CAHIERS D'ENSEIGNEMENT

Principes variationnels

pour des systèmes de dimension finie

Michel Karatchentzeff

VERSION 4.0 — JUIN 2004
FASCICULE VII
TAP 6

Ce fascicule correspond à une partie d'un cours donné de 1984 à 2004 à l'université Pierre et Marie Curie de Paris dans le cadre de la maîtrise de Physique et Applications.

Éditions successives

- version 3.1 année 1997
- version 3.2 année 1999
- version 3.3 année 2001
- version 4.0 année 2004

CHAPITRE 6

Principes variationnels pour des systèmes de dimension finie

On sait qu'un système oscillant peut être caractérisé par les valeurs propres ν_i et les vecteurs propres u_i d'un opérateur H défini positif et hermitien pour un produit scalaire construit sur l'espace de configuration E_N des variables physiques à partir des données du problème.

Le but de ce chapitre est de mettre en évidence ce que l'on peut obtenir en s'affranchissant graduellement de la théorie classique des éléments propres (valeurs propres et vecteurs propres) en suivant le fil conducteur de l'interprétation géométrique de ceux-ci.

6.1. Caractérisation variationnelle récursive des éléments propres.

On cherche une caractérisation "variationnelle" des éléments propres, c'est à dire décrite comme étant le maximum ou le minimum d'une certaine entité réelle. Ce type de problème est appelé "problème variationnel" et son nom provient du *Calcul des Variations* où de tels problèmes sont typiques.

Rappel et notations.

Soit E_N un espace vectoriel réel de dimension N (représentant par exemple l'espace de configuration d'un système dynamique et dans lequel évoluent les paramètres du système), et soit H, un opérateur symétrique (hermitien) sur E_N, défini positif, de vecteurs propres $(u_i\ ,\ i = 1, \ldots, N)$ et de valeurs propres correspondantes ν_i.

Pour simplifier les démonstrations, nous supposerons que les valeurs propres se suivent dans un ordre strictement croissant, c'est à dire que

$$\nu_1 < \nu_2 < \cdots < \nu_N$$

Elles seront alors de plus strictement positives[1] $(0 < \nu_1)$.

Nous définirons par $(e_i = u_i/\|u_i\| \quad , \quad i = 1, \ldots, N)$ la base orthonormée construite à partir des u_i, un vecteur quelconque u de E_N se développera suivant cette base selon

$$u = \sum_i x_i e_i$$

Dans cette base, la forme quadratique associée à H prend la forme particulièrement simple[2]

[1] L'opérateur H étant symétrique (hermitien), ses valeurs propres sont réelles ; comme H est aussi défini positif, elles sont strictement positives.

[2] Comme E_N est de dimension N, par souci de simplification, nous omettrons systématiquement les bornes des sommations sur les indices muets.

$$\begin{aligned}(Hu,u) &= \textstyle\sum_{i,j} x_i x_j (He_i, e_j) && \text{(par bilinéarité)}\\ &= \textstyle\sum_{i,j} \nu_i x_i x_j (e_i, e_j) && \text{(les } e_i \text{ sont vecteurs propres)}\\ &= \textstyle\sum_i \nu_i x_i^2 && \text{(les } e_i \text{ sont orthonormés)}\end{aligned}$$

Sous cette forme, il devient évident que la surface

$$(Hu, u) = 1$$

est un ellipsoïde dont l'équation dans le repère des vecteurs propres s'écrit

$$\sum_i \frac{x_i^2}{\frac{1}{\nu_i}} = 1$$

Les longueurs des "axes" de l'ellipsoïde sont donc données par $\sqrt{1/\nu_i}$, ce qui fait en particulier qu'au plus grand axe est associée la plus petite valeur propre. La direction de chaque axe est parallèle à celles des vecteurs propres correspondants.

Introduction géométrique de la caractérisation récursive.

Cas de la plus petite valeur propre :

Soit maintenant l'un des deux points de l'ellipsoïde[3] par lequel passe son grand axe. Ce point définit *le vecteur propre* u_1 dont le module est par définition la longueur du grand axe, soit $1/\sqrt{\nu_1}$. On peut donc écrire

$$\frac{1}{\nu_1} = (u_1, u_1)$$

Du point de vue de la géométrie, (u_1, u_1) correspond au maximum du carré de la longueur du rayon vecteur lorsqu'on parcourt l'ellipsoïde ; on a donc

$$\begin{aligned}\frac{1}{\nu_1} = (u_1, u_1) &= \max_{(Hu,u)=1} (u, u)\\ &\equiv \left\{ \max_u (u, u) \;\middle|\; (Hu, u) = 1 \right\}\end{aligned}$$

en employant une notation plus longue, mais plus lisible.

Remarque :

On peut également écrire

$$\frac{1}{\nu_1} = \max_{u \neq 0} \frac{(u, u)}{(Hu, u)}$$

[3] Rappelons qu'il y en a deux, symétriques par rapport au centre de l'ellipsoïde.

En effet, les deux définitions coïncident évidemment sur $(Hu,u) = 1$. Par ailleurs le rapport

$$\frac{(u,u)}{(Hu,u)}$$

conserve la même valeur si u est remplacé par ku du fait de la bilinéarité du produit scalaire et de la linéarité de H[4]. Nous pouvons toujours trouver un k qui "ramène" (Hu,u) à 1. Il suffit pour cela de poser :

$$\rho^2 \equiv (Hu,u) \neq 1 \quad \text{et de définir} \quad u = \rho w$$

Par construction, puisque

$$(Hu,u) = \rho^2 (Hw,w) \quad w \text{ sera tel que} \quad (Hw,w) = 1$$

ce qui achève de montrer l'équivalence des deux définitions.

Définition géométrique de la plus petite valeur propre :

Ceci étant, associer au petit axe (ou à la plus petite valeur propre) le maximum d'une expression risque d'être source de confusion ; il est plus pratique d'utiliser la relation[5] :

$$\max f(x) = \frac{1}{\min \dfrac{1}{f(x)}}$$

qui transforme les deux définitions précédentes respectivement en :

$$\begin{aligned} \nu_1 &= \min_{u \neq 0} \frac{(Hu,u)}{(u,u)} \\ &= \min \left\{ (Hu,u) \;\middle|\; (u,u) = 1 \right\} \end{aligned}$$

formes sous lesquelles il est plus facile de les retenir.

Cas de la plus grande valeur propre :

Si nous nous étions intéressés au plus petit axe de l'ellipsoïde, c'est la plus grande valeur propre que nous aurions trouvée avec

$$\begin{aligned} \nu_N &= \max_{u \neq 0} \frac{(Hu,u)}{(u,u)} \\ &= \max \left\{ (Hu,u) \;\middle|\; (u,u) = 1 \right\} \end{aligned}$$

[4] Autrement dit, $(u,u)/(Hu,u)$ prend toujours la même valeur sur un rayon vecteur.
[5] Exercice : la démontrer.

Exercice 1. :

H étant un opérateur symétrique défini-positif sur $\mathbb{R}^N$, sa plus grande valeur propre est définie par le maximum de (Hu, u) lorsque u parcourt la sphère unité $(u, u) = 1$ de $\mathbb{R}^N$ et les vecteurs propres correspondants sont ceux qui permettent d'atteindre ce maximum.

La base canonique étant choisie dans $\mathbb{R}^2$, on suppose que H et u s'écrient sous la forme :

$$H = \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} \; ; \quad u = \begin{pmatrix} x \\ y \end{pmatrix}$$

1. Écrire en utilisant les multiplicateurs de Lagrange les équations correspondant à la définition ci-dessus pour calculer les éléments propres correspondant à la plus grande valeur propre.

2. Les résoudre (expliquer pourquoi le résultat trouvé correspond à la plus grande valeur propre).

3. On pourra également profiter de cet exercice pour tester la méthode de déflation (fascicule viiiTAP 7) pour calculer l'élément propre suivant. Poser les équations, les justifier, les résoudre.

La définition précédente relie également ν_N à la norme de la forme bilinéaire associée à la forme quadratique issue de H (voir cours premier cycle). Il est intéressant de réfléchir sur l'exercice suivant :

Exercice 2. :

On considère la forme quadratique définie par

$$g(x, y) = 8x^2 + 4xy + 5y^2$$

1. Montrer que l'on peut toujours trouver un nombre α réel tel que

$$8x^2 + 4xy + 5y^2 \leq \alpha(x^2 + y^2)$$

2. Vérifier que l'on a toujours

$$8x^2 + 4xy + 5y^2 \leq 9(x^2 + y^2)$$

et donner, s'il existe, le domaine où l'inégalité se transforme en égalité.

