

1 Les opérateurs linéaires.

1.1 Introduction

Une des tâches que l'on rencontre régulièrement en physique est de résoudre des équations différentielles linéaires. De façon générale, nous pouvons représenter ces équations par $\mathcal{L}y = f$, où y est la fonction inconnue à rechercher, f une fonction connue, et \mathcal{L} un opérateur différentiel. Par exemple, l'équation de la chaleur peut s'écrire $\mathcal{L}u = q(x, t)$, où $\mathcal{L} = (\partial_t - D\partial_x^2)$ et $q(x, t)$ est le terme de source. A priori, la recherche des solutions est du domaine de l'analyse. Nous allons voir cependant que nous pouvons ramener la résolution de ces équations dans le domaine de l'algèbre matricielle des systèmes à n équations et n inconnus¹ du genre $AX = B$. Le très grand avantage est que pour faire de l'algèbre, nous n'avons, en gros, que besoin d'addition et de multiplication². Les transformées de Fourier et de Laplace que nous avons rencontrés dans ce cours étaient des exemples particuliers d'outils bien adaptés à une certaine classe d'équations qui nous permettaient de ramener l'analyse à l'algèbre. Nous allons généraliser cette approche et voir toute la puissance de feu que cela nous procure.

Depuis le début de ce cours, nous insistons fortement sur le concept de vecteur. Au premier chapitre, nous avons vu que dans un espace vectoriel, nous pouvons définir des bases : cela nous permet de manipuler les vecteurs à l'aide de colonnes (ou de lignes) de nombres. Nous avons également vu que si nous disposons d'un produit scalaire, cela facilite grandement la tâche de trouver les coefficients d'un vecteur dans une base orthogonale.

Un vecteur peut être un objet aussi simple qu'un vecteur usuel du plan euclidien, ou un objet beaucoup plus complexe tel qu'une fonction. Prenons le cas d'une fonction f . Une fonction est une *machine* qui prend un nombre en entrée et produit un nombre en sortie. Nous pouvons choisir plusieurs représentations pour une même fonction. Par exemple, si nous choisissons la base de Fourier avec les vecteurs de la base $\exp(ikx)$, f s'écrira comme une superposition de ces fonctions, chacun avec un poids $\tilde{f}(k)$ (nous avons absorbé ici le facteur $1/2\pi$ dans la définition de \tilde{f}) :

$$f(x) = \int_{-\infty}^{+\infty} \tilde{f}(k) \exp(ikx) dk$$

L'intégral ici n'effectue rien d'autre que la superposition des vecteurs de base avec leurs poids correspondants. Il est également usuel de représenter la fonction f par un tableau qui à chaque entrée numérique, associe un nombre, et on note cela $f(x)$ (la notation est

1. n étant infini en l'occurrence, mais ceci n'induit pas de difficultés particulières

2. Voir également le chapitre ?? sur la signification d'une équation différentielle

1 Les opérateurs linéaires.

un peu confuse). En réalité, cela revient à représenter une fonction f sur la base des δ de Dirac, où chaque valeur $f(x)$ est le poids associé à un Dirac centré sur x :

$$f(x) = \int_{-\infty}^{+\infty} f(y)\delta(x-y)dy$$

A nouveau, l'intégral ne fait rien d'autre que de superposer des vecteurs de la base.

Revenons maintenant au concept général de vecteur. Soit l'espace vectoriel \mathcal{E} . Nous pouvons définir des opérations qui transforment un vecteur en un autre, et cela de façon linéaire. Dans l'espace des vecteurs du plan euclidien, la rotation ou la projection sur un axe sont de telles opérations. Par exemple, la rotation de la somme de deux vecteurs égale la somme de leurs rotations : $R(e_1 + e_2) = Re_1 + Re_2$. Dans l'espace des fonctions infiniment dérivables, l'opération dérivation D est une opération linéaire : elle transforme un vecteur (une fonction) en un autre, et cela de façon linéaire. De même, l'opération intégration $I[f] = \int_0^x f(y)dy$. De façon général, nous appelons, dans l'espace des fonctions, un opérateur linéaire comme une machine qui prend en entrée une fonction et produit en sortie une autre fonction, et fait cela de façon linéaire :

$$O[\lambda f_1 + \mu f_2] = \lambda O[f_1] + \mu O[f_2]$$

où λ et μ sont des scalaires et f_1 et f_2 des fonctions.

Exercices. Soit l'opérateur X qui prend une fonction f en entrée et produit la fonction multipliée par x en sortie : $X[f(x)] = xf(x)$ [exemple : $X[\sin(x)] = x \sin(x)$]. Démontrer que c'est un opérateur linéaire. Qu'en est-il de l'opérateur X^2 : $X^2[f(x)] = x^2f(x)$? Soit $V(X)$ un opérateur tel que $V(X)[f(x)] = V(x)f(x)$. Est-ce que ce dernier est linéaire? Même question pour l'opérateur 1 : $1[f(x)] = f(x)$. L'opérateur 0 est l'opérateur qui associe à n'importe quelle fonction la fonction 0 . Est-il linéaire? De façon général, l'opérateur λ associe à une fonction f la fonction λf (il faut avouer que la notation est vraiment confuse entre l'opérateur λ et le scalaire λ ; on s'y habitue vite cependant). Soit l'opérateur de translation $T_\epsilon[f(x)] = f(x + \epsilon)$; démontrer qu'il est linéaire.

Note sur les notations. Pour manipuler les opérateurs linéaires, la coutume est de laisser tomber les signes du genre $()$ et $[]$. Ainsi, nous écrivons Of où même pire, $Of(x)$ à la place de $O[f(x)]$. La confusion est gênante quand on écrit par exemple, $Xf(x) = xf(x)$. X ici est un opérateur, $f(x)$ et $xf(x)$ sont des fonctions³; $Xf(x)$ est la fonction qui résulte de l'application de X à f . Pour éviter un peu ces confusions, la convention que nous suivrons dans ce cours et de toujours noter les opérateurs par des lettres majuscules.

