

Chapitre 5. Formes quadratiques et matrices symétriques.

1. Formes bilinéaires, formes quadratiques

1.1. Formes bilinéaires et quadratiques

On a déjà rencontré la notion de forme multilinéaire (Chap. 2). Sur un espace vectoriel E , on appelle *forme bilinéaire* réelle une application qui fait correspondre à toute paire de vecteurs $X, Y \in E$ un nombre réel $f(X, Y)$, cette application étant linéaire en X et en Y , donc

$$\begin{aligned} f(\lambda_1 X_1 + \lambda_2 X_2, Y) &= \lambda_1 f(X_1, Y) + \lambda_2 f(X_2, Y) \\ f(X, \mu_1 Y_1 + \mu_2 Y_2) &= \mu_1 f(X, Y_1) + \mu_2 f(X, Y_2) . \end{aligned} \quad (1.1)$$

La forme bilinéaire est dite *symétrique* si $f(X, Y) = f(Y, X)$.

Exemples. Le produit scalaire $\vec{X} \cdot \vec{Y}$ dans l'espace euclidien $\mathbb{R}^n$ est une forme bilinéaire symétrique. La composante sur un axe donné du produit vectoriel $\vec{X} \wedge \vec{Y}$ dans l'espace $\mathbb{R}^3$ est une forme bilinéaire, mais pas symétrique (elle est en fait antisymétrique !). Si g et h sont deux fonctions d'une variable réelle, intégrables sur un intervalle (a, b) , $f(g, h) = \int_a^b g(x)h(x) dx$ est une forme bilinéaire symétrique en g et h .

Le premier exemple suggère la définition suivante : Étant donnée une forme bilinéaire symétrique f , on dit que X et Y sont *orthogonaux* pour f si $f(X, Y) = 0$.

Étant donnée la forme bilinéaire $f(X, Y)$, on lui associe une *forme quadratique* par

$$Q(X) = f(X, X) . \quad (1.2)$$

Bien sûr, cette forme quadratique n'est pas linéaire : $Q(\lambda X) = \lambda^2 Q(X)$. Inversement pour toute forme quadratique Q , on peut construire une forme bilinéaire symétrique f telle que $Q(X) = f(X, X)$ par l'opération de *polarisation* : on écrit simplement, grâce à la bilinéarité

$$f(X + Y, X + Y) = f(X, X) + f(X, Y) + f(Y, X) + f(Y, Y) \quad (1.3)$$

et si on fait l'hypothèse que f est symétrique, $f(X, Y) = \frac{1}{2}(f(X + Y, X + Y) - f(X, X) - f(Y, Y)) = \frac{1}{2}(Q(X + Y) - Q(X) - Q(Y))$.

Exemples. Reprenons les deux exemples ci-dessus. Au produit scalaire dans $\mathbb{R}^n$ correspond la forme quadratique $\|\vec{X}\|^2 = \vec{X} \cdot \vec{X}$ qui est la norme carrée (la longueur carrée) du vecteur $\vec{X}$. De même, $\int_a^b f^2(x) dx$ est une norme carrée pour les fonctions (de carré intégrable) sur (a, b) .

Théorème de Pythagore. Soit f une forme bilinéaire symétrique, Q la forme quadratique associée, on a pour toute paire de vecteurs orthogonaux

$$\forall X, Y : f(X, Y) = 0 \implies Q(X + Y) = Q(X) + Q(Y) , \quad (1.4)$$

qui découle de (1.3).

1.2. Formes définies positives

On dit que la forme quadratique Q est *définie positive* si

$$\forall X \neq 0 \in E \quad Q(X) > 0, \quad (1.5)$$

et donc $Q(X) = 0$ si et seulement si $X = 0$. La forme est *semi-définie positive* si l'inégalité n'est pas stricte : $\forall X \neq 0 \in E \quad Q(X) \geq 0$, elle est *indéfinie* si $Q(X)$ peut prendre un signe ou l'autre selon la valeur de X . Par abus de langage on dit d'une forme bilinéaire qu'elle est définie positive, semi-définie positive, etc, si la forme quadratique associée l'est. *Exemples.* Le produit scalaire habituel dans l'espace euclidien $\mathbb{R}^n$ est défini positif, $Q(\vec{X})$ définissant la norme carrée, c'est-à-dire la longueur carrée du vecteur $\vec{X}$. Au contraire, dans l'espace-temps de la Relativité restreinte (espace de Minkowski), la forme quadratique $Q(X) = c^2 t^2 - x_1^2 - x_2^2 - x_3^2$ est indéfinie : les quadrivecteurs de "genre temps" ont une norme carrée positive, ceux de "genre espace" une norme carrée négative, ceux de "genre lumière" une norme nulle. Dans l'espace $\mathbb{R}^2$, la forme quadratique $Q(X) = x_1 x_2$ est indéfinie et la forme $Q'(X) = (x_1 - x_2)^2$ est semi-définie positive, pourquoi ?

Si la forme symétrique f est définie positive, pour toute paire X, Y de vecteurs non colinéaires et tout réel λ , le vecteur $\lambda X + Y$ n'est pas nul, donc $Q(\lambda X + Y) > 0$ est strictement positif. Or

$$Q(\lambda X + Y) = \lambda^2 Q(X) + 2\lambda f(X, Y) + Q(Y) .$$

est un trinôme du second degré en λ , et le fait qu'il est toujours strictement positif implique que son discriminant est négatif, donc

$$\Delta' = f(X, Y)^2 - Q(X)Q(Y) < 0$$

En revanche si X et Y sont colinéaires, il existe un λ_0 tel que $\lambda_0 X + Y = 0$, et alors $Q(\lambda X + Y) \geq 0$ s'annule en λ_0 mais ne change pas de signe, son discriminant est nul. On obtient ainsi l'**inégalité de Schwarz**

$$|f(X, Y)| \leq (Q(X)Q(Y))^{\frac{1}{2}}, \quad (1.6)$$

avec égalité si et seulement si X et Y sont colinéaires.