3. Quelle est dans la base canonique la matrice M symétrique associée à la forme quadratique g ?

4. Quelle propriété de la matrice M peut-on déduire des questions précédentes ?

Le quotient de Rayleigh.

La quantité

$$\frac{(Hu, u)}{(u, u)}$$

a une grande importance. Elle est appelée le *quotient de Rayleigh*.[6] Nous avons vu que ce quotient est relié à l'énergie cinétique et l'énergie potentielle du système dans le cas des systèmes mécaniques. Le fait que la plus petite (ou la plus grande) valeur propre d'un système soit calculable à partir de ce quotient est souvent dénommé *Principe de Rayleigh*.

Les valeurs propres suivantes...

Géométriquement, le deuxième axe de l'ellipsoïde est défini par l'un des points qui se trouvent

- d'une part à la plus grande distance de l'origine.
- d'autre part tel que le vecteur ainsi defini soit orthogonal à u_1.

Nous devons donc avoir pour les mêmes raisons

$$\begin{aligned}\frac{1}{\nu_2} &= \max_u \Big\{ (u, u) \;\Big|\; (Hu, u) = 1 \;,\; (u, u_1) = 0 \Big\} \\ &= \max_u \left\{ \frac{(u, u)}{(Hu, u)} \;\Big|\; (u, u_1) = 0 \right\}\end{aligned}$$

ou encore

$$\nu_2 = \min_u \left\{ \frac{(Hu, u)}{(u, u)} \;\Big|\; u \neq 0 \;,\; (u, u_1) = 0 \right\}$$

La récursivité est trouvée, on aura de la même manière

$$\nu_3 = \min_u \left\{ \frac{(Hu, u)}{(u, u)} \;\Big|\; u \neq 0 \;,\; (u, u_1) = 0 \;,\; (u, u_2) = 0 \right\}$$

etc.

Justification de la caractérisation variationnelle récursive.

Nous n'avons pour l'instant qu'une image géométrique de la caractérisation variationnelle récursive des éléments propres. Il reste donc à montrer que cette image a un sens mathématique, c'est à dire que le maximum ou le minimum de la quantité

$$\frac{(Hu, u)}{(u, u)}$$

[6] Lord Rayleigh (1842–1918) a été l'un des premiers à introduire dans les théories physiques, en particulier en acoustique, les notions de vecteurs propres et de valeurs propres d'opérateur. L'un de ses livres, *The Theory of Sound* n'arrive toujours pas à vieillir.

existe et que nous pourrons en déduire les éléments propres de H. Formulé ainsi, il s'agit donc bien d'un problème variationnel.

Pour cela, nous devons démontrer avant toute autre chose qu'il existe des solutions extrémales de ce problème, autrement dit, que la fonction réelle de la variable u passe par des *extrema* lorsque u varie dans E_N et qu'il existe au moins, pour chaque extremum, une valeur de u qui permette de l'atteindre. Nous choisirons systématiquement de faire les démonstrations à partir de la plus petite valeur propre ; la transposition au cas de la plus grande, étant évidente, est laissée aux soins du lecteur.

La démonstration se fait en deux temps : il suffit de

1) se placer sur la boule unité et de montrer qu'il existe au moins un maximum et au moins un minimum.

2) puis de montrer qu'en dehors de la boule unité, il ne peut y avoir d'extremum plus petit ou plus grand.

Démonstration :

1) E_N étant de dimension finie, la boule unité est un fermé borné. L'application

$$u \longrightarrow (Hu, u) : \quad \Big\{ u \in E_N \;\Big|\; \|u\| = 1 \Big\} \longrightarrow \mathbf{R}$$

est continue car elle est composée d'applications continues. D'après le théorème de Bolzano–Weierstrass, toute application numérique continue sur un compact atteint ses bornes sur le compact. Donc

$$\begin{array}{llll} \text{il existe} \quad u_1^m & \text{tel que} \quad (Hu_1^m, u_1^m) \le (Hu, u) & \text{pour tout } u \text{ tel que} & \|u\| = 1 \\ \text{il existe} \quad u_N^M & \text{tel que} \quad (Hu_N^M, u_N^M) \ge (Hu, u) & \text{pour tout } u \text{ tel que} & \|u\| = 1 \end{array}$$

2) Soit v tel que $\|v\| \neq 1$. Posons $\rho^2 = \|v\|^2 = (v, v)$ et soit w tel que $v = \rho w$ tel que $(v, v) = \rho^2 (w, w) = \rho^2$ et donc $(w, w) = 1$. Rappelons qu'alors

$$\frac{(Hv, v)}{(v, v)} = \frac{(Hw, w)}{(w, w)}$$

Autrement dit, la valeur du quotient de Rayleigh pour un vecteur quelconque sera la même que pour un vecteur unitaire construit sur ce vecteur. Il n'existe donc pas de vecteur de norme différente de 1 fournissant par exemple un quotient de Rayleigh plus petit.

En effet, supposons qu'il en existe un : v. On a donc

$$(v, v) \neq 1 \quad \text{et} \quad \frac{(Hv, v)}{(v, v)} \equiv \mu < \nu_1 = \min\{(Hu, u) | (u, u) = 1\}$$

Soit w un vecteur associé à v et de norme 1. On aurait donc

$$\frac{(Hv, v)}{(v, v)} = \frac{(Hw, w)}{(w, w)} = \mu < \nu_1$$

Il existerait donc un vecteur w de norme 1 donnant au quotient de Rayleigh une valeur plus petite que ν_1, ce qui est contraire à l'hypothèse.

Une démonstration analogue s'applique pour le maximum.

En conclusion :

Pour $u \in E_N$, le quotient de Rayleigh passe par un minimum et il existe au moins une valeur de u pour laquelle ce minimum est atteint. De plus tous les éléments de l'espace vectoriel engendré par u permettent d'atteindre ce minimum.

(Conclusion analogue pour le maximum).

Application aux éléments propres suivants.

Pouvons-nous étendre la conclusion précédente aux autres quantités définies géométriquement ? La réponse est oui.

Soit $\{u_1\}$ des vecteurs minimisant $(Hu, u)/(u, u)$. Nous voulons démontrer que

$$\left\{ \frac{(Hu, u)}{(u, u)} \;\middle|\; u \neq 0 \ , \ (u, u_1) = 0 \right\}$$

possède un plus petit et un plus grand élément. La démonstration se fait toujours en deux étapes.

1) on se restreint à

$$\{(Hu, u) | (u, u) = 1 \ , \ (u, u_1) = 0\}$$

L'ensemble de définition est un sous-ensemble du précédent ; c'est l'intersection d'un compact et d'un fermé. C'est donc un fermé dans un compact. C'est donc un compact.

(Hu, u) n'a pas été modifié et donc atteint ses bornes sur le compact.

2) La démonstration est identique.

Ce résultat s'étend donc, de proche en proche (c'est la récursivité) pour tous les quotients de Rayleigh dont on a donné précédemment une définition géométrique.

Note historique :

La première caractérisation récursive des éléments propres est due à H. Weber (Math. Ann. **1**, 1869, p.1–36) qu'il ne faut pas confondre avec

- le physicien allemand Wilhelm Eduard Weber (1804–1891) fort connu pour ses travaux sur le magnétisme et dont le nom a été donné à une unité de flux magnétique.
- le physicien Ernst Weber (1901–) spécialiste des ondes micrométriques.
- le compositeur Karl Maria von Weber (1786–1826).

Lien entre la caractérisation variationnelle des quotients de Rayleigh et les élément propres de H.

