M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique

M.-L. Mazure

L'addition parallèle d'opérateurs interprétée comme inf-convolution de formes quadratiques convexes

M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique, tome 20, nº 3 (1986), p. 497-515

<http://www.numdam.org/item?id=M2AN_1986__20_3_497_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

(Vol 20, n° 3, 1986, p 497 à 515)

L'ADDITION PARALLÈLE D'OPÉRATEURS INTERPRÉTÉE COMME INF-CONVOLUTION DE FORMES QUADRATIQUES CONVEXES (*)

par M.-L. MAZURE [1]

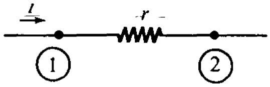
Communiqué par P J LAURENT

Résumé — *Étant donné deux circuits en parallèle, de résistances représentées par les matrices semi-définies positives A et B respectivement, la résistance équivalente est exprimée par ce que l'on appelle la somme parallèle des opérateurs A et B On interprète cette opération comme l'inf-convolution des formes quadratiques associées à A et B (principe variationnel sous-jacent) Les propriétés de l'addition parallèle sont alors déduites de celles de l'inf-convolution de fonctions convexes*

Abstract — *Given two parallels circuits with resistance determined by semi-definite positive matrices A, B, the equivalent resistance of the circuit is expressed in terms of A and B by the so-called parallel sum of the operators A and B We interpret this operation as the inf-convolution of the quadractic forms associated to A and B The properties of the parallel sum are then deduced from those of the inf-convolution of convex functions*

I. MODÈLE PHYSIQUE SOUS-JACENT

Un circuit électrique de résistance r, traversé par un courant i, obéit à la loi d'Ohm $v = ri$ où v est la tension mesurée entre les points (1) et (2) du circuit :



r est un nombre réel strictement positif, ou, si l'on admet les courts-circuits, un réel positif ou nul.

(*) Reçu en juillet 1985

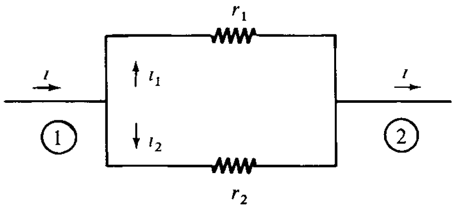
[1] Laboratoire d'Analyse Numérique, Université Paul-Sabatier, 118, Route de Narbonne, 31062 Toulouse Cedex

Deux résistances r_1 et r_2 peuvent être ajoutées :

1) *En série,*

Dans ce cas, la résistance équivalente à l'ensemble est $r = r_1 + r_2$.

2) *En parallèle.*



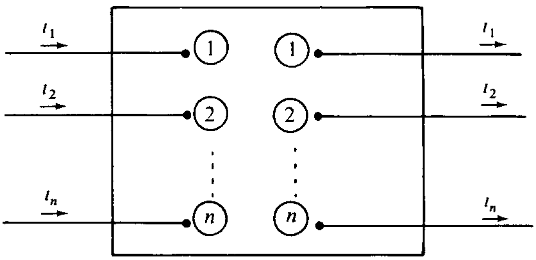
Le courant i se partage en i_1 (à travers r_1) et i_2 (à travers r_2) suivant la loi de Kirchhoff : $i = i_1 + i_2$.

Si r est la résistance équivalente à l'ensemble et v la tension entre les points ① et ②, la loi d'Ohm nous donne :

$$v = ri = r(i_1 + i_2) = r_1 i_1 = r_2 i_2$$

ce qui permet d'écrire, dans le cas où r_1 et r_2 sont strictement positifs, la loi d'addition des résistances en parallèle : $\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2}$, c'est-à-dire $r = (r_1^{-1} + r_2^{-1})^{-1}$ que nous noterons $r_1 /\!\!/ r_2$.

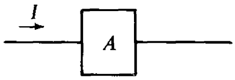
Si r_1 (ou r_2) est nul, il est clair que $r = 0$, que nous noterons encore $r_1 /\!\!/ r_2 = 0$. Nous avons ainsi obtenu « l'addition parallèle » de scalaires positifs ou nuls. Généralisons en considérant maintenant un réseau électrique d'ordre n, c'est-à-dire un ensemble de n couples « entrée-sortie » que nous représentons de la manière suivante :



Si l'on mesure la tension v_k entre « l'entrée » k et la « sortie » k (pour $k = 1, \ldots, n$), le courant correspondant étant i_k, un tel réseau obéit encore à la loi d'Ohm généralisée que nous écrirons :

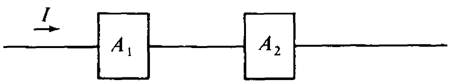
$V = AI$ où, maintenant, la « tension » s'écrit $V = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$, le « courant » $I = \begin{pmatrix} i_1 \\ \vdots \\ i_n \end{pmatrix}$ et où A, *résistance généralisée* du réseau, est une matrice symétrique semi-définie positive (généralisation naturelle d'un nombre réel positif ou nul).