Exemple : dans l'exemple précédent de l'espace euclidien $\mathbb{R}^3$, cette inégalité nous dit que

$$|\vec{X} \cdot \vec{Y}| \leq \|\vec{X}\| \|\vec{Y}\|$$

ou encore, si on se rappelle la formule de trigonométrie $\cos \theta = \frac{\vec{X} \cdot \vec{Y}}{\|\vec{X}\| \|\vec{Y}\|}$, que $|\cos \theta| \leq 1$, avec égalité ssi $\theta = 0$ ou π donc $\vec{X}$ et $\vec{Y}$ colinéaires. Plus généralement, pour toute forme bilinéaire définie positive, l'inégalité de Schwarz (1.6) nous permet de définir (au signe près et à 2π près) l'angle θ entre deux vecteurs X et Y par $\cos \theta = f(X, Y)/(Q(X)Q(Y))^{\frac{1}{2}}$.

1.3. Représentations matricielles

Supposons que l'on a choisi une base e_i dans l'espace E . Dans cette base, on écrit les vecteurs $X = \sum_i x_i e_i$ et $Y = \sum_i y_i e_i$, donc la forme bilinéaire

$$f(X, Y) = \sum_{ij} x_i y_j f(e_i, e_j) = \sum_{ij} x_i b_{ij} y_j,$$

où la matrice B de la forme bilinéaire (dans la base choisie e_i) est définie par

$$B = (b_{ij}) \quad b_{ij} = f(e_i, e_j). \quad (1.7)$$

Cette matrice est symétrique, $b_{ij} = b_{ji}$, si la forme bilinéaire l'est. Utilisant la même notation X et Y pour les matrices colonnes des composantes de X et Y , on voit que l'on peut écrire

$$f(X, Y) = X^T B Y.$$

Supposons maintenant que l'on effectue un changement de base $e_i \rightarrow e'_j = \sum_i e_i a_{ij}$ (cf Chap 1, (2.4)). Comme on l'a vu au chapitre 1, les composantes X et X' d'un vecteur donné dans l'ancienne et la nouvelle base sont reliées par $X = A X'$ (Chap 1, (2.5)). Par conséquent la forme bilinéaire s'exprime maintenant selon $f(X, Y) = X^T B Y = X'^T A^T B A Y'$ donc à l'aide de la matrice $B' = A^T B A$ (et non pas selon $A^{-1} B A$ comme pour une application linéaire, comparer avec Chap 1, (4.4) !)

2. Réduction d'une forme quadratique

Dans toute cette section on supposera que les formes bilinéaires et les matrices associées sont symétriques.

2.1. Vecteurs orthogonaux, vecteurs orthonormés

Définition : Si f est une forme bilinéaire symétrique définie positive, on dit que des vecteurs $X_1, \dots, X_k$ sont *orthonormés* (pour la forme f) si

$$f(X_i, X_j) = \delta_{ij}$$

autrement dit si ces vecteurs sont deux à deux orthogonaux : $f(X_i, X_j) = 0$ si $X_i \neq X_j$ et s'ils sont normés $Q(X_i) = 1$.

Lemme 1 : Si les vecteurs $X_1, \dots, X_k$ sont orthonormés (pour la forme f), ils sont nécessairement linéairement indépendants.

La preuve (élémentaire !) est laissée en exercice.

2.2. Procédé d'orthonormalisation de Schmidt

Soit f une forme bilinéaire symétrique définie positive.

Théorème 1 : À partir de tout système de k vecteurs linéairement indépendants $X_1, \dots, X_k$, on peut construire un système de k vecteurs orthonormés $\tilde{X}_1, \dots, \tilde{X}_k$, combinaisons linéaires des $X_1, \dots, X_k$.

Preuve par récurrence sur k . Pour $k = 1$, on dispose d'un vecteur X_1 non nul, donc de norme non nulle. Le vecteur $\tilde{X}_1 = X_1/Q(X_1)^{\frac{1}{2}}$ est bien normé. Supposons alors la propriété vraie pour tout système de $k - 1$ vecteurs, et considérons le système de k vecteurs linéairement indépendants $X_1, \dots, X_k$. Le sous-système $X_1, \dots, X_{k-1}$ remplit la condition de récurrence, on peut donc construire un système de $k - 1$ vecteurs orthonormés $\tilde{X}_1, \dots, \tilde{X}_{k-1}$, combinaisons linéaires des $X_1, \dots, X_{k-1}$. Le k -ième vecteur X_k est indépendant de $X_1, \dots, X_{k-1}$ donc aussi de $\tilde{X}_1, \dots, \tilde{X}_{k-1}$. Cherchons un $Y = X_k + \sum_{i=1}^{k-1} \lambda_i \tilde{X}_i$ orthogonal à $\tilde{X}_1, \dots, \tilde{X}_{k-1}$: en prenant le produit scalaire par f entre cet Y et les autres : $f(Y, \tilde{X}_i) = f(X_k, \tilde{X}_i) + \lambda_i$, on détermine $\lambda_i = -f(X_k, \tilde{X}_i)$. Finalement ce vecteur Y étant non nul (sans quoi X_k ne serait pas linéairement indépendant des $\tilde{X}_1, \dots, \tilde{X}_{k-1}$), il suffit de le normer pour obtenir $\tilde{X}_k = Y/f(Y, Y)^{\frac{1}{2}}$ et terminer la preuve par récurrence.

Ce théorème a comme corollaire que l'on peut toujours trouver une base orthonormale dans l'espace vectoriel E .

Bien comprendre que ce théorème, sous l'hypothèse de l'existence d'une forme bilinéaire définie positive, nous ramène sur le terrain bien connu de la géométrie euclidienne. Dans la base orthonormée, la forme bilinéaire prend l'allure familière du produit scalaire

“en coordonnées rectangulaires”, $f(X, Y) = \sum_i x_i y_i$, et la norme carrée $Q(X) = \sum_i x_i^2$. Un espace vectoriel doté d’une forme bilinéaire définie positive est appelé *espace euclidien*.