Il reste à montrer que les ν_r et u_r définis de façon variationnelle sont en fait "identiques" aux éléments propres de H définis par

$$Hu_r = \nu_r u_r$$

Commençons par la plus petite valeur propre. On a le
Théorème : *Soit*

$$\nu_1 = \frac{(Hu_1, u_1)}{(u_1, u_1)} = \min\left\{\frac{(Hu, u)}{(u, u)} \;\middle|\; u \neq 0\right\}$$

Alors ν_1 et u_1 sont éléments propres de H, autrement dit

$$Hu_1 = \nu_1 u_1$$

Démonstration : Les hypothèses s'écrivent

$$(Hu_1, u_1) - \nu_1(u_1, u_1) = 0 \quad , \quad (Hu, u) - \nu_1(u, u) \geq 0$$

soit encore

$$\big(\big[H - \nu_1 I\big]u_1, u_1\big) = 0 \quad , \quad \big(\big[H - \nu_1 I\big]u, u\big) \geq 0$$

Autrement dit, l'opérateur

$$L_1 = H - \nu_1 I$$

est hermitien et positif (c'est ce qui est écrit ci–dessus). On utilise alors le
Lemme : *Si L est un opérateur hermitien positif, alors*

$$\Big|(Lu, v)\Big|^2 \leq (Lu, u)(Lv, v)$$

De plus s'il existe un x tel que $(Lx, x) = 0$, alors $Lx = 0$.
Démonstration : C'est une simple généralisation de l'inégalité de Schwarz. Comme L est positif, le membre de droite de l'inégalité a un sens, ce qui veut dire que puisqu'il est réel et positif, l'inégalité a un sens. Pour tout ρ réel, nous pouvons écrire

$$0 \leq (L(u + \rho v), u + \rho v) = (Lu, u) + \rho(Lv, u) + \rho(Lu, v) + \rho^2(Lv, v)$$

L est hermitien et réel (symétrique), donc

$$(Lv, u) = (v, Lu) = (Lu, v)$$

et l'inégalité s'écrivant

$$0 \leq (Lu, u) + 2\rho(Lu, v) + \rho^2(Lv, v)$$

peut s'interpréter comme un trinôme du second degré en ρ ; ce dernier est, pour tout ρ, du signe du coefficient de ρ^2. Son discriminant est donc négatif, soit

$$\Big|(Lu, v)\Big|^2 \leq (Lu, u)(Lv, v)$$

L'hypothèse d'existence d'un x tel que $(Lx, x) = 0$, appliquée à l'inégalité ci-dessus, conduit à :

$$\Big|(Lx, y)\Big|^2 \leq (Lx, x)(Ly, y) = 0$$

pour tout y. Donc $(Lx, y) = 0$ quelque soit y. Il suffira de choisir en particulier $y = Lx$ pour obtenir $Lx = 0$ et achever la démonstration du lemme.

Revenons à la démonstration du théorème. L_1, étant hermitien positif, satisfait à l'inégalité ci-dessus ; le choix $u = u_1$ la transforme en :

$$\Big|(L_1 u_1, v)\Big|^2 \leq (L_1 u_1, u_1)(Lv, v)$$

Or par hypothèse, $(L_1 u_1, u_1) = 0$. Donc $(L_1 u_1, v) = 0$ pour tout v et en particulier pour $v = L_1 u_1$. Donc $L_1 u_1 = 0$, soit de façon explicite

$$Hu_1 - \nu_1 u_1 = 0$$

ν_1 et u_1 sont donc des éléments propres de H.

Q.E.D.

Remarques.

• Par cette proposition, nous venons de démontrer l'existence d'éléments propres pour une matrice hermitienne (symétrique) indépendamment de toute notion de polynôme caractéristique.

• De plus nous avons montré que s'il existe un U_1 tel que

$$\min \frac{(Hu, u)}{(u, u)} = \frac{(Hu_1, u_1)}{(u_1, u_1)} = \nu_1$$

u_1 est tel que $Hu_1 = \nu_1 u_1$.

La valeur ν_1 est atteinte par tous les vecteurs propres correspondants de H. Si, en effet, il existe un autre u tel que

$$(Hu, u) = \nu_1(u, u)$$

u jouit des propriétés de u_1. Donc $Hu = \nu_1 u$ et u est un vecteur propre. Il y a donc unicité de l'espace propre correspondant à ν_1.

• Le théorème aurait pu être démontré de façon légèrement différente en remarquant que la fonction

$$u \longrightarrow (L_1 u, u)$$

a un minimum nul en $u = u_1$. Définissons donc pour tout v

$$\begin{aligned} J(\rho) &= \big(L_1(u_1 + \rho v), u_1 + \rho v\big) \\ &= (L_1 u_1, u_1) + 2\rho(L_1 u_1, v) + \rho^2(L_1 v, v) \\ &= 0 + 2\rho(L_1 u_1, v) + \rho^2(L_1 v, v) \end{aligned}$$

par hypothèse. Toujours par hypothèse, cette fonction doit avoir un minimum nul à l'origine. D'où

$$\left.\frac{dJ}{d\rho}\right|_{\rho=0} = 0 + 2(L_1u_1, v) + 0$$

Donc $(L_1u_1, v) = 0$ et on poursuit de même la démonstration.[7]

Théorème : *Soient les quantités ν_r et u_r définies par la r-ième récurrence*

$$\nu_r = \frac{(Hu_r, u_r)}{(u_r, u_r)} = \min\left\{\frac{(Hu, u)}{(u, u)} \;\middle|\; u \neq 0\ ,\ (u, u_1) = \cdots = (u, u_{r-1} = 0\right\}$$

alors ν_r et u_r sont éléments propres de H.

Démonstration

Le r-ième minimum et l'un des r-ièmes vecteurs minimisant satisfont donc aux relations

$$(Hu_r, u_r) - \nu_r(u_r, u_r) = 0 \quad , \quad (Hu, u) - \nu_r(u, u) \geq 0$$

pour tout u tel que

$$(u, u_1) = \cdots = (u, u_{r-1}) = 0$$

ou encore

$$\big([H - \nu_r I]u_r, u_r\big) = 0 \quad , \quad \big([H - \nu_r I]u, u\big) \geq 0$$

pour tout u orthogonal aux $r-1$ espaces propres précédents.

Les lemmes déjà utilisés s'appliquent de la même façon pour $H - \nu_r I \equiv L_r$ et montrent que

$$(L_r u_r, u) = \big([H - \nu_r I]u_r, u\big) = 0$$

pour tout u orthogonal aux $r-1$ espaces propres précédents.

$L_r u_r$ est donc orthogonal à tout vecteur orthogonal à l'espace engendré par $\{u_1, \ldots u_{r-1}\}$. Appelons ce dernier E_{r-1}. $L_r u_r$ appartient donc à cet espace. Or pour tout vecteur de base, on a

$$\begin{aligned}(L_r u_r, u_i) &= (Hu_r, u_i) - \nu_r(u_r, u_i) \quad i = 1, \ldots, r-1\\ &= \nu_r(u_r, u_i) - \nu_r(u_r, u_i)\\ &= 0\end{aligned}$$

$L_r u_r$ est donc orthogonal à tout vecteur de E_{r-1} ; c'est donc le vecteur nul. Donc

$$L_r u_r = 0 \qquad Hu_r - \nu_r u_r = 0$$

Q.E.D.

Remarques.

• Par la suite, nous verrons que les nombres $\nu_1, \ldots, \nu_N$ sont rangés en ordre croissant car les conditions qui les définissent

---

[7] A bien regarder, ces deux démonstrations sont fondamentalement les mêmes.

$$\nu_1 = \min\left\{ \frac{(Hu,u)}{(u,u)} \;\middle|\; u \neq 0 \right\} \quad \text{qui définit} \quad u_1$$
$$\nu_2 = \min\left\{ \frac{(Hu,u)}{(u,u)} \;\middle|\; u \neq 0\ ,\ (u,u_1) = 0 \right\} \quad \text{qui définit} \quad u_2$$
$$\cdots = \quad \cdots$$
$$\nu_r = \min\left\{ \frac{(Hu,u)}{(u,u)} \;\middle|\; u \neq 0\ ,\ (u,u_1) = \cdots = (u,u_{r-1} = 0 \right\} \quad \text{qui définit} \quad u_r$$

sont de plus en plus restrictives. Nous aurons donc

$$\nu_1 \leq \nu_2 \leq \cdots \leq \nu_r \leq \cdots \leq \nu_N$$

• Ce qui a été fait dans un sens aurait aussi bien se faire dans l'autre en partant de

$$\nu_N = \max\left\{ \frac{(Hu,u)}{(u,u)} \;\middle|\; u \neq 0 \right\} \quad \text{qui définit} \quad u_N$$

6.2. Contraintes sur un système oscillant.

L'un des très nets avantages de la définition variationnelle des éléments propres apparaît lorsqu'on étudie les systèmes oscillants sous contrainte. On va montrer successivement que, non seulement cette contrainte se "glisse" naturellement dans les expressions, mais que de plus une généralisation immédiate de cette notion permet une nouvelle définition des éléments propres de façon non récursive.

