$, et sont donc représentés par les deux coordonnées sphériques angulaires θ et φ . Un vecteur de $\mathbb{L}_2(S)$

est donc une fonction de θ et de φ . Sur la sphère, l'élément de surface est $d\sigma = \frac{1}{4\pi} \sin \theta d\theta d\varphi$, donc la norme $\|f\|_{\text{sph}}$ d'une fonction $f \in \mathbb{L}_2(S)$ sera donnée par

$$\|f\|_{\text{sph}}^2 = \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi |f(\theta, \varphi)|^2 \sin \theta d\theta d\varphi$$

Le produit scalaire est

$$\langle f | g \rangle_{\text{sph}} = \int_0^{2\pi} \int_0^\pi f(\theta, \varphi) \overline{g(\theta, \varphi)} \sin \theta d\theta d\varphi$$

Cette section est essentiellement consacrée à la construction d'une base orthogonale de l'espace $\mathbb{L}_2(S)$.

Définition. On appelle harmoniques sphériques les restrictions à la sphère $S = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$ des polynômes harmoniques homogènes des trois variables x, y, z . Une harmonique sphérique de degré n sera la restriction à la sphère d'un polynôme harmonique homogène de degré n .

Rappelons qu'un polynôme homogène est un polynôme dont tous les termes ont le même degré : par exemple $x^3 + 5xy^2 - x^2z + 2xyz$ ou $x^6 + y^6 + x^4z^2 - 3xy^3z^2$ sont homogènes, tandis que $x + yz$ ou $x^3 - z^2 + 2xy^4$ ne le sont pas. Par ailleurs, une fonction est dite harmonique si son laplacien est nul.

Les harmoniques sphériques forment un espace vectoriel, qui est un sous-espace dense de l'espace de Hilbert $\mathbb{L}_2(S)$

L'expression la plus générale d'un polynôme homogène de degré n des trois variables x, y, z est

$$P(x, y, z) = \sum_{i+j+k=n} \sum_{j,k} a_{ijk} x^i y^j z^k$$

La dimension de l'espace vectoriel \mathcal{W}_n des polynômes homogènes de degré n est donc le nombre de coefficients de cette expression, qui est aussi le nombre de multi-indices (i, j, k) tels que $i + j + k = n$, soit $\frac{1}{6}(n+1)(n+2)(n+3)$. Cette dimension est évidemment plus petite que la dimension $\frac{1}{6}(n+1)(n+2)(n+3)$ de tous les polynômes de degré $\leq n$.

La dimension des polynômes homogènes *harmoniques* de degré n sera encore plus petite, puisque la condition $\Delta P = 0$ implique des relations linéaires entre les coefficients a_{ijk} . Pour la connaître, il faut voir de combien de coefficients primitifs indépendants dépendent effectivement les a_{ijk} , ce qu'on peut faire comme suit. D'après la formule de Taylor, on a

$$a_{ijk} = \frac{\partial^n P}{\partial x^i \partial y^j \partial z^k} \quad (7.1.)$$

Opérateurs

On notera qu'il est inutile de préciser qu'il s'agit des dérivées en $(x, y, z) = (0, 0, 0)$, puisque ces dérivées sont constantes. Or, si le polynôme P est harmonique, on a

$$\frac{\partial^2 P}{\partial z^2} = -\frac{\partial^2 P}{\partial x^2} - \frac{\partial^2 P}{\partial y^2} \quad (7.2)$$

ce qui signifie que toutes les dérivées partielles d'ordre pair par rapport à z pourront s'exprimer linéairement en fonction des seules dérivées par rapport à x et y ; quant à celles d'ordre impair par rapport à z , on pourra les ramener à des dérivées par rapport à x , y et z , celle restant par rapport à z étant d'ordre 1. Si on fait cela dans 7.1, on aura exprimé tous les a_{ijk} uniquement à partir des $n+1$ dérivées $\partial^n P / \partial x^i \partial y^j$ (nombre de couples i, j tels que $i+j=n$), plus les n dérivées $\partial^n P / \partial x^i \partial y^j \partial z$ (nombre de couples i, j tels que $i+j+1=n$), soit $2n+1$.