Exemple. Considérons l’espace E des polynômes de degré $\leq n$ dans la variable x et définissons la forme bilinéaire $f(p, q) = \int_{-1}^1 p(x)q(x) dx$. Cette forme est évidemment symétrique et définie positive. À partir de la base naturelle $\{1, x, x^2, \dots, x^n\}$ de l’espace E , on peut, grâce au procédé d’orthonormalisation de Schmidt, construire une base orthonormée $p_k(x)$. Ce sont les polynômes $p_k(x) = (k + \frac{1}{2})^{\frac{1}{2}} P_k(x)$, avec P_k les “polynômes de Legendre” $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = \frac{1}{2}(3x^2 - 1)$, etc. (Voir TP1.) Ces polynômes seront rencontrés par la suite dans le cours de mécanique quantique, où ils jouent un rôle important dans la description du moment angulaire.

2.3. Matrices orthogonales

Considérons un espace E doté d’une forme bilinéaire définie positive, donc euclidien. On notera dans la suite $X.Y = f(X, Y)$ et $\|X\|^2 = Q(X)$. Soient e_i une base orthonormée, x_i, y_i les composantes de deux vecteurs X et Y dans cette base : $X = \sum_i x_i e_i, Y = \sum_i y_i e_i$.

Le produit scalaire et la norme carrée y prennent donc les expressions familières

$$X.Y = \sum_i x_i y_i \quad \|X\|^2 = \sum_i x_i^2, \quad (2.1)$$

et les composantes x_i, y_i s’expriment en termes de produits scalaires avec les vecteurs de base

$$x_i = e_i.X, \quad y_i = e_i.Y. \quad (2.2)$$

En termes des vecteurs colonnes des composantes de X et Y

$$X.Y = X^T Y = Y^T X \quad \|X\|^2 = X^T X. \quad (2.3)$$

Définition : On appelle *matrice orthogonale* toute matrice carrée $n \times n$ telle que

$$O^T O = I \iff O^{-1} = O^T \iff O O^T = I, \quad (2.4)$$

En écrivant explicitement ces conditions d’orthonormalité, on obtient

$$\sum_k O_{ki} O_{kj} = \delta_{ij} \quad \sum_\ell O_{i\ell} O_{j\ell} = \delta_{ij} \quad (2.5)$$

qui expriment que les vecteurs colonnes d’une part, les vecteurs lignes de l’autre, de la matrice O sont orthonormés.

Exemples. Vérifier que les matrices $\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ et $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} \end{pmatrix}$ sont orthogonales. Plutôt que d’écrire le produit matriciel $O.O^T$, il suffit de calculer (mentalement

ou avec un crayon !) le produit scalaire de chaque colonne (ou ligne) avec elle-même et avec les autres.

Proposition 1 : *Les transformations par des matrices orthogonales laissent invariant le produit scalaire.*

En effet si $X' = OX$, $Y' = OY$, alors $X'^T Y' = X^T O^T O Y = X^T Y$. En particulier, si X et Y sont deux vecteurs orthogonaux, leurs transformés par une matrice orthogonale $X' = OX$, $Y' = OY$ le sont aussi : $X^T Y = 0 \implies X'^T Y' = 0$.

Dans l'espace euclidien à 3 dimensions, les transformations définies par ces matrices sont telles que $\overrightarrow{OM} = X \mapsto \overrightarrow{OM'} = X' = OX$ avec $|OM| = |OM'|$, ce sont des *isométries* et on démontre que ce sont des rotations ou des réflexions par rapport à un plan, ou leurs composées (leurs produits). Le produit de deux matrices orthogonales est une matrice orthogonale (le vérifier !), ce qui signifie que la composée de deux isométries est une isométrie. La transformation $X \mapsto -X$ qui correspond à la matrice $O = -I$ est le produit d'une rotation de π autour d'un axe quelconque Δ passant par l'origine O par la réflexion dans le plan orthogonal en O à Δ . *Exercice :* pour s'en convaincre, (1) faire la figure, (2) écrire les deux matrices qui effectuent ces transformations.

Proposition 2 : *Tout changement de base orthonormée définit une matrice orthogonale. Réciproquement, toute matrice orthogonale transforme une base orthonormée en une autre base orthonormée.*

Preuve : soient e_i et f_i deux bases orthonormées, $e_i \cdot e_j = \delta_{ij}$, $f_i \cdot f_j = \delta_{ij}$. Formons la matrice de changement de base, cf. equ. (2.4) du chapitre 1, $f_j = \sum_i e_i \Omega_{ij}$, Soient en utilisant les produits scalaires orthonormés : $\Omega_{ij} = e_i \cdot f_j$. Mais selon l'observation faite en (2.2), Ω_{ij} représente la composante de e_i sur f_j (ou vice versa), et selon (2.5), l'orthonormalité de ces composantes est équivalente à l'orthogonalité de la matrice Ω , cqfd. La réciproque découle de la Proposition 1.

Remarque : en géométrie usuelle dans l'espace euclidien à 3 dimensions, cette proposition dit simplement qu'on passe d'un repère orthonormé à un autre par une rotation ou une "rotation-réflexion". Qu'en est-il de l'espace euclidien à deux dimensions ?

Proposition 3 : *Tout changement de base par une matrice orthogonale transforme une matrice symétrique en matrice symétrique.*

La preuve est immédiate : si $B = B^T$, pour toute matrice orthogonale O , $B' = O^T B O$ satisfait bien $B'^T = (O^T B O)^T = O^T B^T O = O^T B O = B'$ cqfd.

2.4. Diagonalisation d'une matrice symétrique

Supposons qu'on a dans E une autre forme bilinéaire symétrique f . (On ne la suppose pas définie positive.) Elle est définie par la donnée dans la base e_i (orthonormée pour le produit scalaire euclidien) de la matrice symétrique B , $B_{ij} = f(e_i, e_j)$.

On démontre alors l'important théorème

Théorème 2 : *Toute matrice symétrique réelle B peut se diagonaliser par un changement de base orthogonal*

$$\exists O \quad B = O \Lambda O^T, \quad \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n) \quad \lambda_i \in \mathbb{R}. \quad (2.6)$$

Autrement dit, toute matrice symétrique réelle B possède n valeurs propres réelles et un système de n vecteurs propres orthonormés.