Définition d'une contrainte.

Une contrainte impose une relation sur les coordonnées du système ; celle-ci s'écrit

$$f(q_1, \ldots, q_N) = 0$$

Mais pour qu'elle ait un sens, il faut qu'elle soit compatible avec la position d'équilibre du système. Dans le cas des systèmes oscillants, nous avons choisi l'origine comme position d'équilibre ; dans cette hypothèse, la contrainte doit satisfaire à

$$f(0, \ldots, 0) = 0$$

Admettant ceci et nous bornant à l'étude des petites oscillations, nous pouvons remplacer cette contrainte dans un voisinage de l'origine par son développement en série de Taylor (en admettant les hypothèses adéquates sur f) pour obtenir :

$$f(q_1, \ldots, q_N) = \left(\frac{\partial f}{\partial q_1}\right)_0 q_1 + \cdots + \left(\frac{\partial f}{\partial q_N}\right)_0 q_N + o(q) = 0$$

De sorte que, si les termes supérieurs à ceux du premier ordre sont négligés, une contrainte peut être représentée par une équation du type

$$\alpha_1 q_1 + \cdots + \alpha_N q_N = 0 \quad \text{avec dans notre cas :} \quad \alpha_i = \left(\frac{\partial f}{\partial q_i}\right)_0$$

.

Nous prendrons donc une équation de ce type comme définition d'une contrainte. Géométriquement, c'est l'équation d'un plan passant par l'origine dans E_N.

Comme, analytiquement, les solutions du système sans contrainte appartenaient à E_N, celles du système avec contrainte sont obtenues à partir de ces dernières en leur imposant la condition de contrainte, ce qui revient à les lier par une relation linéaire et donc abaisser de un leur degré de liberté.

Géométriquement, les solutions appartenant au plan ont un degré de liberté de moins, puisque l'hyperplan est de dimension $N-1$ dans E_N.

Le système avec contrainte aura donc $N-1$ degrés de liberté et donc $N-1$ éléments propres normalisés que nous noterons

$$\widetilde{\nu}_1, \widetilde{\nu}_2, \ldots, \widetilde{\nu}_{N-1} \quad \widetilde{u}_1, \widetilde{u}_2, \ldots, \widetilde{u}_{N-1} \quad \text{normalisés}$$

l'espace de configuration étant noté, quant-à lui $\widetilde{E}_{N-1}$.

Remarque :

Dans le cas général, il n'est pas évident que le système continue à avoir des éléments propres lorsqu'il lui est ajouté une contrainte, c'est à dire qu'il est toujours possible de lui associer un opérateur hermitien. Mais c'est évident du point de vue expérimental comme le montrent les schémas de la figure 6.1.

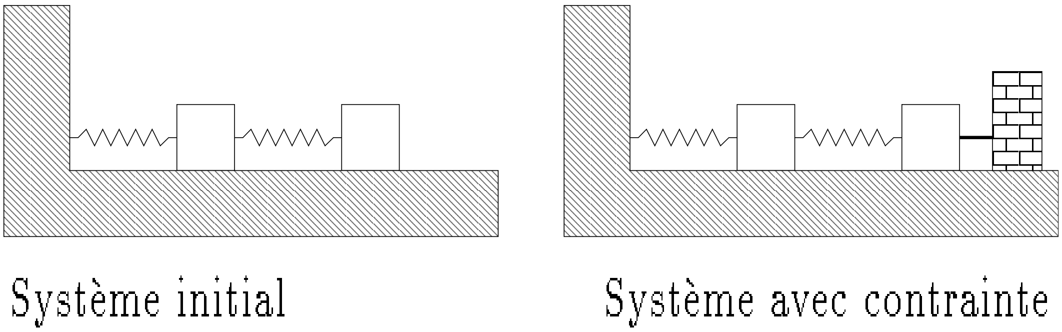

Figure 6.1. La contrainte du second schéma "bloque" la deuxième masse sur sa position d'équilibre.

Dans les deux cas, il est évident que les systèmes oscillent. Par contre, il est moins évident que la contrainte impose de passer par la coordonnée d'équilibre du système initial.

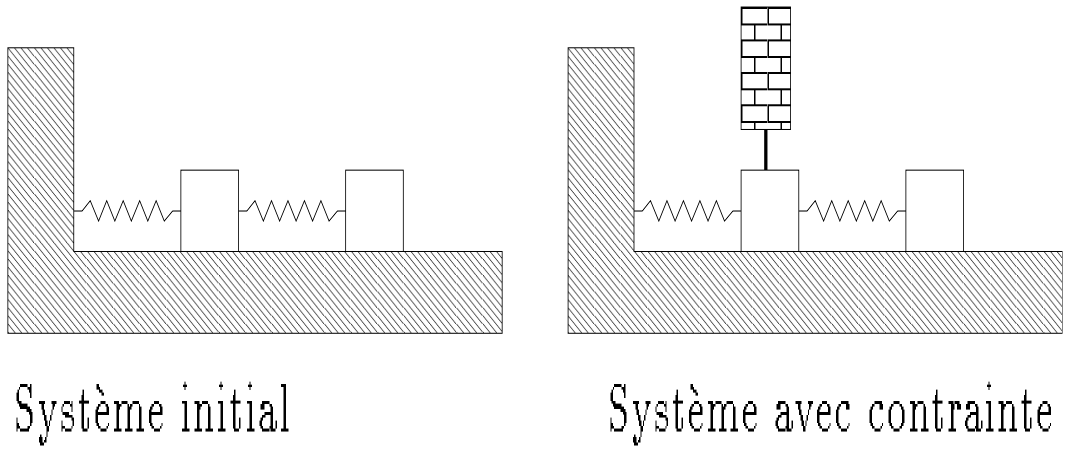

Figure 6.2. Ici, la contrainte maintient le premier ressort sur sa position d'équilibre.

Vecteur de contrainte

Soit dans E_N, l'équation du plan

$$\alpha_1 q_1 + \cdots + \alpha_N q_N = 0$$

de contrainte. Cette équation définit donc l'espace vectoriel $\widetilde{E}_{N-1}$ des solutions du système avec contrainte.

Soit φ un vecteur normalisé et orthogonal à $\widetilde{E}_{N-1}$. Ce vecteur φ sera par définition le *vecteur de contrainte*. A l'aide de ce vecteur et du produit scalaire, la condition de contrainte peut s'écrire

$$(u, \varphi) = 0$$

et définit *le plan de contrainte*.[8]

Géométriquement, il est intuitif que les solutions devront être choisies à l'intersection de l'ellipsoïde sur lequel sont définies les solutions du système libre dans E_N $(Hu, u) = 1$ avec le “plan” $(u, \varphi) = 0$.

Il s'agit encore d'un “ellipsoïde” dont les axes principaux fourniront les directions cherchées. Il est alors tout aussi intuitif de dire que le plus grand axe de ce nouvel ellipsoïde aura une longueur inférieure à celle du plus grand axe, mais sera supérieure à celle de l'axe suivant, ce qui doit entraîner

$$\nu_1 \leq \widetilde{\nu}_1 \leq \nu_2$$

C'est ce que démontrera le théorème de Rayleigh dans le cas d'une contrainte.

Remarque :

La condition de contrainte

$$(u, \varphi) = 0$$

[8] Si le produit scalaire choisi n'est pas le produit scalaire euclidien usuel, u sera le vecteur B-normal associé à grad φ par le théorème de Riesz.

est un produit scalaire. Dans la définition de la seconde valeur propre

$$\nu_2 = \min\left\{ \frac{(Hu,u)}{(u,u)} \;\middle|\; u \neq 0 \;,\; (u,u_1) = 0 \right\} \quad \text{qui définit également} \quad u_2$$

le produit scalaire (u, u_1) peut donc s'interpréter comme une contrainte définie par le premier vecteur propre.

Opérateur associé à une contrainte.