En conclusion, la dimension de l'espace vectoriel des polynômes homogènes harmoniques de degré n est $2n+1$.

Cela implique que pour en trouver une base, il suffit d'exhiber $2n+1$ polynômes homogènes harmoniques linéairement indépendants.

Pour cela, on considère les intégrales de la forme

$$f(x, y, z) = \int_{-\pi}^{+\pi} (ix \cos t + iy \sin t + z)^n u(t) dt \quad (7.3.)$$

où $u(t)$ est une fonction quelconque, disons continue, sur $[-\pi, +\pi]$. Ces fonctions sont visiblement des polynômes homogènes de degré n . Si on dérive, on obtient

$$\Delta f(x, y, z) = \int_{-\pi}^{+\pi} n(n-1) [1 - \cos^2 t - \sin^2 t] (ix \cos t + iy \sin t + z)^{n-2} u(t) dt$$

ce qui est nul en vertu de l'identité $\cos^2 t + \sin^2 t = 1$; la dérivation sous le signe d'intégration est justifiée par les théorèmes généraux, puisque les dérivées d'ordre supérieur (par exemple quatre) sont toujours des fonctions continues et bornées de t sur $[-\pi, +\pi]$.

Ainsi les fonctions 7.3 sont toutes des polynômes homogènes harmoniques de degré n , et les fonctions de θ, φ qu'on en déduit en prenant $r=1$ sont alors des harmoniques sphériques de degré n :

$$\begin{aligned} f(\theta, \varphi) &= \int_{-\pi}^{+\pi} [i \sin \theta \cos \varphi \cos t + i \sin \theta \sin \varphi \sin t + \cos \theta]^n u(t) dt \\ &= \int_{-\pi}^{+\pi} [i \sin \theta \cos(t - \varphi) + \cos \theta]^n u(t) dt \end{aligned} \quad (7.4.)$$

Les fonctions de cette forme ne représentent pas toutes les harmoniques sphériques de degré n , mais suffisent pour en construire une base. En effet, les fonctions

$$Y_{n,\ell}(\theta, \varphi) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} [i \sin \theta \cos(t - \varphi) + \cos \theta]^n e^{i\ell t} dt \quad (7.5.)$$

pour $\ell = -n, -n+1, -n+2, \dots, n-1, n$, qui correspondent aux cas où $u(t) = \frac{1}{2\pi} e^{i\ell t}$ sont linéairement indépendantes et au nombre de $2n+1$.

Exercice. Prouver que pour tout n fixé, les $2n+1$ fonctions $Y_{n,\ell}$, pour $\ell = -n, -n+1, -n+2, \dots, n-1, n$, sont linéairement indépendantes (Effectuer dans l'intégrale de 7.5 le changement de variable $s = t - \varphi$ et tenir compte du fait que tout est périodique). En déduire qu'elles forment une base de l'espace des harmoniques sphériques de degré n .

Appelons \mathcal{Y}_n l'espace vectoriel (de dimension $2n+1$) des harmoniques sphériques de degré n . La somme directe des \mathcal{Y}_n pour n allant de zéro à l'infini sera un espace de dimension infinie, qu'on désignera par \mathcal{Y} .

On peut démontrer un théorème du même type que les théorèmes de Weierstrass (cf. chap. **X**, section **5**):

Théorème 11. *Toute fonction continue sur la sphère et à valeurs réelles ou complexes, est la limite uniforme d'une suite d'harmoniques sphériques.*

Nous ne donnons pas de démonstration détaillée, mais il n'y aurait rien de vraiment nouveau par rapport aux théorèmes de la section **X.5**. Le théorème de Weierstrass concernant les polynômes algébriques (théorème 2 du chapitre **X**), s'appliquait à l'intervalle $[0, 1]$, ou (moyennant un changement de variable linéaire) à n'importe quel intervalle borné, par exemple $[-1, +1]$. Mais on peut étendre la démonstration du théorème aux polynômes à deux ou trois variables⁽⁹⁾ (cela ne fait essentiellement que compliquer l'écriture), et prouver ainsi que les fonctions continues sur le cube $[-1, +1] \times [-1, +1] \times [-1, +1]$ sont des limites uniformes de polynômes. Or la sphère est contenue dans ce cube; en outre, il se trouve que la restriction d'un polynôme quelconque à la sphère est automatiquement aussi la restriction d'un polynôme homogène harmonique (ce point, au demeurant élémentaire, serait en fait la seule nouveauté).