Établissons d'abord le

Lemme 2 : *Si B est une matrice symétrique réelle, (i) deux vecteurs propres de B correspondant à deux valeurs propres distinctes sont orthogonaux ; (ii) B n'a que des valeurs propres réelles.*

Les deux propriétés du Lemme découlent du même argument :

(i) Soient λ et μ deux valeurs propres distinctes de B , de vecteurs propres respectifs X et Y . On a donc $BX = \lambda X$, $BY = \mu Y$. Calculons de deux façons $Y^T BX$

$$\begin{aligned} Y^T BX &= \lambda Y^T X \\ &= (Y^T BX)^T = X^T B^T Y = X^T BY = \mu X^T Y = \mu Y^T X, \end{aligned}$$

(où on a utilisé le fait que le nombre $Y^T BX$, considéré comme un matriciel 1×1 , est égal à son transposé, et il en est de même de $Y^T X$). Il en découle que $(\lambda - \mu)Y^T X = 0$, donc puisque $\lambda \neq \mu$, $Y^T X = 0$, les vecteurs propres X et Y sont orthogonaux (pour le produit scalaire euclidien (2.1)).

(ii) Soit λ une valeur propre de B pour le vecteur propre X . Il se pourrait que λ et X soient complexes. Écrivant $BX = \lambda X$ et sa conjuguée (B est réelle) $BX^* = \lambda^* X^*$, on voit qu'on est dans les conditions du point (i) précédent, mais cette fois X et X^* ne peuvent être orthogonaux, car $XX^* = \sum_{i=1}^n |x_i|^2$, somme des modules carrés des composantes de X , qui ne peut s'annuler que si $X = 0$. On en conclut que $\lambda \neq \lambda^*$ est impossible, donc que λ est réelle.

Preuve du Théorème. Démonstration par récurrence sur la dimension n de la matrice (ou de l'espace vectoriel). Pour $n = 1$, le théorème est évident : dans l'espace de dimension 1, le vecteur de base peut être normé. Supposons le théorème vrai pour tout espace de dimension $\leq n - 1$. Dans l'espace E de dimension n , doté d'une base e_i orthonormée pour le produit scalaire euclidien (2.1), B a au moins un vecteur propre X_1 de valeur propre λ_1 , que nous supposons normé. Par le lemme, λ_1 et X_1 sont réels, et $X_1 = \sum_i x_i e_i$. Supposons par exemple que $x_1 \neq 0$. Le système de vecteurs $\{X_1, e_2, \dots, e_n\}$ est une base, et par le procédé d'orthonormalisation de Schmidt, on peut construire une base orthonormée $X_1, \tilde{e}_2, \dots, \tilde{e}_n$. Soit F le sous-espace engendré par les vecteurs $\tilde{e}_2, \dots, \tilde{e}_n$. C'est l'espace orthogonal au vecteur propre X_1 . Montrons que B laisse le sous-espace F invariant. En effet pour tout $j = 2, \dots, n$, on a $X_1 \cdot B\tilde{e}_j = \tilde{e}_j \cdot BX_1 = \lambda_1 \tilde{e}_j \cdot X_1 = 0$, $B\tilde{e}_j$ est orthogonal à X_1 donc appartient au sous-espace F . B est représentée dans F par une matrice symétrique (Proposition 3 du § 2.3). On peut maintenant appliquer à F , espace de dimension $n - 1$, l'hypothèse de récurrence : il possède une base orthonormée de vecteurs propres $X_2, \dots, X_n$ de B . Par la Proposition 2 du § 2.3, on passe de la base initiale e_i à cette nouvelle base $X_1, X_2, \dots, X_n$ par un changement de base orthogonal, cqfd.

2.5. Réduction d'une forme quadratique

Soient Q une forme quadratique, B la matrice réelle symétrique qui la représente dans une base,

$$Q(X) = \sum_{ij} B_{ij} x_i x_j .$$

Selon le théorème 2 du § précédent, on peut trouver un changement de base par une matrice orthogonale qui diagonalise la matrice B : $B = O\Lambda O^T$, où O a pour vecteurs-colonnes les vecteurs propres (orthonormés) de B , cf chapitre 4, § 2.2. Récrivant $Q(X) = X^T B X = X^T O\Lambda O^T X$, on voit que le changement de coordonnées $X = OX'$, soit $x_i = \sum_j O_{ij} x'_j$, diagonalise la forme quadratique Q : $Q = X'^T \Lambda X'$, c'est-à-dire l'exprime comme somme de carrés avec comme coefficients les valeurs propres de B .

Théorème 3 : Pour toute forme quadratique $Q(X) = \sum_{ij} B_{ij} x_i x_j$, il existe un changement de coordonnées orthogonal $x'_i = \sum_j O_{ji} x_j$ qui la diagonalise

$$Q(X) = \sum_{i=1}^n \lambda_i x_i'^2 = \sum_{i=1}^n \lambda_i \left(\sum_j O_{ji} x_j \right)^2 . \quad (2.7)$$

On appelle cette expression la *forme réduite* de Q .

Corollaire 1 : Une forme quadratique est définie positive (resp. semi-définie positive) ssi les valeurs propres de sa matrice sont strictement positives (resp. positives ou nulles). Elle est indéfinie si sa matrice possède des valeurs propres des deux signes.