1.2 L'algèbre des opérateurs.

Se donner une algèbre est se donner un ensemble \mathcal{E} ou l'on définit les deux opérations $+$ et \cdot (produit) entre ses membres, avec toutes les propriétés d'associativités usuelles

3. A vrai dire, c'est encore pire : f est une fonction, $f(x)$ est un nombre. A nouveau, on a l'habitude de ne pas toujours distinguer explicitement les deux choses et laisser le boulot au cerveau. Dans ce cas, le cerveau agit comme un vulgaire compilateur C , testant constamment le *type* des notations qu'on utilise.

1 Les opérateurs linéaires.

que vous connaissez : soit $a, b, c \in \mathcal{E}$, alors

$$\begin{aligned} a + b, a.b &\in \mathcal{E} \\ a + b &= b + a \\ a(b + c) &= ab + ac; (a + b)c = ac + bc \\ (ab)c &= a(bc) \end{aligned}$$

Nous devons avoir quelques propriétés de plus pour mériter le nom d'algèbre. Il faut qu'il existe des éléments neutre vis à vis des deux opérations, qu'on appelle 0 et 1 : $a + 0 = 0$ et $a1 = 1a = a$. De plus, les inverses des éléments vis à vis de + et de . doivent exister : pour chaque élément a , il doit exister un élément unique, qu'on appelle $-a$, tel que $a + (-a) = 0$; de même, il doit exister un élément unique qu'on note $1/a$ ou a^{-1} tel que $a.(1/a) = (1/a)a = 1$ (pour $a \neq 0$).

L'ensemble des nombres (rationnels ou réel), équipé de + et de . usuel, constitue une algèbre. C'est un cas un peu particulier, puisqu'en plus, la multiplication y est commutative ($ab = ba$). Mais tous les théorèmes qui ont été démontrés pour l'algèbre des nombres sans invoquer la commutativité du produit sont valable pour n'importe quelle autre algèbre.

Nous pouvons définir une algèbre pour les opérateurs linéaires. Nous devons d'abord préciser le sens de l'égalité entre opérateurs. Nous dirons que $O_1 = O_2$ si le résultat de l'application de ces deux opérateurs à une fonction est le même, quel que soit la fonction ⁴ : $\forall f, O_1f = O_2f$.

L'opération + entre opérateurs hérite directement sa définition de l'opération + entre les fonctions⁵ : L'opérateur $O_1 + O_2$ est l'opérateur qui, appliqué à une fonction, produit la fonction $O_1f + O_2f$ (bien noter que là, l'addition est entre deux fonctions). L'opération . est la combinaison d'opérateur : O_1O_2 est l'opérateur qui à la fonction f , associe la fonction $O_1[O_2[f]]$. On peut affirmer maintenant que l'ensemble des opérateurs linéaires muni de + et de . constitue une algèbre. Les opérateurs 0 et 1 sont les éléments neutres de l'addition et de la multiplication. Dans la suite de ce cours, l'ensemble des fonctions auxquels ces opérateurs s'appliquent est l'ensemble des fonctions \mathcal{L}_2 au moins deux fois dérivable.

Toutes les notations que nous utilisons dans l'algèbre classique peuvent être utilisées pour l'algèbre des opérateurs. Par exemple, X^2 est effectivement $X.X$. Si l'on désigne par ∂_x et ∂_y les opérateurs de dérivation par rapport à x et y , alors $\partial_x^2 - \partial_y^2 = (\partial_x - \partial_y)(\partial_x + \partial_y)$. Il faut juste faire attention à la commutativité : en général, deux opérateurs ne commutent pas : $O_1O_2 \neq O_2O_1$. Cela ne veut pas dire que l'on ne peut pas trouver deux opérateurs

4. Il est évident que $\forall f$ est une condition trop exigeante et n'a pas de sens en général. Si par exemple, l'opérateur contient des dérivées d'ordre n , nous devons comprendre $\forall f$ comme : quelque soit f " n -fois dérivable". A chaque fois, nous supposons l'ensemble des fonctions comme compatible avec la définition de l'opérateur. Nous n'entrerons pas plus dans le détail lors de ce cours, pour ne pas alourdir chaque assertion par un train de précautions et de conditions d'applicabilité. Dans la majorité des cas, nous supposons que nous travaillons avec l'ensemble des fonctions $\mathcal{L}^2[-\infty, +\infty]$ au moins deux fois continuellement dérivable. Ceci surtout impose à nos fonctions qu'elles et leurs dérivées $\rightarrow 0$ quand leur argument $\rightarrow \infty$.

5. De même que dans l'espace des fonctions, l'opération + est héritée de l'addition entre les nombres : la fonction $f + g$ est la fonction qui associe au nombre x le nombre $f(x) + g(x)$.

1 Les opérateurs linéaires.

qui commutent, l'exemple précédent de ∂_x et ∂_y le montre. Simplement, il ne faut pas le présumer à l'avance.

Exemple 1. Démontrons que $DX - XD = 1$ (bien noter que ceci est une égalité entre opérateurs). Nous avons $DX[f(x)] = (d/dx)(xf(x)) = xf'(x) + f(x)$ et $XD[f(x)] = xf'(x)$. Donc $(DX - XD)[f(x)] = f(x)$: Le résultat d'application de $DX - XD$ à une fonction est le même que l'application de l'opérateur 1, d'où l'égalité entre les opérateurs.

Définition. On appelle commutateur de deux opérateurs O_1 et O_2 , l'opérateur $O_1O_2 - O_2O_1$. Il est usuel de noter ce dernier $[O_1, O_2]$.

Le commutateur joue un très grand rôle en mécanique quantique. Depuis le livre de Paul Dirac en 1930, la mécanique quantique est formulée à travers les relations de commutations.