La question se pose alors de savoir quel est l'opérateur $\widetilde{H}$ correspondant au système avec contrainte, c'est à dire quel est l'opérateur qui a

$$\widetilde{\nu}_1, \widetilde{\nu}_2, \ldots, \widetilde{\nu}_{N-1} \quad , \quad \widetilde{u}_1, \widetilde{u}_2, \ldots, \widetilde{u}_{N-1} \quad \text{(normalisés)}$$

pour éléments propres.

Pour répondre à cette question, il est logique d'introduire P l'opérateur de projection sur $\widetilde{E}_{N-1}$. P s'exprime aisément à l'aide du vecteur de contrainte φ. On a

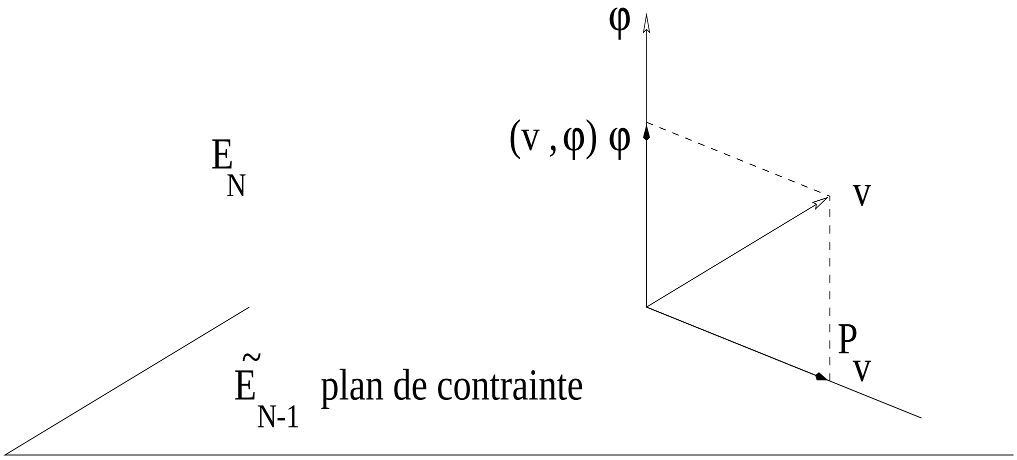

Figure 6.3. Contrainte, plan de contrainte et vecteur de contrainte.

$$\begin{aligned} &Pv = v - (v,\varphi)\varphi \quad \text{puisque} \quad E_N = \widetilde{E}_{N-1} \oplus \mathrm{Vect}\,\{\varphi\} \\ &(\varphi,\varphi) = 1 \end{aligned}$$

où Vect $\{\varphi\}$ est l'espace vectoriel engendré par φ.

Si $u \in \widetilde{E}_{N-1}$, on a bien $Pu = u$ puisque $(u, \varphi) = 0$ par hypothèse.

Si $u = k\varphi$, on a bien $Pu = 0$ car

$$\begin{aligned} Pu &= k\varphi - (k\varphi,\varphi)\varphi = k\varphi - k(\varphi,\varphi)\varphi \\ &= k\varphi - k\varphi = 0 \end{aligned}$$

puisque φ est normalisé.

Les principales propriétés du projecteur P sont les suivantes :

• P est positif. En effet

$$\begin{aligned}(Pv, v) &= (v, v) - (v, \varphi)^2 \\ &= (v, v)(\varphi, \varphi) - (v, \varphi)^2 && (\text{car } \|\varphi\| = 1) \\ &\geq 0 && (\text{d'après l'inégalité de Schwarz})\end{aligned}$$

• P n'est pas défini puisque

$$(P\varphi, \varphi) = \varphi - (\varphi, \varphi)\varphi = 0 \quad \text{avec} \quad \varphi \neq 0$$

• Par contre, $P^2 = P$ car

$$P^2 v = Pv - (v, \varphi)P\varphi = Pv \quad \text{puisque} \quad P\varphi = 0$$

• P est hermitien car

$$\begin{aligned}(Pv, u) &= (v - (v, \varphi)\varphi, u) = (v, u) - (v, \varphi)(\varphi, u) \\ &= (v, u) - (v, \varphi(\varphi, u)) = (v, u - (\varphi, u)\varphi) \\ &= (v, Pu)\end{aligned}$$

• Enfin, pour tout $v \in E_N$ et tout $u \in \widetilde{E}_{N-1}$, on a $(v, u) = (Pv, u)$, qui n'est qu'une conséquence de la somme directe orthogonale.

$$\begin{gathered}v = Pv + (v, \varphi)\varphi \\ (v, u) = (Pv, u) + ((v, \varphi)\varphi, u)\end{gathered}$$

Or $(v, \varphi)\varphi$ appartient à l'orthogonal de $\widetilde{E}_{N-1}$; le produit scalaire $((v, \varphi)\varphi, u)$ est donc nul.

Revenons à notre problème. Par définition,

$$\widetilde{\nu}_1 = (H\widetilde{u}_1, \widetilde{u}_1) = \min\left\{(Hu, u) \;\middle|\; u \in \widetilde{E}_{N-1}\ ,\ (u, u) = 1\right\}$$

car $H\widetilde{u}_1$ appartient à E_N et $\widetilde{u}_1$ à $\widetilde{E}_{N-1}$. En appliquant le théorème précédent

$$\widetilde{\nu}_1 = (PH\widetilde{u}_1, \widetilde{u}_1) = \min\left\{(Hu, u) \;\middle|\; u \in \widetilde{E}_{N-1}\ ,\ (u, u) = 1\right\}$$

nous définissons un opérateur de $\widetilde{E}_{N-1} \quad \longrightarrow \quad \widetilde{E}_{N-1}$.

Reste à montrer que cet opérateur est hermitien par rapport au produit scalaire induit, ce qui est vrai puisque pour tout u et v dans $\widetilde{E}_{N-1}$,

$$\begin{aligned}(PHu, v) &= (Hu, Pv) = (Hu, v) \quad \text{car} \quad v \in \widetilde{E}_{N-1} \\ &= (HPu, v) \quad \text{car} \quad u \in \widetilde{E}_{N-1}\end{aligned}$$

Dans certains cas, il peut être plus intéressant de définir H comme opérateur, non sur $\widetilde{E}_{N-1}$, mais sur E_N. Alors

$$\begin{aligned}\widetilde{\nu}_1 = (H\widetilde{u}_1, \widetilde{u}_1) &= \min\Big\{(Hu,u) \;\Big|\; u \in \widetilde{E}_{N-1}\ ,\ (u,u)=1\Big\}\\ &= \min\Big\{(HPu,Pu) \;\Big|\; (Pu,Pu)=1\Big\}\\ &= \min\Big\{(PHPu,u) \;\Big|\; (P^2u,u)=(Pu,u)=1\Big\}\end{aligned}$$

ce qui fait apparaître l'opérateur PHP qui est hermitien sur E_N.
Ce résultat s'étend à toutes les valeurs propres pour lesquelles est défini le même PHP. L'ensemble des valeurs propres formant un système complet, PHP est complètement défini.
Remarque : L'opérateur P se généralise aisément à k contraintes ; dans ce cas, il s'écrira

$$Pu = u - (u,\varphi_1)\varphi_1 - (u,\varphi_2)\varphi_2 - \cdots (u,\varphi_k)\varphi_k$$

si les φ_i sont les vecteurs de contrainte normalisés.

6.3. Principe de comparaison ou de monotonie.

Enoncé :

Si, aux conditions imposées sur les solutions possibles d'un problème variationnel, on ajoute des conditions supplémentaires, on obtient un "problème variationnel modifié".