Exemples. Au vu de ce corollaire, il n'est pas difficile de voir que dans l'espace de dimension 3, la forme $Q_1(X) = x_1^2 + x_2^2 + x_3^2 - x_1 x_2 - x_2 x_3$ est définie positive, que $Q_2(X) = x_1^2 + x_2^2 + x_3^2 - x_1 x_2 - x_2 x_3 - x_1 x_3$ est semi-définie positive, et que $Q_3(X) = x_1^2 - 2x_2 x_3$ est indéfinie. Vérifier que les valeurs propres de leurs matrices sont respectivement $B_1 : \{1 \pm \frac{\sqrt{2}}{2}, 1\}$, $B_2 : \{\frac{3}{2}, \frac{3}{2}, 0\}$ et $A_3 : \{1, 1, -1\}$ et que leurs formes réduites s'écrivent effectivement

$$\begin{aligned} Q_1(X) &= \frac{2+\sqrt{2}}{8} (x_1 - \sqrt{2}x_2 + x_3)^2 + \frac{1}{2} (x_1 - x_3)^2 + \frac{2-\sqrt{2}}{8} (x_1 + \sqrt{2}x_2 + x_3)^2 \\ Q_2(X) &= \frac{3}{2} \left(\frac{x_1 - x_3}{\sqrt{2}} \right)^2 + \frac{3}{2} \left(\frac{-x_1 + 2x_2 - x_3}{\sqrt{6}} \right)^2 \\ Q_3(X) &= x_1^2 + \left(\frac{x_2 - x_3}{\sqrt{2}} \right)^2 - \left(\frac{x_2 + x_3}{\sqrt{2}} \right)^2 . \end{aligned}$$

2.6. Diagonalisation simultanée de deux matrices symétriques commutantes

Le théorème 2 du paragraphe § 2.4 possède encore un autre corollaire intéressant.

Corollaire 2 : Si A et B sont deux matrices réelles symétriques qui commutent, $AB = BA$, on peut les diagonaliser simultanément par un même changement de base orthogonal.

Noter que le théorème 2 en est un cas particulier, quand $A = I$, la matrice identité. Noter aussi que ce résultat prolonge (pour des matrices symétriques) celui obtenu à la Prop. 7 du Chap. 4. Ici on n'a pas besoin de supposer les valeurs propres de l'une des matrices distinctes.

Avant de prouver ce Corollaire, démontrons le

Lemme 3 : Si A et B commutent, un espace propre de A est invariant par B .

En effet si X est un vecteur propre de A , $AX = \lambda X$, et $ABX = BAX = \lambda BX$. Ou bien $BX = 0$, ou bien il est vecteur propre de A . Dans les deux cas, il appartient à l'espace propre.

Preuve du corollaire : on diagonalise d'abord A selon le théorème. B laisse invariant tout sous-espace propre F de A , c'est une matrice symétrique dans F , on peut donc l'y diagonaliser toujours selon le théorème par un nouveau changement de base orthogonal (qui n'affecte pas la diagonalité de A). En procédant ainsi dans chaque sous-espace propre de A , on construit une base orthonormale dans laquelle les deux matrices sont diagonales, cqfd.

3. Extension aux formes sesquilineaires et matrices hermitiennes

Toute la discussion qui précède se généralise à des formes à valeurs complexes dites formes sesquilineaires et aux matrices "hermitiennes" qui leur sont associées. Cette situation est importante pour le physicien en particulier dans les applications à la physique quantique. Nous nous bornerons ici à de brèves indications.

Si E désigne maintenant un espace vectoriel sur les nombres complexes (c'est-à-dire les combinaisons linéaires de vecteurs peuvent se faire avec des nombres complexes), on définit une *forme sesquilineaire* $f(X, Y)$ comme une forme (une application dans les nombres complexes) qui est linéaire dans son second argument Y , mais "antilinéaire" dans le premier X . Ce qui remplace (1.1) est donc

$$\begin{aligned} f(\lambda_1 X_1 + \lambda_2 X_2, Y) &= \lambda_1^* f(X_1, Y) + \lambda_2^* f(X_2, Y) \\ f(X, \mu_1 Y_1 + \mu_2 Y_2) &= \mu_1 f(X, Y_1) + \mu_2 f(X, Y_2) , \end{aligned} \quad (3.1)$$

où λ^* est le complexe conjugué de λ . La forme est dite hermitienne si $f(Y, X) = f(X, Y)^*$.

Exemples. (1) Considérons l'espace $\mathbb{C}^n$ des vecteurs X à n composantes x_i complexes. La forme $\sum_i x_i^* y_i$ est une forme sesquilineaire hermitienne. (2) Soit à nouveau g et h des fonctions d'une variable réelle $x \in (a, b)$, mais cette fois à valeurs complexes (par exemple e^{ix}). L'expression $\int_a^b g(x)^* h(x) dx$ est une forme sesquilineaire hermitienne.

On définit comme précédemment la forme $Q(X) = f(X, X)$ associée à une forme sesquilineaire, qui peut être selon les cas définie positive, semi-définie positive, indéfinie, etc. On a à nouveau une inégalité de Schwarz, qui s'exprime de la même façon. Dans une base, une forme hermitienne s'exprime à l'aide d'une matrice $A = (a_{ij})$, $a_{ij} = f(e_i, e_j)$, et la matrice est hermitienne ce qui signifie que

$$A = A^{T*} \iff a_{ij} = a_{ji}^* . \quad (3.2)$$

Si $X = \sum_i x_i e_i$ et $Y = \sum_j y_j e_j$, $f(X, Y) = \sum_{ij} x_i^* y_j f(e_i, e_j) = \sum_{ij} x_i^* y_j a_{ij}$.

On définit en général la conjugaison hermitique d'une matrice B par

$$B^\dagger = B^{T*} . \quad (3.3)$$

Les matrices hermitiennes sont donc les matrices égales à leur conjuguée hermitique. On démontre que leurs valeurs propres sont toutes réelles, c'est l'analogue du Lemme 2 du § 2.4.

Par ailleurs on définit les matrices *unitaires*. Ce sont les matrices carrées U à éléments complexes qui satisfont

$$U^\dagger U = I \iff U^{-1} = U^\dagger . \quad (3.4)$$

On démontre alors le théorème

Théorème 4 : Toute matrice hermitienne peut se diagonaliser par un changement de base unitaire

$$A = U \Lambda U^\dagger$$

qui implique le

Corollaire 3 : *Toute forme sesquilinéaire peut se mettre sous forme diagonale par un changement de base unitaire,*

$$f(X, Y) = X^\dagger A Y = \tilde{X}^\dagger \Lambda \tilde{Y} = \sum_i \lambda_i \tilde{x}_i^* \tilde{y}_i \quad \tilde{X} = U^\dagger X, \quad \tilde{Y} = U^\dagger Y$$

et la forme est définie positive ssi les valeurs propres sont toutes (réelles) positives.