On en déduit alors aisément (comme cela avait été fait pour le théorème 3 du chapitre **X**) que les harmoniques sphériques sont denses dans l'espace $\mathbb{L}_2(S)$. On peut énoncer :

⁽⁹⁾ Weierstrass a lui-même donné cette extension du théorème à la fin de son article publié en 1885, déjà cité (cf. note 3 du chapitre **X**).

Opérateurs

Théorème 12. *La famille infinie des fonctions $Y_{n,\ell}$ pour $n = 0, 1, 2, \dots$ et $-n \leq \ell \leq +n$, forme une base orthogonale de l'espace de Hilbert $\mathbb{L}_2(S)$.*

Démonstration. Pour montrer qu'une famille de fonctions forme une base hilbertienne, il faut vérifier qu'elle est orthogonale et que le sous-espace des combinaisons linéaires finies de ces fonctions est dense. La seconde partie de ce programme est déjà réglée par l'extension du théorème de Weierstrass mentionnée ci-dessus. Il ne reste donc plus qu'à calculer les intégrales

$$\int_0^{2\pi} \int_0^\pi \sin \theta Y_{n\ell}(\theta, \varphi) Y_{n'\ell'}(\theta, \varphi) d\theta d\varphi$$

C'est l'objet des questions 3 et 6 de l'exercice suivant.

Exercice. 1. On rappelle la définition des polynômes de Legendre (chap. X, section 4A) :

$$P_n(x) = \frac{1}{n! 2^n} \frac{d^n([x^2 - 1]^n)}{dx^n}$$

Montrer que l'on peut écrire aussi

$$P_n(x) = \frac{1}{2i\pi} \int_\gamma \frac{(z-1)^n (z+1)^n}{2^n (z-x)^{n+1}} dz$$

où γ est un lacet simple du plan complexe qui entoure le point $z = x$.

2. On suppose ici que $-1 < x < +1$. Prendre pour γ un cercle de centre x et de rayon $R = \sqrt{1-x^2}$. Paramétrer ce lacet avec un paramètre angulaire t qui varie de $-\pi$ à $+\pi$, et montrer que l'on a

$$P_n(x) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} [x + \sqrt{1-x^2} \cos t]^n dt$$

En déduire que $Y_{n,0}(\theta, \varphi) = P_n(\cos \theta)$.

3. Montrer que pour tout $n \geq 0$ on a

$$\int_0^{2\pi} \int_0^\pi Y_{n\ell}(\theta, \varphi) Y_{nk}(\theta, \varphi) \sin \theta d\theta d\varphi = 0$$

si $k \neq \ell$. Conclure que pour tout n fixé, les $2n+1$ fonctions $Y_{n\ell}$ ($-n \leq \ell \leq +n$) forment une base orthogonale de \mathcal{Y}_n . Remarquer que l'orthogonalité provient uniquement de l'intégration par rapport à φ et que la coordonnée θ n'y joue aucun rôle.

4 Montrer que les fonctions $f(x, y, z) = r^n Y(\theta, \varphi)$ (où r, θ, φ sont les coordonnées sphériques de x, y, z) sont des polynômes homogènes harmoniques.

5. On rappelle l'expression du laplacien en coordonnées sphériques :

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \Delta_{\text{sph}}$$

où Δ_{sph} est la partie sphérique du laplacien, qui n'opère que sur les coordonnées θ et φ :

$$\Delta_{\text{sph}} = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2}$$

Montrer que $\Delta_{\text{sph}} Y_{n\ell} = -n(n+1) Y_{n\ell}$.