• *Si ces problèmes variationnels correspondent à des* minima, *le minimum solution du problème modifié ne peut être plus petit que le minimum défini par le problème originel.*

• *Si ces problèmes variationnels correspondent à des* maxima, *le maximum solution du problème modifié ne peut être plus grand que le maximum défini par le problème originel.*

Démonstration :

Soit f une application définie sur un ensemble E et à valeurs réelles. Soient

$$A = \Big\{u \;\Big|\; u \in E \text{ satisfait à la condition } C1\Big\} \quad \text{et}$$
$$B = \Big\{u \;\Big|\; u \in E \text{ satisfait à la condition } C2\Big\}$$

Alors

$$\Big\{u \;\Big|\; \text{u satisfait aux conditions } C1 \text{ et } C2\Big\} = A \cap B \subset A$$

Donc

$$f\big(A \cap B\big) \subset f\big(A\big)$$

et puisque f est à valeurs réelles et que $A \cap B \neq \emptyset$ (sinon le problème variationnel n'aurait pas de sens)

$$\min f\big(A\big) \leq \min f\big(A \cap B\big) \quad \text{et} \quad \max f\big(A\big) \geq \max f\big(A \cap B\big)$$

Comme le minimum (resp. maximum) de $f\big(A\big)$ correspond au problème originel et comme le minimum (resp. maximum) de $f\big(A \cap B\big)$ correspond au problème modifié, la proposition est démontrée.

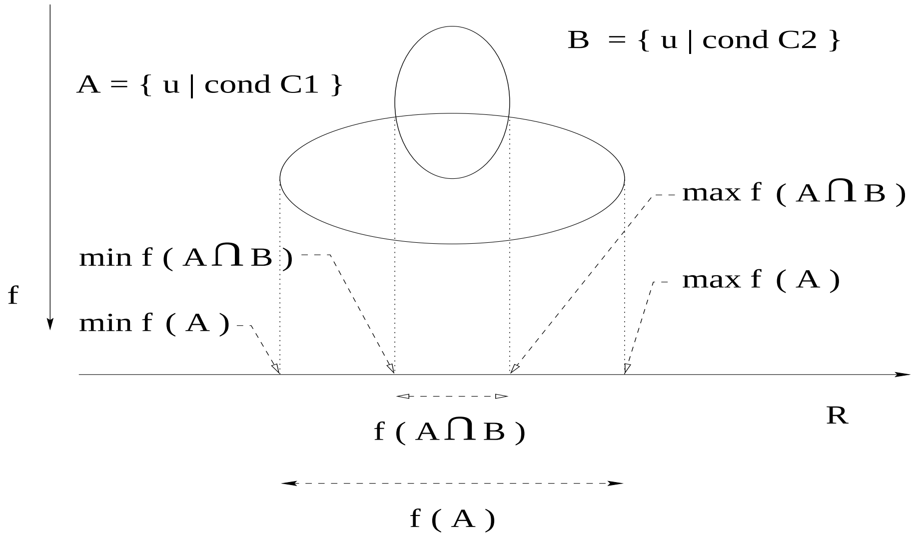

Figure 6.4. Le principe de monotonie.

Ce principe illustre simplement le fait que le minimum croit de façon monotone avec les contraintes.

De son côté, le maximum décroît de façon monotone avec les mêmes contraintes.

Exemple :

Les étudiants des MPA de toutes les universités de France subissent une même épreuve. Celle-ci est corrigée (sans cela, elle n'a aucun sens) et une note est attribuée à chacun.

La meilleure note est définie par

$$N_2 = \max\Big\{\text{notes des étudiants} \;\Big|\; \text{étudiants} \in \text{MPA}\Big\}$$

Nous pouvons ajouter, par exemple, la condition que les étudiants appartiennent à l'Université Pierre et Marie Curie et définir

$$N_1 = \max\Big\{\text{notes des étudiants} \;\Big|\; \text{étudiants} \in \text{MPA et étudiants} \in \text{UPMC}\Big\}$$

Le *principe de monotonie* dit simplement que

$$N_1 \leq N_2$$

On aurait pu prendre la moins bonne note, mais cela eût été faire injure aux étudiants de Pierre et Marie Curie qui, comme chacun sait, sont tous des cracks.

6.4. Le théorème de Rayleigh pour une contrainte.

Originalité :

D'après la définition habituelle des valeurs propres, il n'est pas du tout évident de prédire de quelle façon les valeurs propres d'un système sont affectées par l'adjonction de contraintes.

Il est par contre aisé, à partir de leurs propriétés variationnelles, de prouver le fameux théorème de Rayleigh qui indique en particulier que les valeurs propres du système avec contraintes ne peuvent être plus petites que celles du système originel. Plus précisément :

Enoncé du théorème de Rayleigh :

La plus basse valeur propre d'un système soumis à une *contrainte se trouve comprise entre la première et la seconde valeur propre du système originel. Soit avec les notations que nous utilisons*

$$\nu_1 \leq \widetilde{\nu}_1 \leq \nu_2$$

ll existe bien entendu un énoncé correspondant pour le maximum.

Remarque géométrique :

L'interprétation géométrique des éléments propres rend ce résultat évident :

Les directions propres correspondant à $1/\nu_1$ et $1/\nu_2$ sont associées aux deux plus grands axes de l'ellipsoïde

$$(Hu, u) = 1$$

$1/\widetilde{\nu}_1$ correspond au plus grand axe de l'ellipsoïde intersection du précédent avec le plan $(u, \varphi) = 0$. Cet axe ne pourra donc pas être plus grand que $1/\nu_1$, mais ne pourra être plus petit que $1/\nu_2$.

Démonstration :

Puisque

$$\nu_1 = \min \left\{ \frac{(Hu, u)}{(u, u)} \;\middle|\; u \in E_N \right\}$$

$$\widetilde{\nu}_1 = \min \left\{ \frac{(Hu, u)}{(u, u)} \;\middle|\; u \in E_N \ , \ (\varphi, u) = 0 \right\}$$

l'application du principe de comparaison donne

$$\nu_1 \leq \widetilde{\nu}_1$$

C'est le second volet de l'inégalité qui est un peu plus délicat à démontrer : nous voulons montrer que si $\widetilde{\nu}_1$ est supérieur à ν_1, il ne peut en fait dépasser ν_2. Pour cela, il faut répondre à la question : quelle est la valeur maximum que peut prendre $\widetilde{\nu}_1 = \widetilde{\nu}_1(\varphi)$?

Si nous choisissons en particulier $\varphi = u_1$, le plan de contrainte est orthogonal au premier vecteur propre, de sorte que, par définition de la deuxième valeur propre :

$$\widetilde{\nu}_1(\varphi = u_1) = \nu_2$$

Donc, pour au moins un choix de vecteur de contrainte (celui de la première valeur propre), on a

$$\widetilde{\nu}_1 = \nu_2$$

Reste à démontrer que cette valeur ne peut être dépassée.

$\diamond$ Tout d'abord si $u \in \{\varphi\}^\perp$, comme

$$\widetilde{\nu}_1 = \min\left\{ \frac{(Hu,u)}{(u,u)} \;\middle|\; u \in E_N \ , \ u \neq 0 \ , \ (\varphi, u) = 0 \right\}$$

on aura toujours par définition du minimum

$$\widetilde{\nu}_1 \leq \frac{(Hu,u)}{(u,u)}$$

$\diamond$ Reste à majorer $(Hu,u)/(u,u)$. Soit $u \in \text{Vect}\,\{u_1, u_2\}$. Le développement de u sur u_1 et u_2 supposés normés s'écrit :

$$u = a_1 u_1 + a_2 u_2 \quad \text{où} \quad \Big(a_i = (u_i, u) \ ; \ i = 1, 2\Big)$$

et donc

$$(Hu,u) = a_1^2 (Hu_1, u_1) + 2 a_1 a_2 (Hu_1, u_2) + a_2^2 (Hu_2, u_2) \quad .$$

Or $(Hu_i, u_j) = \nu_i (u_i, u_j) = \nu_i \delta_{ij}$, $i, j = 1, 2$, donc

$$(Hu,u) = a_1^2 \nu_1 + a_2^2 \nu_2 \leq (a_1^2 + a_2^2) \nu_2 = \nu_2 (u,u) \quad \text{car} \quad \nu_1 \leq \nu_2$$

Donc

$$\frac{(Hu,u)}{(u,u)} \leq \nu_2$$

$\diamond$ Définissons maintenant A par

$$A = \left\{ u \;\middle|\; u \in \{\varphi\}^\perp \quad \text{et} \quad u \in \text{Vect}\,\{u_1, u_2\} \right\}$$

Cet ensemble n'est pas réduit au vecteur nul. En effet

$$\begin{aligned} &dim \quad \{\varphi\}^\perp = N - 1 \\ &dim \quad \text{Vect}\,\{u_1, u_2\} = 2 \end{aligned}$$

Comme ce sont deux sous-espaces vectoriels de E_N (de dimension n), leur intersection ne peut être nulle. Soit donc $u \in A$; d'après 1 et 2 :

$$\widetilde{\nu}_1 \leq \frac{(Hu,u)}{(u,u)} \leq \nu_2$$

Q.E.D.