Finalement le corollaire de la fin du paragraphe précédent admet une extension : si A et B sont deux matrices hermitiennes qui commutent, elles peuvent être diagonalisées simultanément.

Cette dernière propriété joue un rôle fondamental en mécanique quantique : toute quantité observable y est représentée par un opérateur (une matrice, dans une base) hermitien(ne). Ses valeurs propres sont les valeurs que peut donner la mesure de cette observable. Le fait que deux “observables commutent” signifie qu’on peut les mesurer simultanément.

Récapitulons le parallèle entre formes bilinéaires et formes sesquilinéaires	
forme bilinéaire symétrique $f(X, Y)$	$\leftrightarrow$ forme sesquilinéaire hermitienne $f(X, Y)$
forme quadratique $Q(X) = f(X, X)$	$\leftrightarrow$ norme carrée réelle de vecteurs complexes $Q(X) = f(\overline{X}, X)$
matrice symétrique $A = A^T$	$\leftrightarrow$ matrice hermitienne $A = A^\dagger$
matrice orthogonale $O^{-1} = O^T$	$\leftrightarrow$ matrice unitaire $U^{-1} = U^\dagger$
diagonalisation par une matrice orthogonale	$\leftrightarrow$ diagonalisation par une matrice unitaire .

Toutes ces propriétés, étendues éventuellement à des espaces de dimension infinie, seront très utiles en mécanique quantique...

4. Applications physiques

4.1. Tenseur d’inertie

On connaît du cours de Mécanique les notions de moment cinétique $\vec{J}$ et de moment d’inertie I d’un point matériel M de masse m en rotation de vitesse angulaire ω autour d’un axe $\vec{u}$. Si r est la distance du point M à l’axe $\vec{u}$, sa vitesse est $v = r\omega$, (ou mieux $\vec{v} = \omega \vec{u} \wedge \vec{r}$), son moment cinétique (par rapport à $\vec{u}$) est $J_u = mvr = mr^2\omega$, et son moment d’inertie (par rapport à $\vec{u}$) est $I_u = mr^2$, tel que

$$J_u = I_u \omega. \quad (4.1)$$

Ce moment d’inertie (et J_u) s’annule seulement si la masse m est sur l’axe $\vec{u}$. L’énergie cinétique de rotation est $T = \frac{1}{2}mv^2 = \frac{1}{2}I_u\omega^2 = \frac{1}{2}\frac{J_u^2}{I_u}$. Généralisons maintenant ces formules au cas d’un solide en rotation.

Soit un corps solide indéformable dont la distribution de masse est décrite par une densité $\rho(\vec{r})$ dans un certain volume V . Quand ce corps a un mouvement de rotation de vitesse angulaire ω autour d’une direction $\vec{u}$ passant par un point O , le point M de coordonnées $\vec{r} = \overrightarrow{OM}$ a une vitesse $\vec{v} = \omega \vec{u} \wedge \vec{r}$ et le vecteur moment cinétique $\vec{J}$ par rapport à O est la superposition des contributions $\overrightarrow{OM} \wedge \rho \vec{v}$ sommées sur tout le solide :

$$\vec{J} = \int d^3r \rho(\vec{r}) \vec{r} \wedge (\vec{\omega} \wedge \vec{r}). \quad (4.2)$$

C'est donc un vecteur dont les composantes sont des fonctions linéaires de celles de $\vec{\omega} = \omega \vec{u}$, ce qui généralise (4.1) et amène à définir le “tenseur d'inertie” I (ne pas confondre avec la matrice identité !)

$$\vec{J} = I \cdot \vec{\omega}, \quad (4.3)$$

c'est-à-dire $J_i = \sum_j I_{ij} \omega_j$. En utilisant la formule du double produit vectoriel ⁵ $\vec{a} \wedge (\vec{b} \wedge \vec{c}) = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})$, on calcule $J^i = \int d^3r \rho(\vec{r}) \sum_{j=1}^3 (r^2 \delta_{ij} - r^i r^j) \omega^j$, et le tenseur d'inertie I est décrit par la matrice symétrique 3×3

$$I_{ij} = \int d^3r \rho(\vec{r}) (r^2 \delta_{ij} - r^i r^j), \quad (4.4)$$

ou encore,

$$I = \int dx dy dz \rho(x, y, z) \begin{pmatrix} y^2 + z^2 & -xy & -xz \\ -yx & x^2 + z^2 & -yz \\ -zx & -zy & x^2 + y^2 \end{pmatrix} = \begin{pmatrix} A & -F & -E \\ -F & B & -D \\ -E & -D & C \end{pmatrix}. \quad (4.5)$$

(Les notations $A, B, \dots, F$ sont traditionnelles.)

Noter que le moment cinétique n'est en général pas colinéaire au vecteur rotation $\vec{\omega}$. Mais la matrice I_{ij} est symétrique, donc diagonalisable par changement de coordonnées orthogonal, (Théorème 2). Dans les nouveaux axes (OX, OY, OZ) , $I = \text{diag}(I_X, I_Y, I_Z)$. Le long de ces “axes principaux d'inertie”, le moment cinétique est colinéaire au vecteur rotation.

L'énergie cinétique de rotation se calcule aussi aisément. Comme $\vec{v}^2 = (\vec{\omega} \wedge \vec{r})^2 = \vec{\omega} \cdot (\vec{r} \wedge (\vec{\omega} \wedge \vec{r}))$

$$T = \int dx dy dz \frac{1}{2} \rho(x, y, z) \vec{v}^2 = \frac{1}{2} \vec{\omega} \cdot \vec{J} = \frac{1}{2} \vec{\omega} \cdot I \cdot \vec{\omega}.$$

L'énergie cinétique de rotation est donc donnée par la forme quadratique Q associée à I . Cette forme est évidemment définie positive, en tant qu'énergie cinétique. Cela peut se voir aussi d'autres façons : les valeurs propres I_X, I_Y, I_Z sont les “moments d'inertie principaux” par rapport aux axes (OX, OY, OZ) et sont donc positifs. Enfin cette propriété résulte de l'inégalité de Schwarz. En effet, pour tout vecteur $\vec{w}$, $Q(\vec{w}) = \vec{w} \cdot I \cdot \vec{w} = \int d^3r \rho(\vec{r}) (r^2 w^2 - (\vec{r} \cdot \vec{w})^2)$; or $r^2 w^2 - (\vec{r} \cdot \vec{w})^2 \geq 0$ avec égalité seulement si $\vec{r}$ et $\vec{w}$ sont colinéaires, (cf (1.6)), donc la somme $\int d^3r \rho(\vec{r}) (r^2 w^2 - (\vec{r} \cdot \vec{w})^2)$ avec $\rho \geq 0$ est aussi ≥ 0 (et ne s'annulerait que pour une distribution linéaire de masses le long de la direction $\vec{w}$).