6. Vérifier à l'aide d'intégrations par parties que $\langle \Delta_{\text{sph}} Y_{n\ell} | Y_{n\ell} \rangle_{\text{sph}} = \langle Y_{n\ell} | \Delta_{\text{sph}} Y_{n\ell} \rangle_{\text{sph}}$. En déduire (en utilisant le résultat obtenu en 5) que si $n \neq m$, $Y_{n\ell}$ est orthogonale à Y_{mk} (indépendamment de k et ℓ). Autrement dit, les sous-espaces \mathcal{Y}_n sont orthogonaux dans $\mathbb{L}_2(S)$.

La question 6 de cet exercice consiste à prouver que la restriction à \mathcal{Y}_n de l'opérateur Δ_{sph} est égale à $-n(n+1)I$; cela montre que les nombres $\lambda_n = -n(n+1)$ sont des valeurs propres de Δ_{sph} et les $Y_{n\ell}$ des vecteurs propres. Comme par ailleurs les $Y_{n\ell}$ forment (pour l'ensemble des indices $n = 0, 1, 2, \dots$ et $-n \leq \ell \leq +n$) une base de $\mathbb{L}_2(S)$, ils fournissent forcément la totalité des vecteurs propres possibles (s'il existait une valeur propre autre que les $\lambda_n = -n(n+1)$, le sous-espace propre correspondant serait orthogonal à tous les $Y_{n\ell}$, ce qui ne peut pas se produire puisque justement les $Y_{n\ell}$ forment une base. Par le même argument que dans la démonstration du théorème 2, on voit aussi que le spectre est réduit à ces valeurs propres. On peut donc énoncer :

Théorème 13. *L'opérateur Δ_{sph} sur l'espace $\mathbb{L}_2(S)$ a un spectre formé exclusivement des valeurs propres $\lambda_n = -n(n+1)$, de multiplicité $2n+1$; les vecteurs propres associés sont les fonctions $Y_{n\ell}$ données en 7.5.*

Commentaire. L'opérateur Δ_{sph} est évidemment discontinu ; son domaine est celui des fonctions $f = \sum \sum a_{n\ell} Y_{n\ell}$ telles que la série $\sum \sum (2n+1) |a_{n\ell}|^2$ converge. On peut aussi caractériser ce domaine comme suit : les dérivations partielles de l'opérateur étant entendues au sens des distributions, le domaine est l'ensemble des fonctions $f \in \mathbb{L}_2(S)$ telles que $\Delta_{\text{sph}} f \in \mathbb{L}_2(S)$.

8. Théorie de l'atome d'hydrogène.

En dehors de l'oscillateur quantique, un autre exemple traité par Schrödinger dans les mémoires de 1926⁽¹⁰⁾ est celui d'un électron dans un

⁽¹⁰⁾ Pages 3 à 12 dans la traduction française de A. Proca rééditée récemment par Jacques Gabay.

Opérateurs

champ coulombien. Il s'agit du problème avec *un seul* électron dans le champ coulombien, car s'il y en a deux ou plus on ne sait pas résoudre l'équation. Le système physique qui correspond à ce problème est celui d'un atome avec un seul électron, donc l'atome d'hydrogène, ou un atome d'hélium ionisé positivement. Le champ coulombien est alors celui du noyau, et son potentiel est $V(x, y, z) = -\epsilon^2/r$ (ϵ étant la charge élémentaire, notée en gothique pour ne pas confondre avec $e = 2.71828\dots$, et $r = \sqrt{x^2 + y^2 + z^2}$)

Jusqu'ici, nous avons presque toujours considéré des problèmes à une seule variable; ici il y en a trois, et l'espace de Hilbert est $\mathcal{H} = \mathbb{L}_2(\mathbb{R}^3)$. Mais cela ne change rien au principe.

Dans le problème de l'oscillateur quantique (section 3), la base de vecteurs propres, les fonctions d'Hermite, a été fournie gratuitement. Ici, on ne pourra pas fournir une *base* de vecteurs propres car il n'en existe pas, mais on peut fournir un système orthogonal *non complet* de fonctions propres. Pour l'oscillateur quantique, on a pu démontrer que le spectre ne contenait que les valeurs propres $\lambda_n = \hbar\omega(n + \frac{1}{2})$ parce que les vecteurs propres associés formaient une base complète. Ici, on ne pourra pas arriver à cette conclusion car il n'y a pas de base complète de vecteurs propres, et il se trouve que le spectre n'est pas formé que de valeurs propres: il contient en outre la demi-droite $[0, +\infty[$, qui constitue le spectre essentiel, et il n'y a aucune valeur propre > 0 .