Fonctions d'opérateurs. Beaucoup de fonctions usuelles sont définies à l'aide de séries algébriques, comme par exemple, $\exp(x) = \sum_n x^n/n!$. Puisque l'on dispose d'une algèbre pour les opérateurs linéaires, nous pouvons faire de même et définir des fonctions d'opérateurs, comme par exemple $\exp(O)$ ou $\log(1 + O)$, qui sont elles mêmes des opérateurs linéaire. Par exemple, pour l'opérateur D et le scalaire ϵ ,

$$\exp(\epsilon D) = \sum_{n=0}^{\infty} (1/n!) \epsilon^n D^n$$

et le résultat de son application à une fonction $f(x)$ produit la fonction

$$\sum_{n=0}^{\infty} (1/n!) \epsilon^n f^{(n)}(x) = f(x + \epsilon)$$

L'opérateur $\exp(\epsilon D)$ n'est donc rien d'autre que l'opérateur de translation T_ϵ vu plus haut. Bien sûr, dès que l'on parle de suites et de séries infinies, nous avons besoin de la notion de *convergence*. La convergence dans l'espace des opérateurs hérite sa définition de la convergence dans l'espace des fonctions : nous dirons que la suite O_n converge vers O si la suite des *fonctions* $O_n f$ converge vers la fonction $O f$ quelque soit f (cf la note 4).

Disposer des fonctions d'opérateurs nous permet de résoudre symboliquement nombre d'équations à dérivées partielles (EDP). Prenons d'abord l'équation différentielle ordinaire

$$dy/dt = ay \tag{1.1}$$

avec la condition initiale $y(t = 0) = y_0$ où a et y_0 sont des scalaires ne dépendant pas du temps. La solution bien connue est $y(t) = \exp(ta)y_0$. La valeurs de la fonction à un temps ultérieur t s'obtient en appliquant (multipliant) le scalaire $\exp(ta)$ à la condition initiale.

1 Les opérateurs linéaires.

Soit maintenant l'EDP de premier ordre

$$\partial_t \psi - \partial_x \psi = 0 \quad (1.2)$$

où $\psi(x, t)$ est une fonction des deux variables x, t ; la condition initiale étant $\psi(x, t = 0) = \psi_0(x)$, $\psi_0(x)$ étant une fonction connue de la variable x . Nous pouvons écrire cette équation sous forme opératorielle

$$\partial \psi / \partial t = D \psi$$

où $D = \partial / \partial x$ est un opérateur qui ne dépend pas de t . En s'inspirant de l'exemple de l'équation (1.1), nous pouvons donner la solution comme

$$\psi(x, t) = \exp(t.D) \psi_0(x)$$

Nous savons par ailleurs que l'opérateur $\exp(tD)$ n'est rien d'autre que l'opérateur "translation d'une quantité t ". La solution s'écrit donc

$$\psi(x, t) = \psi_0(x + t)$$

Ceci est la solution exacte de l'EDP (1.2) que nous pouvons obtenir soit par des transformées de Fourier-Laplace, soit par la méthode des caractéristiques du chapitre ???. Ceci n'est pas une analogie. Les mêmes règles d'algèbre (et d'analyse) que nous appliquons aux fonctions ont été définies pour les opérateurs et nous donnent le droit de les manipuler symboliquement de la même façon. Voyons cela d'un peu plus près. La fonction $f(t) = \exp(ta)$ est donnée comme la série

$$\exp(ta) = \sum_{n=0} t^n a^n / n!$$

sa dérivée (qui coïncide avec la dérivée terme à terme de la série) s'écrit (en jouant sur l'indice de sommation)

$$\begin{aligned} f'(t) &= \sum_{n=0} t^n a^{n+1} / n! \\ &= a \sum_{n=0} t^n a^n / n! \\ &= a f(t) \end{aligned}$$

et c'est pour cela que cette fonction est la solution générale de (1.1). En utilisant les mêmes règles de manipulation pour l'algèbre d'opérateurs, nous voyons que l'opérateur $\exp(tD)$ possède comme dérivée par rapport au temps l'opérateur $D \exp(tD)$. La fonction $\psi(x, t) = \exp(tD) \psi_0(x)$ a donc pour dérivée par rapport au temps la fonction $D \exp(tD) \psi_0(x)$, c'est à dire $D \psi(x, t)$.

De façon générale, nous pouvons avoir des EDP du genre

$$\partial \psi / \partial t = H \psi$$

1 Les opérateurs linéaires.

où H est un opérateur spatial (ne dépendant que de la variable x) plus compliqué que le simple opérateur de dérivation spatiale D , mais la discussion ci-dessus reste valide et nous pouvons donner la solution comme

$$\psi(x, t) = \exp(tH)\psi(x, t = 0)$$

L'opérateur $\exp(tH)$ n'est plus alors une simple translation, mais le principe reste le même : la fonction solution à un temps ultérieur t est donnée par l'application de l'opérateur $\exp(tH)$ à la fonction "condition initiale". L'opérateur $\exp(tH)$ est appelé, à juste titre, l'opérateur de l'évolution temporelle. Cette façon de présenter les choses est appelée, en mécanique quantique l'interprétation d'Heisenberg (voir plus bas pour une digression historique). Évidemment, cette façon de résoudre l'équation ne nous sert à rien si nous ne savons pas calculer $\exp(tH)$. Nous verrons plus bas les outils que les mathématiciens ont développé pour calculer efficacement ce genre de fonctions d'opérateurs.

Exercices.

1. Identité de Jacobi. Démontrer que pour trois opérateurs A, B, C , nous avons

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$$

2. Démontrer que $D^2 + X^2 = (D + X)(D - X) + 1 = (D - X)(D + X) - 1$. En déduire $[D - X, D + X]$ et $[D^2 + X^2, D \pm X]$. [indication : Utiliser la relation $[D, X] = 1$].

3. Commutateur des opérateurs. Démontrer que les opérateurs O et $f(O)$ commutent. $f(x)$ est une fonction analytique dans le voisinage de $x = 0$. Même chose pour $g(O)$ et $f(O)$. Démontrer que si A et B commutent, alors $f(A)$ et $g(B)$ commutent.

4. Dans l'espace des opérateurs linéaires sur les fonctions à trois variables, nous définissons $L_z = X\partial_y - Y\partial_x$; L_x et L_y sont définis cycliquement à partir de cette dernière définition. Calculer $[L_\alpha, L_\beta](\alpha, \beta = x, y, z)$. Donner la définition de L_α^2 et de $L^2 = \sum_\alpha L_\alpha^2$. Calculer $[L^2, L_\alpha]$.