⁵ que l'on vérifie par le calcul direct : $\vec{b} \wedge \vec{c} = (b_y c_z - b_z c_y, \dots)$, $\vec{a} \wedge (\vec{b} \wedge \vec{c}) = (a_y (b_x c_y - b_y c_x) - a_z (b_z c_x - b_x c_z), \dots) = (b_x (a_x c_x + a_y c_y + a_z c_z) - c_x (a_x b_x + a_y b_y + a_z b_z), \dots)$

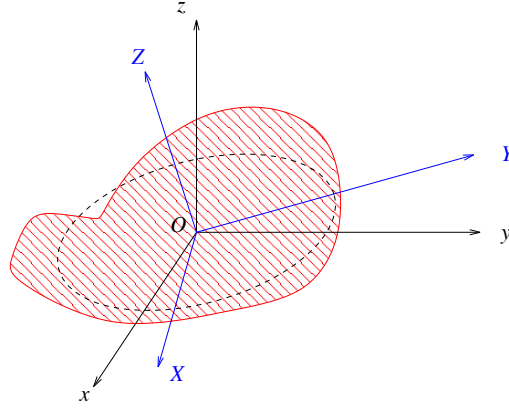


Fig. 13: Ellipsoïde d'inertie en O (ligne brisée) et axes principaux d'inertie.

Soient α, β, γ les coordonnées du vecteur $\vec{w}$. La forme quadratique s'écrit $Q(\vec{w}) = A\alpha^2 + B\beta^2 + C\gamma^2 - 2D\beta\gamma - 2E\gamma\alpha - 2F\alpha\beta$ et l'équation $Q(\vec{w}) = 1$ définit une surface dans l'espace à trois dimensions. C'est un ellipsoïde, dit *ellipsoïde d'inertie au point O* . Les axes principaux d'inertie sont les axes de cet ellipsoïde, voir figure 12.

Toutes ces notions sont importantes dans l'étude de la rotation du corps considéré et de la stabilité de ce mouvement de rotation.

Comment ce tenseur I est-il modifié si on change le point de référence $O \mapsto O'$? En particulier si on choisit le centre de gravité G ? Se rappeler que par définition de G , si $\vec{r}' = \overrightarrow{GM}$, $\int d^3r' \rho(\vec{r}') \vec{r}' = 0$, tandis que $\int d^3r' \rho(\vec{r}') = M$, la masse du corps. Réponse : soit $\vec{g} = \overrightarrow{OG}$, alors $I_O = I_G + M(g^2 \delta_{ij} - g_i g_j)$.

Soit I_G le tenseur d'inertie par rapport à son centre de gravité G dont la vitesse (instantanée) est $\vec{V}$. L'énergie cinétique totale s'écrit $T = \int d^3r \frac{1}{2} \rho(\vec{r}) \vec{v}^2$ mais $\vec{v}(\vec{r}) = \vec{V} + \vec{\omega} \wedge \vec{r}$ donc $\vec{v}^2 = V^2 + 2\vec{V} \cdot (\vec{\omega} \wedge \vec{r}) + (\vec{\omega} \wedge \vec{r})^2$. Le terme "croisé" $\vec{V} \cdot (\vec{\omega} \wedge \vec{r})$ s'annule par intégration sur $\vec{r}$, (puisque à nouveau, pour le centre de gravité, $\int d^3r \rho(\vec{r}) \vec{r} = 0$), et $(\vec{\omega} \wedge \vec{r})^2 = \vec{\omega} \cdot (\vec{r} \wedge (\vec{\omega} \wedge \vec{r}))$, comme plus haut, d'où $\int d^3r \rho(\vec{r}) (\vec{\omega} \wedge \vec{r})^2 = \vec{\omega} \cdot I_G \cdot \vec{\omega}$. Finalement l'énergie est la somme de l'énergie cinétique de translation du centre de gravité et de l'énergie cinétique de rotation autour de G

$$T = \frac{1}{2} M \vec{V}^2 + \frac{1}{2} \vec{\omega} \cdot I_G \cdot \vec{\omega}. \quad (4.6)$$

4.2. Tenseur de conductivité

Dans un conducteur à trois dimensions (et non plus nécessairement filaire), la loi de proportionnalité des intensités aux différences de potentiel (loi d'Ohm), ou des densités de courant $\vec{j}$ au champ électrique $\vec{E}$, prend une forme matricielle $\vec{j} = \sigma \vec{E}$, soit en composantes

$$\begin{pmatrix} j_1 \\ j_2 \\ j_3 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix}.$$

On admettra que le tenseur σ est symétrique $\sigma_{ij} = \sigma_{ji}$. Selon le théorème 2 de diagonalisation, il existe un repère orthonormé de coordonnées tel que σ soit diagonal. Dans ce repère, les nouvelles composantes de $\vec{j}$ sont proportionnelles à celles de $\vec{E}$. On retrouve la loi d'Ohm habituelle, mais avec en général des conductivités différentes dans les différentes directions !

4.3. Stabilité des extrema d'un potentiel

Considérons un système mécanique à N degrés de liberté. Ce peut être un système de n particules massives en interaction dans l'espace $\mathbb{R}^3$, chacune dotées de 3 coordonnées x, y, z , auquel cas le nombre de degrés de liberté est $N = 3n$; ce peut être un solide indéformable dont la position dans l'espace $\mathbb{R}^3$ est repérée par les trois coordonnées de son centre de gravité G et par trois angles de rotation autour de G , d'où $N = 6$, etc. On appellera x_i , $i = 1, \dots, N$ les coordonnées de ces N degrés de liberté.