Le système orthogonal non complet de fonctions propres s'obtient comme suit. On introduit d'abord la famille $f_{nk}(r)$ des *fonctions de Laguerre*, définies pour $t > 0$ par

$$f_{nk}(t) = Q_{n-k-1}^{2k+1}(t) t^k e^{-t/2}$$

qui sont représentées sur la figure 5 du chapitre X. Les Q_n^j sont les polynômes de Laguerre déjà introduits au chapitre X, section 4C. Pour tout j , Q_n^j est de degré n .

Le système cherché est alors celui des fonctions

$$\psi_{nk\ell}(x, y, z) = f_{nk}\left(\frac{2m\epsilon^2}{\hbar^2 n} r\right) Y_{k\ell}(\theta, \varphi) \quad (8.1.)$$

pour les indices entiers $n = 1, 2, 3, \dots$, $0 \leq k \leq n - 1$, et $-n \leq \ell \leq +n$, qui seront les vecteurs propres associés aux valeurs propres

$$E_n = -\frac{m\epsilon^4}{2\hbar^2 n^2} \quad (8.2.)$$

L'importance historique de ce résultat provient de ce qu'il correspond à la loi de Bohr, qui au lieu d'être devinée à partir des observations, est

déduite mathématiquement du *principe de correspondance*, c'est-à-dire de l'hypothèse que la Mécanique “classique” serait l'optique géométrique d'un phénomène ondulatoire.

Voici ce qu'écrivait Schrödinger dans une présentation d'ensemble de ses mémoires :

L'analogie entre la mécanique et l'optique signalée par HAMILTON ne concerne que l'optique *géométrique*; en effet, d'après cette analogie, à toute *trajectoire* du point représentatif dans l'espace de configuration correspond un *rayon de lumière* et cette dernière notion ne peut être définie sans ambiguïté qu'en optique géométrique. *L'analyse de cette image optique*, du point de vue ondulatoire, conduit à l'abandon de la notion de *trajectoire du système*, dès que les dimensions sont *petites* par rapport à la longueur d'onde. Cette notion, et avec elle toute la mécanique classique, subsistent comme approximations, valables seulement dans le cas où les dimensions de la trajectoire sont grandes par rapport à la longueur d'onde. (...) Par analogie avec ce qui se passe en optique, les équations fondamentales doivent être remplacées par une seule équation d'ondes dans l'espace de configuration. (...) Cette équation contient un “paramètre de valeurs propres” E , qui correspond à l'énergie mécanique. (...) En général, cette équation des ondes ne possède de solutions finies, continues, à détermination unique et dont les dérivées satisfassent aux mêmes conditions⁽¹¹⁾, que pour certaines valeurs particulières du paramètre E , qu'on appelle les *valeurs propres*. Les valeurs propres sont, ou bien identiques aux “niveaux d'énergie” de la théorie quantique admise jusqu'à présent, ou bien elles s'en écartent d'une manière qui est en parfait accord avec les résultats de l'expérience.

(Relire aussi les passages cités au début de la section 3.)

On peut donc conclure que cet accord, constaté dans de nombreux cas dès 1926 (et jamais démenti depuis) est une très forte confirmation du principe de correspondance.

On va maintenant vérifier que les fonctions $\psi_{nkl}(x, y, z)$ de 8.1 sont bien les fonctions propres de l'opérateur

$$H = -\frac{\hbar^2}{2m}\Delta - \frac{e^2}{r}$$

Prouvons d'abord que les polynômes de Laguerre $Q_m^j(t)$ qui interviennent dans l'expression de la fonction ψ_{nkl} satisfont la relation

$$t Q''(t) + (j + 1 - t) Q'(t) + m Q(t) = 0 \quad (8.3.)$$

Rappelons que les polynômes $Q_m^j(t)$ sont définis par la dérivée d'ordre m de la fonction $t^{m+j}e^{-t}$ (cf. chap. X, section 4C) :

$$\frac{d^m(t^{m+j}e^{-t})}{dt^m} = Q_m^j(t) t^j e^{-t}$$

⁽¹¹⁾ La vraie nature mathématique de cette condition a été comprise plus tard par J. von Neumann : la solution doit être dans l'espace \mathbb{L}_2 .