Plus schématiquement, nous symboliserons le réseau ci-dessus par :



Ici encore, nous pouvons ajouter deux réseaux d'ordre n de résistances généralisées A_1 et A_2 soit :

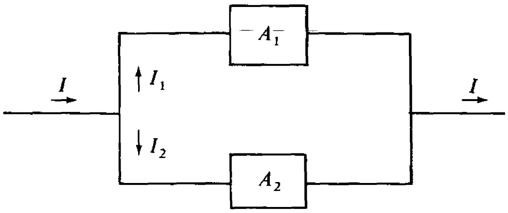
1) *En série.*



résistance équivalente $A_1 + A_2$

soit

2) *En parallèle.*



$I = I_1 + I_2$

Si A_1 et A_2 sont inversibles (c'est-à-dire ici définies positives) un calcul analogue au cas scalaire montre que la matrice résistance équivalente à l'ensemble, est $A = (A_1^{-1} + A_2^{-1})^{-1}$.

Dans le cas non inversible, nous prouverons que, lorsque $A_1 I_1 = A_2 I_2$ ce terme ne dépend que de $I = I_1 + I_2$. Plus précisément $A_1 I_1 = A_2 I_2 = (A_1 /\!/ A_2)(I)$ où $A_1 /\!/ A_2$ est une matrice symétrique semi-définie positive.

Cette « *somme parallèle de A_1 et A_2* » est une généralisation de la formule $(A_1^{-1} + A_2^{-1})^{-1}$ et représente la matrice résistance équivalente à l'ensemble des deux réseaux en parallèle.

De plus, nous noterons que le partage d'un courant I — qui obéit à la loi de Kirchhoff $I = I_1 + I_2$ — se fait de façon à minimiser la puissance (principe de Maxwell).

Remarque : Dans l'exposé du modèle physique, nous avons considéré des réseaux ne comportant que des résistances. Dans ce cas général (résistance + inductance + capacité) le terme de « résistance généralisée » serait à remplacer par « impédance généralisée », la matrice A étant alors hermitienne semi-définie positive.

Les résultats obtenus dans le cas complexe seraient identiques. Nous nous sommes limités ici au cas réel par souci de simplicité; pour toute la suite de l'exposé, nous nous placerons dans l'espace vectoriel $\mathbb{R}^n$ — que nous noterons X — muni du produit scalaire usuel $\langle .,. \rangle$.

II. QUELQUES OUTILS DE L'ANALYSE LINÉAIRE ET DE L'ANALYSE CONVEXE

II.1. Propriétés de base de l'inf-convolution en analyse convexe (voir [6] et [7])

Étant donné deux fonctions f et g de $X = \mathbb{R}^n$ dans $\overline{\mathbb{R}} =]-\infty, +\infty]$ *l'inf-convolution de f et g* est la fonction $f \nabla g : X \to \overline{\mathbb{R}}$ définie, pour tout x de X par :

$$f \nabla g(x) = \operatorname*{Inf}_{\substack{x_1, x_2 \in X \\ x_1 + x_2 = x}} \{ f(x_1) + g(x_2) \} .$$

L'opération d'inf-convolution est :

— commutative : $f \nabla g = g \nabla f$

— associative : $(f \nabla g) \nabla h = f \nabla (g \nabla h)$

— positivement homogène : pour tout $\alpha > 0$ $(\alpha f) \nabla (\alpha g) = \alpha) f \nabla g)$

— monotone : $f \geqslant g \Rightarrow f \nabla h \geqslant g \nabla h$.

De plus $\operatorname{dom}(f \nabla g) = \operatorname{dom} f + \operatorname{dom} g$ (où $\operatorname{dom} f = \{ x \in X / f(x) < +\infty \}$) et $f \nabla g$ est convexe dès que f et g le sont.

D'autre part, à toute fonction $f : X \to \overline{\mathbb{R}}$, non identiquement égale à $+\infty$, on associe sa *fonction conjuguée* $f^* : X \to \overline{\mathbb{R}}$ définie pour tout x^* de X par :

$$f^*(x^*) = \operatorname*{Sup}_{x \in X} \{ \langle x^*, x \rangle - f(x) \} .$$

L'opération d'inf-convolution peut être considérée comme duale de la somme, en ce sens que, si f et g sont deux fonctions de X dans $\overline{\mathbb{R}}$ non identiquement égales à $+\infty$:

1) on a toujours $(f \nabla g)^* = f^* + g^*$

2) sous certaines hypothèses (*cf.* par exemple théorème 6.5.8 de [6]) $(f+g)^* = f^* \nabla g^*$.