5. Opérateur rotation. En utilisant les règles de dérivation en chaîne, démontrer que

$$\frac{\partial}{\partial \phi} = L_z$$

que représente alors l'opérateur $\exp(\alpha L_z)$? $[(r, \theta, \phi)$ sont les coordonnées du point en coordonnées sphérique]

6. Exponentiel d'un opérateur. démontrer que

$$\frac{d}{dt} \exp(tA) = A \exp(tA)$$

où A est un opérateur linéaire. **Help :** Utiliser le développement de l'exponentiel et les règles habituelles de la dérivation.

7. Exponentiel d'un opérateur (encore). Démontrer que $\exp(P^{-1}AP) = P^{-1} \exp(A)P$ où A, P sont deux opérateurs linéaires.

8. Commutation et exponentiel. Soit la matrice

$$A = \begin{pmatrix} 0 & -x \\ x & 0 \end{pmatrix}$$

Calculer A^2 . En déduire une expression générale pour A^n selon que n est pair ou impair. Démontrer alors que

$$\exp(A) = \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}$$

Help : Décomposer la somme en termes pairs et impairs, et utiliser le développement en série des fonctions sin et cos. Soit maintenant les deux matrices

$$C = \begin{pmatrix} 0 & -x \\ 0 & 0 \end{pmatrix}, D = \begin{pmatrix} 0 & 0 \\ x & 0 \end{pmatrix}$$

Démontrer que $C^2 = D^2 = 0$ et en déduire $e^C \cdot e^D$. Que peut on dire de $e^C e^D$ et e^{C+D} ? Est ce que C et D commutent?

9. e^{A+B} Pour deux opérateurs A et B qui ne commutent pas à priori, démontrer que

$$e^{(A+B)t} = e^{At} + \int_0^t e^{A(t-s)} B e^{(A+B)s} ds$$

Cette relation, appelé en mécanique quantique relation de Dyson, est très utile pour évaluer l'exponentiel d'un opérateur, si $\exp(At)$ est connu et que la matrice de l'opérateur B est très creuse. [indication : Quel est la solution de l'équation $y' - ay = f(x)$? Obtenir une équation analogue pour l'opérateur $\exp(A+B)t$]. En profiter pour donner l'expression de $\exp(A+B)$.

10. Equation d'onde. Résoudre par la méthode des opérateur l'équation

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0$$

avec les conditions initiales $u(x, 0) = f(x)$ et $\partial_t u(x, t)|_{t=0} = g(x)$ en l'écrivant sous la forme symbolique $\partial^2 u / \partial t^2 - c^2 D^2 u = 0$ et en vous inspirant de la solution de l'équation ordinaire $u'' - a^2 u = 0$.

1.3 Représentation matricielle des opérateurs.

Le grand avantage de disposer d'une base est de pouvoir manipuler les vecteurs à l'aide des nombres. Le même avantage est obtenu pour les opérateurs. Par exemple, étant donné un vecteur du plan euclidien, nous n'avons pas à nous munir de compas et de rapporteur pour effectuer une rotation, mais seulement à effectuer des additions et des multiplications sur des colonnes de chiffres. Voyons comment cela fonctionne.

Soit un opérateur linéaire R dans un espace \mathcal{E} que nous avons équipé d'une base $\{e_1, \dots, e_n\}$. Soit un vecteur v quelconque. Notre but est de pouvoir calculer le résultat de l'application de R à v , qui est un autre vecteur de \mathcal{E} . Nous pouvons décomposer v dans la base donnée :

$$v = \sum_j a_j e_j$$

et comme R est linéaire,

$$R.v = \sum_j a_j (R.e_j)$$

Pour entièrement caractériser l'opérateur R , nous avons simplement besoin de connaître l'action qu'il effectue sur chaque élément de la base. Par ailleurs, chaque $R.e_i$ (pour $i = 1, \dots, n$) est un vecteur de l'espace \mathcal{E} , et donc décomposable sur la base $\{e_1, \dots, e_n\}$:

$$R.e_j = \sum_i r_{ij} e_i \quad (1.3)$$

Donc, pour entièrement connaître R , il suffit de connaître les $n \times n$ nombres r_{ij} . Nous pouvons alors connaître le résultat de l'application de R à n'importe quel vecteurs de l'espace.

$$R.v = \sum_{i,j} r_{ij} a_j e_i \quad (1.4)$$

Il est habituel de représenter la décomposition de $R.e_j$ comme la j -ième colonne d'un tableau de n lignes et appeler le tableau une matrice.

Exemple 1. Rotation dans le plan euclidien. Caractérisons la rotation de $\pi/2$ dans la base orthonormée habituelle du plan euclidien (e_x, e_y) . Nous savons que $R.e_x = e_y = 0e_x + 1e_y$. La première colonne de la représentation matricielle de R dans cette base sera donc $(0, 1)$. De même, $R.e_y = -1e_x + 0e_y$. La représentation de R est donc

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Il est très important de faire la distinction entre l'opérateur R et sa représentation matricielle. La représentation matricielle est comme une photo de l'opérateur : elle dépend de la base choisie, comme la photo dépend du point de vue du photographe et de l'humeur de la personne qui se fait photographier. Mais la photo *n'est pas* la personne.

1 Les opérateurs linéaires.

Exemple 2. Dérivation dans l'espace des fonctions. Choisissons la base de Fourier $\{e_q = \exp(iqx)\}$. La base est bien sûr infinie, (et même très infinie!), mais cela n'a pas d'importance. Nous avons

$$D.e_q = iq e_q$$

La représentation matricielle de D dans la base de Fourier ne comporte que des éléments diagonaux. Les éléments $d_{qq'}$ de la matrice de D sont donc donnés par $d_{qq'} = 0$ si $q \neq q'$ et par $d_{qq} = iq$. Le fait que D soit diagonale dans la base de Fourier est la raison de son utilisation pour la résolution de l'équation d'onde ou de la chaleur. L'application de D à une fonction quelconque donne

$$Df(x) = \int_{-\infty}^{+\infty} iq \tilde{f}(q) \exp(iqx) dq$$

ce qui n'est rien d'autre qu'une réécriture de l'expression (1.4).