Remarque :

Si on avait choisi Vect $\{u_i\}$, l'intersection aurait toujours été réduite à 0 sauf pour le choix de $\varphi \not\equiv u_i$. dans ce cas, la définition du minimum n'a pas de sens la plupart du temps. On est donc obligé de choisir deux vecteurs

$$\text{Vect}\,\{u_i, u_j\} \quad j > i$$

Par la même démonstration, on vérifierait que

$$\frac{(Hu,u)}{(u,u)} \leq \nu_j$$

ce qui est toujours vrai puisque $j \geq 2$. Les vecteurs à choisir "serrant" le plus l'inégalité sont donc u_1 et u_2.

Remarque :

Ce qui a été fait pour les plus petites valeurs propres peut également se faire pour les plus grandes pour lesquelles le théorème de Rayleigh s'énonce :

La plus grande valeur propre $\widetilde{\nu}_{N-1}$ d'un système soumis à une contrainte se trouve comprise entre les deux plus grandes valeurs propres du système originel, soit

$$\nu_{N-1} \leq \widetilde{\nu}_{N-1} \leq \nu_N$$

6.5. Le théorème du Min-Max ou la caractérisation variationnelle non-récursive des éléments propres.

Introduction.

Nous venons de montrer dans le théorème de Rayleigh que, pour une contrainte quelconque φ, on avait

$$\nu_1 \leq \widetilde{\nu}_1 \leq \nu_2$$

ν_2 est donc une borne supérieure des ν_1 qui est atteinte pour des contraintes particulières, à savoir celles dont les vecteurs de contrainte sont dans l'espace propre associé à ν_1. Donc pour toute φ

$$\min_u \left\{ \frac{(Hu,u)}{(u,u)} \;\middle|\; u \in E_N \;,\; u \neq 0 \;,\; (u,\varphi) = 0 \right\} \leq \nu_2$$

et puisque ν_2 est atteint

$$\nu_2 = \max_\varphi \left\{ \min_u \left\{ \frac{(Hu,u)}{(u,u)} \;\middle|\; u \in E_N \;,\; u \neq 0 \;,\; (u,\varphi) = 0 \right\} \right\}$$

Nous pouvons de même écrire

$$\nu_{N-1} = \min_{\varphi}\Big\{\max_{u}\Big\{\frac{(Hu,u)}{(u,u)} \;\Big|\; u \in E_N \ , \ u \neq 0 \ , \ (u,\varphi) = 0\Big\}\Big\}$$

et remarquer que ν_2 (ou ν_{N-1}) est caractérisée sans qu'il soit fait référence au vecteur propre associé à ν_1 (ou à ν_N). Il s'agit donc d'une caractérisation *non récurrente* que nous allons généraliser sous le nom de *théorème du min-max.*

Remarques historiques.

◇ Lord Rayleigh regardait l'inégalité

$$\nu_1 \leq \widetilde{\nu}_1 \leq \nu_2$$

comme fournissant des propriétés de $\widetilde{\nu}_1$.

◇ Richard Courant (1888–1972) concevait la même inégalité comme donnant des renseignements sur ν_2.

◇ Le *théorème du min–max* donnant la définition non-récursive des valeurs propres a été donnée par E. Fischer (Monatshefte für Math. und Phys. **16**, 1905, 234–249) à la page 249 d'un article donnant une définition purement analytique de la signature d'une forme quadratique sans faire mention des vibrations.

◇ C'est l'article de R. Courant, écrit 15 ans plus tard, sans connaître, semble-t-il le travail de Fischer (Math. Zeitschr, **7**, 1920, 1–57) qui met en avant l'importance de cette définition dans les systèmes physiques.

Théorème du min–max.

Nous l'énonçons dans le cas *max–min.*

La r–ième valeur propre d'un système oscillant est le maximum que peut prendre $(Hu,u)/(u,u)$ quand on lui ajoute $r-1$ contraintes.

Autrement dit, si $\varphi_1, \ldots, \varphi_{r-1}$ sont $r-1$ vecteurs de contrainte, on a

$$\nu_r = \max_{(\varphi_i)}\Big\{\min_{u}\Big\{\frac{(Hu,u)}{(u,u)} \;\Big|\; (u,\varphi_i) = 0, \quad i = 1, \ldots, r-1\Big\}\Big\}$$

(Il y a bien entendu un énoncé “dans l'autre sens”).

Démonstration

Elle est calquée sur le théorème de Rayleigh avec une contrainte. Soit

$$\widetilde{\nu}(\varphi_1, \ldots, \varphi_{r-1}) = \min_{u}\Big\{\frac{(Hu,u)}{(u,u)} \;\Big|\; (u,\varphi_i) = 0 \ , \ i = 1, \ldots, r-1\Big\}$$

D'après le principe de comparaison, ces contraintes élèvent le “niveau” du système, c'est à dire que $\nu_1 \leq \widetilde{\nu}$. Le problème est de savoir jusqu'où ?

En choisissant comme système des φ_i particulier celui des $r-1$ premiers vecteurs propres,

$$\varphi_1 = u_1, \ldots, \varphi_{r-1} = u_{r-1}$$

nous obtenons

$$\widetilde{\nu}_1(u_1, \ldots, u_{r-1}) = \nu_r$$

Montrons que cette valeur ne peut être dépassée.

◇ 1. Si u appartient à l'hyperplan $(u, \varphi_1) = \cdots = (u, \varphi_{r-1}) = 0$ et comme, par définition

$$\widetilde{\nu}_r(\varphi_1, \ldots, \varphi_{r-1}) = \min_u \left\{ \frac{(Hu, u)}{(u, u)} \;\middle|\; (\varphi_i, u) = 0 \right\}$$

La définition du minimum conduit à

$$\widetilde{\nu}_r(\varphi_1, \ldots, \varphi_{r-1}) \leq \frac{(Hu, u)}{(u, u)}$$

◇ 2. Par ailleurs, supposant $u_1, \ldots, u_r$ orthonormés, $u \in \text{Vect}\,\{u_1, \ldots, u_r\}$ entraîne

$$u = k_1 u_1 + \cdots + k_r u_r \quad k_i = (u_i, u)$$

Donc

$$\begin{aligned}(Hu, u) &= (H \sum_i k_i u_i, \sum_j k_j u_j) \\ &= \sum_{ij} k_i k_j (Hu_i, u_j) \\ &= \sum_{i,j \nu_i k_i k_j (u_i, u_j)} = \sum_i \nu_i k_i^2\end{aligned}$$

Comme les ν_i sont en ordre croissant

$$(Hu, u) = \sum_i \nu_i k_i^2 \leq \nu_r \sum_i k_i^2 = \nu_r (u, u)$$

$$\frac{(Hu, u)}{(u, u)} \leq \nu_r$$

◇ 3. Soit

$$A = \left\{ u \;\middle|\; u \in \{\varphi_1, \ldots, \varphi_{r-1}\}^{\perp} \quad \text{et} \quad u \in \text{Vect}\,\{\varphi_1, \ldots, \varphi_{r-1}\} \right\}$$

Comme

$$\begin{aligned}\dim\{\varphi_1, \ldots, \varphi_{r-1}\} &= N - r - 1 \quad \text{et} \\ \dim\text{Vect}\,\{\varphi_1, \ldots, \varphi_{r-1}\} &= r\end{aligned}$$

l'intersection des deux est non vide et il est possible de prendre un élément non nul dedans. Pour cet élément u

$$\widetilde{\nu}_r(\varphi_1, \ldots, \varphi_r) \leq \frac{(Hu, u)}{(u, u)} \leq \nu_r$$

Comme ν_r peut être atteint par un choix particulier de $\varphi_1, \ldots, \varphi_{r-1}$

$$\nu_r = \max_{\varphi} \Big\{ \frac{\min_u (Hu, u)}{(u, u)} \;\Big|\; (u, \varphi_i) = 0 \Big\}$$

Q.E.D.

Les mêmes remarques que dans le cas du théorème de Rayleigh peuvent être faites ici.

Applications.