Supposons la dynamique de ce système décrite par un potentiel $V(x_1, \dots, x_N)$. Cela implique que pour chaque degré de liberté, l'équation de la dynamique s'écrit $m_i \ddot{x}_i = -\frac{\partial V(x)}{\partial x_i}$. Les points stationnaires du système sont les extrema $x^{(0)}$ du potentiel (force nulle)

$$\left. \frac{\partial V(x)}{\partial x_i} \right|_{x^{(0)}} = 0. \quad (4.7)$$

La question est de savoir lesquels de ces extrema correspondent à des positions d'équilibre stable. Pour cela on effectue un développement limité au deuxième ordre de $V(x)$ au voisinage de l'un des $x^{(0)}$ en posant $x_i - x_i^{(0)} = \xi_i$. Selon la formule de Taylor à plusieurs variables

$$V(x) = V(x^{(0)}) + \frac{1}{2} \sum_{i=1}^N \xi_i^2 \left. \frac{\partial^2 V(x)}{\partial x_i^2} \right|_{x^{(0)}} + \sum_{1 \leq i < j \leq N} \xi_i \xi_j \left. \frac{\partial^2 V(x)}{\partial x_i \partial x_j} \right|_{x^{(0)}} + \dots \quad (4.8)$$

où le terme d'ordre 1 est absent en vertu de (4.7). Au terme constant $V(x^{(0)})$ près, qui ne joue aucun rôle dans la discussion, on voit que dans cette approximation de petits écarts ξ à l'extremum $x^{(0)}$, le potentiel est une forme quadratique dans les N variables ξ_i . Écrivons sa forme réduite (au sens du § 2.5)

$$V(x) = V(x^{(0)}) + \frac{1}{2} \sum_{i=1}^N \lambda_i \tilde{\xi}_i^2 \quad (4.9)$$

avec λ_i les valeurs propres de B et $\tilde{\xi}_i$ les modes propres. La question du signe des λ_i va être cruciale :

– s'ils sont tous strictement positifs, c'est-à-dire si la forme quadratique est définie positive, (cf Théorème 3 et Corollaire), l'extremum $x^{(0)}$ est un *minimum local* : au voisinage de $x^{(0)}$, $V(x)$ est strictement supérieur à $V(x^{(0)})$. Posons $B_{ij} = \left. \frac{\partial^2 V(x)}{\partial x_i^2} \right|_{x^{(0)}}$: les ξ_i satisfont $\ddot{\xi}_i + \sum_j B_{ij} \xi_j = 0$, soit des équations d'oscillateurs harmoniques couplés, comme

le montre mieux encore le passage aux modes propres $\tilde{\xi}_i$: $\ddot{\tilde{\xi}}_i + \omega_i^2 \tilde{\xi}_i = 0$, où $\lambda_i = \omega_i^2$ est la valeur propre de B correspondant au mode $\tilde{\xi}_i$. Le système effectue donc de petites oscillations au voisinage de $x^{(0)}$, l'équilibre est **stable** ;

– si par contre la forme possède une ou plusieurs valeurs propres négatives ou nulles, le système est instable. Pour chaque $\tilde{\xi}_i$ de valeur propre $\lambda_i \leq 0$, l'équation de la dynamique est $\ddot{\tilde{\xi}}_i = \lambda_i \tilde{\xi}_i$, avec une force de rappel nulle ou négative : il n'y a plus d'oscillation, il y a même croissance de $\tilde{\xi}_i$ si $\lambda_i < 0$. L'extremum est un point d'équilibre **instable**.

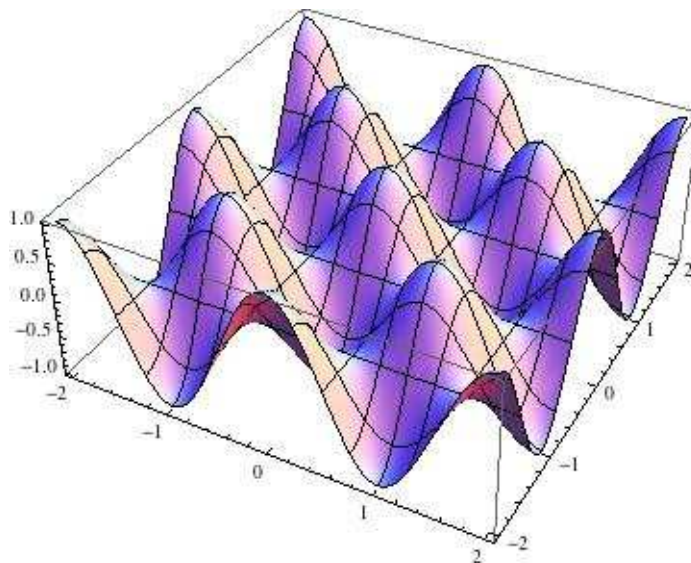


Fig. 14: Le potentiel $V(x, y) = \cos(\pi x) \cos(\pi y)$ avec ses extrema périodiques

Exemple. Ceci est illustré sur un système à deux degrés de liberté x et y , de potentiel $V(x, y) = \cos(\pi x) \cos(\pi y)$. Ce potentiel est représenté sur la figure 14. On lui donne le nom de “boîte à œufs” pour des raisons évidentes. . . . Il est aisé de voir sur la figure et le calcul confirme que le potentiel présente des extrema aux points x, y tous deux entiers ou tous deux demi-entiers. Seuls les points à x et y entiers de parité opposée ($x = 0, y = 1$ par exemple) sont des minima du potentiel, les points où x et y sont entiers de même parité sont des maxima, les points où ils sont tous deux demi-entiers sont des “cols”, avec une valeur propre positive et une négative. Tout cela est bien en accord avec notre intuition : une petite bille lâchée dans ce potentiel oscille au fond d’un creux. . .

On verra en TD un autre exemple.