Nous voyons à travers ce dernier exemple un fait important. Peu importe que notre espace soit fini ou infini, à partir du moment où nous disposons de bases (finies, infinies discrètes ou infinies continues) et que les convergences sont assurées, nous pouvons manipuler les opérateurs linéaires comme des matrices.

Base orthonormale. Nous avons vu au premier chapitre que disposer d'un produit scalaire $(.,.)$ et d'une base orthonormale facilite énormément les choses. Soit $\{e_1, \dots, e_n\}$ une base orthonormale, c'est à dire $(e_i, e_j) = \delta_{i,j}$ et soit la décomposition d'un vecteur quelconque $v : v = \sum a_i e_i$. Alors

$$\begin{aligned} (v, e_k) &= \left(\sum_i a_i e_i, e_k \right) \\ &= \sum_i a_i (e_i, e_k) \\ &= a_k \end{aligned}$$

Le coefficient a_k de la décomposition du vecteur v sur la base $\{e_i\}$ n'est rien d'autre que le produit scalaire entre v et e_k . Le même principe nous donne les coefficients r_{ij} d'un opérateur dans une base orthonormée. En partant de l'expression (1.3), nous voyons que

$$r_{ij} = (e_i, R.e_j)$$

Pour trouver r_{ij} , il suffit d'appliquer R à e_j et ensuite prendre le produit scalaire de ce dernier avec e_i .

Exercices.

1. Soit la base de Fourier $\{1, \cos(2\pi nx/L), \sin(2\pi nx/L)\}$ pour les fonctions définies sur $[0, L]$. En ordonnant correctement les éléments de la base, donner l'expression de l'opérateur D et D^2 dans cette base.

1 Les opérateurs linéaires.

2. Soit les fonctions d'Hermite

$$h_n(x) = C_n(-1)^n \exp(x^2/2) \frac{d^n}{dx^n} \exp(-x^2) \quad (1.5)$$

Les premières fonctions sont (en multipliant par $\exp(-x^2/2)$) $1, 2x, 4x^2 - 2, \dots$. Les coefficients C_n , que nous n'explicitons pas, assurent que les fonctions sont normées. On peut démontrer, avec un peu d'effort, que les fonctions h_n sont deux à deux orthogonales et forment une base.

Démontrer que pour l'opérateur

$$H = -D^2 + X^2$$

nous avons $Hh_n(x) = (2n + 1)h_n(x)$. Donner alors la représentation de H dans la base des h_n .

3. **Opérateur D dans la base des Bessel.** Soit les fonctions de Bessel I_n définies par

$$I_n(x) = \frac{1}{\pi} \int_0^\pi e^{x \cos \theta} \cos(n\theta) d\theta$$

Démontrer que

$$I_n'(x) = (1/2)(I_{n-1}(x) + I_{n+1}(x))$$

Les fonctions $\exp(-x)I_n(x)$ ($n = 0, 1, \dots$) forment une base. Donner l'expression de la matrice de l'opérateur D dans cette base. Cette matrice joue un rôle très important dans les problèmes de matrices tridiagonales.

4. Montrer que si l'opérateur M commute avec tous les autres opérateurs, alors $M = \alpha I$. [Help : donnez vous une base quelconque (e_1, \dots, e_n) et considérer l'opérateur projection sur e_1 : $P_1 e_j = \delta_{i,j} e_1$. En utilisant la commutation de M avec cet opérateur et d'autres opérateurs de projection, démontrer que la matrice de M dans cette base est diagonale. Ensuite, considérer un opérateur de permutation cyclique $\Pi e_i = e_{i+1}$ et considérer sa permutation avec M : en déduire que tous les éléments diagonaux de M sont alors égaux.]

5. **Laguerre et autres.**

1.4 Valeurs et vecteurs propres.

Quand on travail avec un opérateur linéaire et que l'on souhaite résoudre une EDP qui lui est associé, certaines fonctions (ou vecteur) ont un rôle privilégié. On appelle vecteur propre (*eigenvector* en anglais) d'un opérateur O un vecteur $v_n \neq 0$ tel que

$$Ov_n = \lambda_n v_n$$

où λ_n est un scalaire qui est appelé valeur propre du vecteur v_n .

1 Les opérateurs linéaires.

En géométrie plane, pour l'opérateur de projection sur l'axe x , les deux vecteurs \mathbf{e}_x et \mathbf{e}_y sont des vecteurs propres, avec les valeurs propres 1 et 0. En géométrie à trois dimensions, les trois vecteurs \mathbf{e}_x et \mathbf{e}_y et \mathbf{e}_z sont des vecteurs propres de la projection sur l'axe x , le premier associé à $\lambda = 1$, les deux autres à $\lambda = 0$. Dans le plan, Pour l'opération de rotation de π , tous les vecteurs du plan sont des vecteurs propres associés à la valeur $\lambda = -1$ (remarquer cependant que $R_\pi = -I$). Enfin, la rotation de $\pi/2$ n'a pas de vecteurs propres.

Exemple. Dans l'espace des fonctions, $\exp(igx)$ est vecteur propre de l'opérateur $D = \partial_x$ avec la valeur propre ig . Les fonctions $\sin(qx)$ et $\cos(qx)$ sont les vecteurs propres de l'opérateur D^2 avec $\lambda = -q^2$. Les fonctions d'Hermite $h_n(x)$ que nous avons rencontré plus haut sont fonctions propres de l'opérateur $-D^2 + X^2$ avec la valeur propre $\lambda_n = (2n + 1)$.

Si les vecteurs propres sont suffisamment nombreux pour former une base, on les appelle une base propre. Évidemment, la représentation d'un opérateur dans sa base propre est diagonale.