Théorème : *Soient A et B deux matrices hermitiennes, et B définie positive, alors*

$$\nu_k(A+B) \geq \nu_k(A) \quad \textit{pour tout } k = 1, N.$$

démonstration : Par le théorème du maximum, on a :

$$\nu_k(A+B) = \max_{\varphi_1, \ldots, \varphi_{k-1}} \Big\{ \min_u \frac{((A+B)u, u)}{(u, u)} \;\Big|\; (u, \varphi_i) = 0 \quad, i = 1, \ldots, k-1 \Big\}$$

Or

$$\begin{aligned} \frac{((A+B)u, u)}{(u, u)} &= \frac{(Au, u)}{(u, u)} + \frac{(Bu, u)}{(u, u)} \quad \text{par linéarité} \\ &\geq \frac{(Au, u)}{(u, u)} + \min_u \frac{(Bu, u)}{(u, u)} \\ &= \frac{(Au, u)}{(u, u)} + \nu_1 \end{aligned}$$

Donc pour tout $(\varphi_1, \ldots, \varphi_k)$

$$\begin{aligned} \min_u \Big\{ \frac{((A+B)u, u)}{(u, u)} \;\Big|\; (u, \varphi_i) = 0 \quad, i = 1, \ldots, k-1 \Big\} &\geq \min \Big\{ \frac{(Au, u)}{(u, u)} + \frac{(Bu, u)}{(u, u)} \Big\} \\ &\geq \min_u \Big\{ \frac{(Au, u)}{(u, u)} + \nu_1 \Big\} \\ &= \min_u \Big\{ \frac{(Au, u)}{(u, u)} \Big\} + \nu_1 \end{aligned}$$

Les $(\varphi_1, \ldots, \varphi_{k-1})$ sont des paramètres dans les deux membres qui forment une inégalité entre deux fonctions dans $(\varphi_1, \ldots, \varphi_{k-1})$. En prenant le maximum des deux, nous obtenons

$$\max_{\varphi} \Big\{ \min_u \Big\{ \frac{((A+B)u, u)}{(u, u)} \Big\} \Big\} \geq \max_{\varphi} \Big\{ \min_u \frac{(Au, u)}{(u, u)} + \nu_1 \Big\} \Big\}$$

d'où

$$\nu_k(A+B) \geq \nu_k(A) + \nu_1(B)$$

et comme $\nu_1 > 0$ (B défini positif)

$$\nu_k(A+B) \geq \nu_k$$

• Exemple : Déplacement des raies spectrales.

• Remarques :

– Partant de

$$\nu_r(A+B) = \min_{\varphi_1,\ldots,\varphi_{r-1}} \left\{ \max_u \frac{(Au,u)}{(u,u)} \;\middle|\; (u,\varphi_i)=0, \;,\; i=1,\ldots,r-1 \right\}$$

nous aurions trouvé le même résultat.

– Le résultat est obtenu globalement sur toutes les valeurs propres.

6.6. Le théorème de Rayleigh pour un nombre quelconque de contraintes.

Enoncé.

Si sont imposées h contraintes arbitraires

$$(u,\psi_1) = \cdots = (u,\psi_h) = 0$$

où les ψ_i sont les h vecteurs de contrainte indépendants sur un système oscillant dont les valeurs propres sont

$$\nu_1 \leq \cdots \leq \nu_N \quad ,$$

les nouvelles valeurs propres

$$\widetilde{\nu}_1 \leq \widetilde{\nu}_2 \leq \cdots \leq \widetilde{\nu}_N$$

se "glissent" entre les anciennes de sorte que

$$\nu_r \leq \widetilde{\nu}_r \leq \nu_{r+h} \quad \textit{pour tout} \quad r \leq N-h$$

En particulier, cela signifie que s'il n'y a qu'une seule contrainte, tout ceci s'écrit

$$\nu_1 \leq \widetilde{\nu}_1 \leq \nu_2 \leq \widetilde{\nu}_2 \leq \nu_3 \leq \cdots \leq \nu_{N-1} \leq \widetilde{\nu}_{N-1} \leq \nu_N$$

Démonstration.

Démontrer un tel théorème avec les définitions récursives classiques est difficile car il faut comparer deux ensembles de conditions d'orthogonalité différentes

– celui des vecteurs propres.

– celui des conditions de contrainte.

Par contre, la caractérisation min–max conduit à

(1) $$\nu_r = \max_{(\varphi_1,\ldots,\varphi_{r-1})} \left\{ \left\{ \min \frac{(Hu,u)}{(u,u)} \;\middle|\; (u,\varphi_i)=0, i=1,\ldots,r-1 \right\} \right\}$$

$$(2)\qquad \begin{aligned}&\widetilde{\nu}_r(\psi_1,\ldots,\psi_h) = \max_{(\varphi_1,\ldots,\varphi_{r-1})}\\ &\left\{\left\{\min\frac{(Hu,u)}{(u,u)}\ \middle|\ (u,\varphi_i)=0, i=1,\ldots,r-1 \quad , \quad (u,\psi_j)=0, j=1,\ldots,h\right\}\right\}\end{aligned}$$

$$(3)\qquad \nu_{r+h} = \max_{(\varphi_1,\ldots,\varphi_{r+h-1})}\left\{\left\{\min\frac{(Hu,u)}{(u,u)}\ \middle|\ (u,\varphi_i)=0, i=1,\ldots,r+h-1\right\}\right\}$$

Compare (2) et (3). Dans (2), des contraintes sont imposées à des variables qui sont libres dans (3) : le maximum ne peut donc qu'être plus petit, et donc

$$\widetilde{\nu}_r(\psi_1,\ldots,\psi_h) \le \nu_{r+h}$$

Par ailleurs, pour le choix particulier

$$\varphi_1 = u_1, \ldots, \varphi_{r-1} = u_{r-1}$$

(2) s'écrit

$$\begin{aligned}&\min_u\left\{\frac{(Hu,u)}{(u,u)}\ \middle|\ (u,\varphi_i)=0, i=1,\ldots,r-1 \quad \text{et} \quad (u,\psi_j)=0, j=1,\ldots,h\right\}\\ &\qquad \le \widetilde{\nu}_r(\psi_1,\ldots,\psi_h) \le \nu_{r+h}\end{aligned}$$

Comme ν_r peut aussi être caractérisé par le membre de gauche (définition récursive) en enlevant les contraintes $(u,\psi_j)=0$, le principe de comparaison conduit à

$$\begin{aligned}\nu_r &= \min_u\left\{\frac{(Hu,u)}{(u,u)}\ \middle|\ (u,\varphi_i)=0, i=1,\ldots,r-1\right\}\\ &\le \min_u\left\{\frac{(Hu,u)}{(u,u)}\ \middle|\ (u,\varphi_i)=0, i=1,\ldots,r-1 \quad \text{et} \quad (u,\psi_j)=0, j=1,\ldots,h\right\}\\ &\le \widetilde{\nu}_r(\psi_1,\ldots,\psi_h)\end{aligned}$$

et en regroupant le tout

$$\nu_r \le \widetilde{\nu}_r(\psi_1,\ldots,\psi_h) \le \nu_{r+h}$$

Q.E.D.

Remarque.

C'est dans la comparaison des valeurs propres du problème modifié que la caractérisation min–max prend le plus d'importance. Des méthodes de calcul comme celles de Rayleigh–Ritz, et plus récemment de A. Weinstein (1937), consistent à trouver des bornes supérieures et inférieures de certaines valeurs propres en les comparant à celles du système modifié.

CAHIERS D'ENSEIGNEMENT

Du même auteur et dans la même collection

FASCICULE I – Systèmes linéaires. Une introduction à l'algèbre linéaire.

FASCICULE II – (TAP 1) Présentation du cours.

FASCICULE III – (TAP 2) Analyse numérique et Ordinateurs.

FASCICULE IV – (TAP 3) Extrema et Différentielles.

FASCICULE V – (TAP 4) Calcul des variations.

FASCICULE VI – (TAP 5) Formes quadratiques et Opérateurs associés.

FASCICULE VII – (TAP 6) Principes variationnels pour des systèmes de dimension finie.

FASCICULE VIII – (TAP 7) Méthodes 1.

FASCICULE IX – (TAP 8) Méthodes 2.

FASCICULE X – Étude et tracé des courbes du type $f(x, y) = 0$.