Opérateurs

On déduit immédiatement de cette définition que

$$\begin{aligned}
 Q_{m+1}^j(t) t^j e^{-t} &= \frac{d^{m+1}(t^{m+1+j} e^{-t})}{dt^{m+1}} = \\
 &= \frac{d}{dt} \frac{d^m(t^{m+j+1} e^{-t})}{dt^m} = \frac{d}{dt} [Q_m^{j+1}(t) t^{j+1} e^{-t}] = \\
 &= [t \frac{d}{dt} Q_m^{j+1}(t) + (j+1-t) Q_m^{j+1}(t)] t^j e^{-t}
 \end{aligned}$$

d'où

$$Q_{m+1}^j(t) = t \frac{d}{dt} Q_m^{j+1}(t) + (j+1-t) Q_m^{j+1}(t) \quad (8.4.)$$

En dérivant ce résultat il vient

$$\frac{d}{dt} Q_{m+1}^j(t) = t \frac{d^2}{dt^2} Q_m^{j+1}(t) + (j+1-t) \frac{d}{dt} Q_m^{j+1}(t) - Q_m^{j+1}(t) \quad (8.5.)$$

Si on fait une récurrence sur m , on suppose comme hypothèse de récurrence que 8.3 est vraie $\forall j \in \mathbb{R}$ (j n'a aucune raison d'être entier); combinée avec 8.5 cette hypothèse donne

$$\frac{d}{dt} Q_{m+1}^j(t) = -(m+1) Q_m^{j+1}(t)$$

d'où on déduit

$$\begin{aligned}
 t \frac{d^2}{dt^2} Q_{m+1}^j(t) + (j+1-t) \frac{d}{dt} Q_{m+1}^j(t) &= \\
 &= -(m+1) [t \frac{d^2}{dt^2} Q_m^{j+1}(t) + (j+1-t) \frac{d}{dt} Q_m^{j+1}(t)]
 \end{aligned}$$

Revenant à 8.4, on voit que ceci est égal à $-(m+1) Q_{m+1}^j(t)$, ce qui établit la récurrence.

Pour prouver complètement la récurrence, il faut encore s'assurer que 8.3 est vraie pour $m=0$, ce qui est évident puisque les polynômes Q_m^j sont alors constants.

Revenons maintenant à l'équation de Schrödinger. En écrivant le laplacien en coordonnées sphériques

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \Delta_{\text{sph}}$$

on obtient

$$\Delta \psi_{nkl} = \frac{\partial^2 \psi_{nkl}}{\partial r^2} + \frac{2}{r} \frac{\partial \psi_{nkl}}{\partial r} + \frac{1}{r^2} \Delta_{\text{sph}} \psi_{nkl}$$

La partie radiale n'opère que sur la fonction $f_{nk}(\frac{2m\epsilon^2}{\hbar^2 n} r)$, et la partie sphérique Δ_{sph} n'opère que sur les fonctions $Y_{k\ell}(\theta, \varphi)$. La première donne donc

$$[\frac{2m\epsilon^2}{\hbar^2 n}]^2 \cdot \left[f_{nk}''(t) + \frac{2}{t} f_{nk}'(t) \right] Y_{k\ell}(\theta, \varphi)$$

où on a posé $t = \frac{2m\epsilon^2}{\hbar^2 n} r$, et la seconde, d'après le théorème 13, donne

$$-\frac{1}{r^2} k(k+1) f_{nk}(t) Y_{k\ell}(\theta, \varphi) = -[\frac{2m\epsilon^2}{\hbar^2 n}]^2 \frac{1}{t^2} k(k+1) f_{nk}(t) Y_{k\ell}(\theta, \varphi)$$