Disposer d'une base propre accélère (ou rend possible) la solution des problèmes. Prenons le système à deux équations et deux inconnus

$$\begin{aligned} 2x + 3y &= 4 \\ x - 2y &= 1 \end{aligned}$$

formellement, nous pouvons le représenter comme $Au = B$ où A est la matrice

$$\begin{pmatrix} 2 & 3 \\ 1 & -2 \end{pmatrix}$$

$u = (x, y)^T$ et $B = (4, 1)^T$. A est la représentation matricielle d'une application linéaire dans la base $e_1 = (1, 0)^T$ et $e_2 = (0, 1)^T$. La résolution de ce système nécessite quelques opérations d'additions et de combinaisons. Par contre, le système

$$\begin{aligned} 2x &= 4 \\ 3y &= 1 \end{aligned}$$

se résout immédiatement comme deux équations indépendantes. Formellement, on pouvait l'écrire $Au = B$ où cette fois A est la matrice

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

dans la base e_1, e_2 , la matrice de l'opérateur A est diagonale, ce qui permet de ramener la résolution d'un système de deux équations à deux inconnus à la résolution de deux équations à *une* inconnue. Évidemment, un système 2×2 est facile à résoudre, pour un système de 4×4 ou 1000×1000 , le gain est déjà plus appréciable. Dans l'espace de dimension *infinie* de Hilbert, trouver une base propre est souvent la seule possibilité de résoudre un problème, comme nous le verrons plus bas.

1.5 Disposer d'une base propre orthogonale.

Le très grand avantage de disposer d'une base propre orthogonale (ou encore mieux, orthonormale) est de pouvoir résoudre des équations aux dérivées partielles. Nous avons vu une application de cette méthode dans le chapitre sur les séries et transformées de Fourier. Nous allons voir cela dans des cas plus généraux.

Supposons que nous voulons résoudre l'équation aux dérivées partielles

$$\partial_t \psi(x, t) = H\psi(x, t) \quad (1.6)$$

où l'on suppose que l'opérateur H n'a que des composantes spatiales. Les conditions initiales et aux limites de cette équation sont les suivantes :

$$\begin{aligned} \psi(x \rightarrow \pm\infty, t) &= 0 \\ \psi(x, 0) &= g(x) \end{aligned}$$

la fonction tend vers zéro pour les grand x ; à l'instant $t = 0$, la fonction recherchée prend la forme de $f(x)$. Supposons que nous connaissons une base orthonormale $\{f_n(x)\}$ dans laquelle l'opérateur H est diagonale⁶ :

$$Hf_n(x) = \lambda_n f_n(x)$$

A chaque instant t , nous pouvons décomposer la fonction (inconnu) $\psi(x, t)$ sur la base des $\{f_n\}$:

$$\psi(x, t) = \sum_{i=0}^{\infty} a_n(t) f_n(t)$$

Les $a_n(t)$ sont les coefficients de cette décomposition et varient bien sûr d'un instant à un autre. En utilisant cette décomposition dans l'équation (1.6) nous obtenons⁷

$$\sum_{i=0}^{\infty} a'_n(t) f_n(x) = \sum_{i=0}^{\infty} \lambda_n a_n(t) f_n(x)$$

Comme les f_n constituent une base, les coefficients de la décomposition sont uniques et

$$a_n(t) = a_{n,0} \exp(\lambda_n t)$$

Il nous reste juste à trouver les coefficients $a_{n,0}$; ceci s'obtient en utilisant la condition initiale

$$g(x) = \sum_{i=0}^{\infty} a_{n,0} f_n(x)$$

6. Nous supposons que les fonctions f_n sont compatibles avec les conditions aux bords : $f_n(x \rightarrow \pm\infty) = 0$.

7. Nous supposons que l'on peut intervertir les opérations de sommation et de dérivation par rapport au temps.

1 Les opérateurs linéaires.

et donc

$$a_{n,0} = \langle f_n(x), g(x) \rangle$$

Résumons la méthode. On décompose la condition initiale sur la base des fonctions propres. L'amplitude de chaque composante $a_n(t)$ varie exponentiellement avec l'exposant λ_n . Il suffit de recombinaison ces composantes à un temps ultérieurs t pour recalculer la fonction ψ en ce temps là :

$$\psi(x, t) = \sum_{i=0}^{\infty} \langle \psi(x, 0), f_n(x) \rangle \exp(\lambda_n t) f_n(x) \quad (1.7)$$

Cette façon de calculer la fonction ψ est appelé en mécanique quantique l'interprétation de Schrödinger ⁸.

Exemple : oscillateur harmonique. En mécanique quantique, l'équation d'une particule dans un potentiel harmonique s'écrit (en choisissant convenablement les unités)

$$i \frac{\partial \psi}{\partial t} = \left(-\frac{\partial^2}{\partial x^2} + x^2 \right) \psi$$

D'après l'exercice que nous avons vu sur les fonctions de Hermite, il n'est pas difficile d'explicitier la solution sous la forme de

$$\psi(x, t) = \sum a_n e^{-i(2n+1)t} h_n(x)$$

Comparaison Schrödinger-Heisenberg. L'interprétation de Schrödinger et Heisenberg sont bien sûr équivalente. Reprenons l'équation (1.6) . Dans l'interprétation d'Heisenberg, la solution s'écrit

$$\psi(x, t) = e^{tH} \psi(x, 0) \quad (1.8)$$

comme nous l'avons indiqué, la fonction ψ au temps t s'obtient en appliquant l'opérateur $\exp(tH)$ à la fonction ψ au temps 0. Dans sa base propre (les $f_n(x)$ ci-dessus), l'opérateur H est diagonal. En utilisant la définition de l'opérateur

$$\exp(tH) = \sum_{n=0}^{\infty} (1/n!) t^n H^n$$

nous voyons que l'opérateur $\exp(tH)$ est également diagonal dans cette base, et ses éléments diagonaux s'écrivent $((\exp(t\lambda_1), (\exp(t\lambda_2), \dots (\exp(t\lambda_n), \dots)$. En notation matricielle, nous avons

$$H = \begin{pmatrix} \lambda_1 & 0 & \cdot & \cdot \\ 0 & \lambda_2 & 0 & \\ \cdot & 0 & \lambda_3 & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

8. Heisenberg et Schrödinger ont formulé ces deux approches dans les années 1922-28 quand les méthodes d'analyse fonctionnelle n'étaient pas encore popularisées chez les physiciens. La contribution de Von Neumann était de montrer, en important les concepts développés par Hilbert en mathématiques, que les deux approches étaient équivalentes. Dirac a mis la dernière main à l'édifice en 1932 en formulant l'ensemble de façon extrêmement élégante.

1 Les opérateurs linéaires.

et

$$e^{tH} = \begin{pmatrix} e^{t\lambda_1} & 0 & \cdot & \cdot \\ 0 & e^{t\lambda_2} & 0 & \\ \cdot & 0 & e^{t\lambda_3} & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot \end{pmatrix}$$

Si maintenant nous exprimons la fonction $\psi(x, 0)$ dans la même base

$$\psi(x, 0) = \sum_n \langle \psi(x, 0), f_n(x) \rangle f_n(x)$$

nous voyons que l'interprétation d'Heisenberg (1.8) exprime exactement la même chose que l'interprétation de Schrödinger (1.7).

Fonctions d'opérateurs. Le paragraphe précédent nous indique comment calculer une fonction d'opérateur, même quand nous ne connaissons pas la série de la fonction. Soit l'opérateur A , diagonal dans sa base propre (e_1, e_2, \dots) et possédant les valeurs propres $(\lambda_1, \lambda_2, \dots)$. L'opérateur $f(A)$ est définie de telle façon que sa représentation dans la même base soit diagonale, avec les éléments $(f(\lambda_1), f(\lambda_2), \dots)$.

1.6 Opérateurs hermitiens.

Dans la théorie des opérateurs linéaires, une certaine classe d'opérateurs qu'on appelle hermitiens joue un rôle fondamental, puisqu'on peut démontrer que ces opérateurs sont diagonalisables et qui apparaissent très souvent dans les problèmes de physique.

Nous supposons dorénavant que nous disposons d'un produit scalaire (\cdot, \cdot) . On appelle l'adjoint d'un opérateur O l'opérateur O^\dagger tel que quelque soit les deux vecteurs u, v

$$(u, Ov) = (O^\dagger u, v)$$

Supposons que le produit scalaire est défini sur le corps des réels. Alors $(u, v) = (v, u)$. Si dans une base (que nous prenons orthonormale pour plus de simplicité, mais sans nécessité aucune) l'opérateur O est représenté par la matrice $O_{ij} = (e_i, Oe_j)$, on peut démontrer sans difficulté que pour la matrice de l'adjoint, $O_{ij}^\dagger = O_{ji}$: on interverti les lignes et les colonnes d'une matrice pour obtenir celle de son adjoint. La matrice résultante est appelé la transposée. Si le produit scalaire est défini sur le corps des complexes, alors $O_{ij}^\dagger = O_{ji}^*$.

Un opérateur est dit self-adjoint ou hermitien si $O^\dagger = O$.

Exemple. Dans l'espace des fonctions réelles ou complexes \mathcal{L}^2 (qui donc tendent vers 0 pour $x \rightarrow \pm\infty$), l'opérateur D n'est pas hermitien : une intégration par partie montre que $(u, Dv) = -(Du, v)$. Par contre, deux intégrations par partie montrent que l'opérateur D^2 est hermitien. De même, l'opérateur iD est Hermitien pour les fonctions complexes.

Les opérateurs hermitiens ont quelques propriétés remarquables, parmi lesquels les suivantes que nous demandons au lecteur de démontrer/découvrir à travers les exercices.

Exercices.

1. Les valeurs propres d'un opérateurs hermitiens sont réelles : Si $Ax = \lambda x$, $A = A^\dagger$, alors $\lambda \in \mathbb{R}$. [Indication : calculer (x, Ax) et $(x, A^\dagger x)$]
2. Les vecteurs propres associés à deux valeurs propres distinctes sont orthogonales : Si $Ax_1 = \lambda_1 x_1$, $Ax_2 = \lambda_2 x_2$, $A = A^\dagger$, $\lambda_1 \neq \lambda_2$, alors $(x_1, x_2) = 0$. [indication : calculer (x_1, Ax_2) et $(x_1, A^\dagger x_2)$].
3. Démontrer, dans l'espace des fonctions \mathcal{L}^2 que les valeurs propres de l'opérateur $-D^2$ sont positifs. [indication : il suffit de démontrer que $\forall x, (x, -D^2 x) \geq 0$, ce qui peut s'obtenir en effectuant une intégration par partie. De tels opérateurs sont appelés défini positif.].
4. Soit la fonction réelle $V(x)$ possédant un minimum absolu $E : \forall x \in \mathbb{R}, V(x) \geq E$. Démontrer d'abord que l'opérateur V est hermitien [définition : $V.f(x) = V(x)f(x)$]; Démontrer ensuite que toutes ses valeurs propres sont $\geq E$. [indication : il suffit de considérer $(u, (V - E)u)$ pour une fonction u quelconque et démontrer que sa valeur est toujours ≥ 0]. En combinant avec le résultat de la question précédente, que pouvez vous déduire pour l'opérateur

$$H = -D^2 + V$$
5. Pour un opérateur hermitien, le minimum de (x, Ax) avec la contrainte $(x, x) = 1$ est fournit par le vecteur propre de la matrice associée à la plus petite des valeurs propres; la valeur propre est juste un multiplicateur de Lagrange (voir le chapitre sur le calcul variationnel pour la définition des multiplicateurs de Lagrange).
6. Dans l'espace des fonctions \mathcal{L}^2 à trois variables définies sur \mathbb{R}^3 , démontrer que l'adjoint de l'opérateur gradient est l'opérateur ($-$ divergence), et vice et versa, tandis que l'adjoint du rotationnel est lui même.