En regroupant tout, on obtient

$$\Delta\psi_{nk\ell}(x, y, z) = [\frac{2m\epsilon^2}{\hbar^2 n}]^2 \cdot \left[f_{nk}''(t) + \frac{2}{t} f_{nk}'(t) - \frac{1}{t^2} k(k+1) f_{nk}(t) \right] Y_{k\ell}(\theta, \varphi)$$

Par conséquent $H\psi_{nk\ell}$ devient, en convertissant tout à la variable t :

$$H\psi_{nk\ell} = -\frac{2m\epsilon^4}{\hbar^2 n^2} \left[f_{nk}''(t) + \frac{2}{t} f_{nk}'(t) - \frac{1}{t^2} k(k+1) f_{nk}(t) - \frac{n}{t} f_{nk}(t) \right] Y_{k\ell}(\theta, \varphi)$$

Or d'après 8.3, les polynômes $Q_{n-k-1}^{2k+1}(t)$ qui interviennent dans l'expression des fonctions $f_{nk}(t)$ vérifient la relation

$$tQ''(t) + (2k+2-t)Q'(t) + (n-k-1)Q(t) = 0$$

Des calculs élémentaires conduisent alors à la relation correspondante pour les fonctions $f_{nk}(t) = Q_{n-k-1}^{2k+1}(t) t^k e^{-t/2}$:

$$f'' + \frac{2}{t} f' + \left[-\frac{k(k+1)}{t^2} + \frac{n}{t} \right] f = -\frac{1}{4} f \quad (8.6.)$$

Ainsi on voit que

$$H\psi_{nk\ell} = \frac{m\epsilon^4}{2\hbar^2 n^2} \psi_{nk\ell}$$

ce qui montre que les $\psi_{nk\ell}$ sont des vecteurs propres, associés aux valeurs propres $E_n = m\epsilon^4/2\hbar^2 n^2$.

Cela ne prouve cependant en rien qu'il n'existe pas d'autres fonctions propres, car les $\psi_{nk\ell}$ ne forment pas une base.

On voit que la valeur propre E_n ne dépend que de n ; par conséquent le nombre de $\psi_{nk\ell}$ associées à la valeur propre E_n est le nombre des couples d'indices k, ℓ . Or on a vu que ℓ prend les $2k+1$ valeurs comprises entre $-k$ et $+k$, tandis que k prend les valeurs de 0 à $n-1$. Le nombre de couples k, ℓ est donc la somme des nombres $2k+1$ pour k allant de 0 à $n-1$, soit n^2 . Comme les $\psi_{nk\ell}$ sont linéairement indépendantes, la multiplicité de la valeur propre E_n est donc n^2 .

Opérateurs

Pour prouver qu'il n'existe pas d'autres fonctions propres que les ψ_{nkl} , on remarque que s'il en existait, elles se décomposeraient aussi sous la forme $f(t)Y(\theta, \varphi)$, puisque l'opérateur H n'agit en fait que sur la variable radiale. La fonction $f(t)$ vérifierait donc toujours l'équation 8.6, mais cette fois pour des valeurs non nécessairement entières ≥ 1 de n . La forme de l'équation (du type de Fuchs) garantit que les solutions sont des séries de Laurent en t . En cherchant ces séries, on obtient une relation de récurrence pour leurs coefficients, et on se rend compte que le seul cas où elles peuvent donner des polynômes est celui où n est entier ≥ 1 ; dans les autres cas, la récurrence entre les coefficients montre que leur signe reste constant à partir d'un certain rang, et on peut en déduire que ces séries donnent des fonctions qui croissent exponentiellement quand t tend vers $\pm\infty$, ou (puisque'il s'agit de séries de Laurent) qu'elles sont singulières pour $t = 0$. Dans tous ces cas, il est impossible que l'intégrale

$$\int_0^\infty t^2 |f(t)|^2 dt$$

converge, donc les fonctions ψ correspondantes ne seront pas dans $\mathbb{L}_2(\mathbb{R}^3)$.

C'est ainsi qu'on montre que les E_n (les niveaux de Bohr) sont les seules valeurs propres possibles, de multiplicité n^2 . Avec les indications ci-dessus, cela est laissé en exercice (mais les calculs sont longs!).

FIN