1.7 Méthodes opératorielles, algèbre de Lie.

En algèbre classique, pour résoudre un problème, on essaye de le décomposer en termes plus simples. Par exemple, pour résoudre $x^2 - 3x + 2 = 0$, on peut remarquer que l'on peut la ramener à la forme $(x - 1)(x - 2) = 0$; on remarque ensuite que pour que ce produit soit 0, il suffit qu'un des termes soit nul, et qu'il suffit donc de résoudre les deux équations $x - 1 = 0$ et $x - 2 = 0$ pour avoir les solutions du problème. Or, ces deux dernières équations sont solubles par des techniques que nous connaissons déjà.

Les opérateurs linéaires possèdent une algèbre. On peut se demander si on ne peut pas entreprendre la même démarche pour décomposer des EDP compliquées en systèmes plus simples. La réponse est oui si on fait attention aux conditions aux limites et si on garde

en tête que l'algèbre des opérateurs est non commutative. Nous n'allons pas donner ici un exposé de ces méthodes, mais préférons l'illustrer à travers deux exemples.

1.7.1 L'oscillateur harmonique en mécanique quantique.

L'opérateur $H = -D^2 + V(X)$ joue un rôle très important en mécanique quantique. Appelant ψ la fonction d'amplitude, l'équation de Schrödinger est de la forme

$$i \frac{\partial \psi}{\partial t} = H\psi$$

Nous avons vu au §1.6 que H est hermitien et que toutes ses valeurs propres sont supérieures à E_0 , la valeur minimum de la fonction $V(x)$. L'opérateur $-D^2$ joue le rôle de l'énergie cinétique et l'opérateur $V(x)$ celui de l'énergie potentielle ; l'opérateur H est appelé Hamiltonien du système. Nous travaillons dans l'espace $\mathcal{L}^2[-\infty, +\infty[$ et la gestion des conditions aux limites ne pose pas trop de problème ($\psi(x \rightarrow \pm\infty) = 0$). Nous avons vu au §1.5 que résoudre cette équation revient à trouver les valeurs et vecteurs propres de l'opérateur H . Prenons le cas particulier du Hamiltonien

$$H = -D^2 + X^2$$

d'une particule se trouvant dans un potentiel harmonique. Nous savons que toutes les valeurs propres sont > 0 . Nous avons donné en exemple plus haut les fonctions et valeurs propres de cet hamiltonien. Nous allons voir que nous pouvons trouver ces fonctions sans résoudre aucune équation différentielle, un peu comme faire du beurre avec de l'eau. Il suffit juste de manipuler l'algèbre des opérateurs et surtout de leurs commutateurs. L'exemple que nous allons suivre est le prototype de ce genre de calcul et nous le donnons donc avec un peu de détail.

Comme $[D, X] = 1$, nous pouvons décomposer l'hamiltonien :

$$H = (-D + X)(D + X) + 1 = (D + X)(-D + X) - 1$$

Notons que l'opérateur X est hermitien $X = X^\dagger$ et que $D^\dagger = -D$. Nous pouvons donc poser

$$\begin{aligned} A &= D + X \\ A^\dagger &= -D + X \end{aligned}$$

et récrire $H = AA^\dagger - 1 = A^\dagger A + 1$. Ce qui implique naturellement que $[A, A^\dagger] = 2$.

Par ailleurs, on peut démontrer facilement que pour trois opérateurs quelconque, $[FG, F] = F[F, G]$ et $[F, FG] = F[F, G]$. Cette petite gymnastique nous permet d'obtenir

$$\begin{aligned} [H, A^\dagger] &= [A^\dagger A + 1, A^\dagger] \\ &= [A^\dagger A, A^\dagger] \\ &= 2A^\dagger \end{aligned} \tag{1.9}$$

L'équation (1.9) est un cas particulier d'un opérateur d'échelle. Avoir à sa disposition de telles relations est d'un très grand avantage comme nous allons le voir plus bas.

1 Les opérateurs linéaires.

Lemme. Si ψ_E est une fonction propre de H associée à la valeur propre E , alors $A^\dagger\psi_E$ est une fonction propre de H associée à la valeur propre $E + 2$.

Vous voyez ici l'avantage : si on connaît une valeur propre et une fonction propre, on peut en trouver beaucoup d'autre par l'application successive de l'opérateur A^\dagger . La démonstration est immédiate :

$$\begin{aligned} HA^\dagger\psi_E &= (2A^\dagger + A^\dagger H)\psi_E \\ &= (E + 2)A^\dagger\psi_E \end{aligned}$$

Nous pouvons effectuer la même démarche pour A et démontrer que $HA\psi_E = (E - 2)A\psi_E$. Comme nous savons que toutes les valeurs propres sont > 0 , il existe donc forcément une valeur propre minimum que nous appelons ϵ (noter que $0 < \epsilon < 2$). Toutes les autres valeurs propres sont donc de la forme $(2n + \epsilon)$, avec $n \in \mathbb{N}$.

Il nous reste maintenant à déterminer cette valeur ϵ . Appelons ψ_0 la fonction propre associée à ϵ . Nous devons alors obligatoirement avoir

$$A\psi_0 = 0 \tag{1.10}$$

sinon nous aurions des valeurs propres négatives. De là, nous pouvons naturellement déduire que

$$H\psi_0 = \psi_0$$

Et par conséquent, $\epsilon = 1$. Les valeurs propres de l'opérateur H sont donc de la forme $E_n = 2n + 1$. Quand je vous disais qu'on peut fait du beurre avec de l'eau !

Noter que la relation (1.10) est juste une équation différentielle

$$d\psi_0/dx + x\psi_0 = 0$$

avec les conditions $\psi_0 \rightarrow 0$ pour $x \rightarrow \pm\infty$. Cette équation est simple à résoudre

$$\psi_0(x) = C \exp(-x^2/2)$$

Les autres fonctions s'obtiennent à partir de là (cf l'exercice 2).

Exercices.

1. Donner les matrices des opérateurs A et A^\dagger de l'oscillateur harmonique dans la base des fonctions propres de H .
2. En appliquant de façon récursive l'opérateur A^\dagger à ψ_0 , démontrer que la fonction propre ψ_n est bien de la forme donnée par l'équation (1.5).
3. Guider suffisamment le lecteur pour résoudre l'atome d'hydrogène.