II.2. Pseudo-inverse d'un opérateur symétrique

La notion de pseudo-inverse d'un opérateur A de X est une généralisation de la notion d'inverse au cas où l'opérateur A n'est pas bijectif (voir [6], pp. 202-207). Dans le cas qui nous intéresse, c'est-à-dire lorsque l'opérateur A est *symétrique*, X se décompose en somme directe orthogonale : $X = \operatorname{Im} A \oplus \operatorname{Ker} A$, et la restriction $A/\operatorname{Im} A$ de l'opérateur A à son image, est un automorphisme de $\operatorname{Im} A$, donc inversible. Désignant par P la projection orthogonale de X sur le sous-espace $\operatorname{Im} A$, on obtient le schéma suivant qui permet de définir le *pseudo-inverse* A^+ de l'opérateur A :

$$\begin{tikzcd} X \arrow[r, "P"] \arrow[rrr, bend right, "A^+"'] & ImA \arrow[r, "(A/ImA)^{-1}"] & ImA \arrow[r, hook] & X \end{tikzcd}$$

A^+ est un opérateur symétrique de X ayant même image et même noyau que A et vérifiant les propriétés suivantes (qui d'ailleurs caractérisent A^+) :

$$\begin{cases} AA^+ = A^+A = P \\ A^+AA^+ = A^+ . \end{cases}$$

De plus, $(A^+)^+ = A$ et A^+ n'est autre que l'inverse A^{-1} lorsque A est bijectif.

II.3. Conjuguée d'une forme quadratique (resp. partiellement quadratique) convexe

Un opérateur symétrique A de X est dit *semi-défini positif* lorsque $\langle Ax, x \rangle \geqslant 0$ pour tout x de X, ou, de façon équivalente, lorsque la forme quadratique associée $f : X \to \mathbb{R}$, définie par $f(x) = 1/2 \langle Ax, x \rangle$, est une fonction convexe.

Lorsque A, symétrique semi-défini positif (s.s.d.p.), est inversible, on vérifie facilement que pour un x^* donné de X, la fonction concave qui à x associe $\langle x^*, x \rangle - 1/2 \langle Ax, x \rangle$ atteint son maximum pour $x = A^{-1} x^*$. On en déduit que :

$$\forall x^* \in X \quad f^*(x^*) = \frac{1}{2} \langle A^{-1} x^*, x^* \rangle .$$

Dans le cas général, la conjuguée d'une forme quadratique convexe est donnée par (voir [8], p. 108) :

THÉORÈME II.3.1 : *Soit A un opérateur s.s.d.p. et $f(x) = 1/2 \langle Ax, x \rangle$. Alors* :

$$f^*(x^*) = \begin{cases} \frac{1}{2} \langle A^+ x^*, x^* \rangle & si \quad x^* \in \operatorname{Im} A \\ +\infty & si \quad x^* \notin \operatorname{Im} A . \end{cases}$$

Exemple 1 : Prenons pour A la projection orthogonale sur un sous-espace vectoriel H de X. Alors A est s.s.d.p., $A^+ = A$ et $\operatorname{Im} A = H$. Donc :

$$f^*(x^*) = \begin{cases} \frac{1}{2} \| x^* \|^2 & \text{si} \quad x^* \in H \\ +\infty & \text{si} \quad x^* \notin H . \end{cases}$$

Exemple 2 : Prenons pour A un opérateur s.s.d.p. de rang 1. $\operatorname{Im} A$ étant une droite vectorielle, $A/\operatorname{Im} A$ est une homothétie de rapport $\alpha > 0$. Donc :

$$f^*(x^*) = \begin{cases} \frac{1}{2\alpha} \| x^* \|^2 & \text{si} \quad x^* \in \operatorname{Im} A \\ +\infty & \text{si} \quad x^* \notin \operatorname{Im} A . \end{cases}$$

Remarquons que, lorsque A est s.s.d.p., il en est de même de son pseudo-inverse, donc la conjuguée d'une forme quadratique convexe est un exemple de ce que nous appellerons « forme partiellement quadratique convexe » selon la définition suivante :

DÉFINITION II.3.2 : *Nous dirons qu'une fonction $h = X \to \overline{\mathbb{R}}$ est une forme partiellement quadratique convexe* (*) *s'il existe un opérateur A, s.s.d.p. et un sous-espace vectoriel H de X tels que* :

$$h(x) = \begin{cases} \frac{1}{2} \langle Ax, x \rangle & si \quad x \in H \\ +\infty & si \quad x \notin H . \end{cases}$$

(*) Les_formes partiellement quadratiques convexes sur $\mathbb{R}^n$ sont exactement les fonctions h $\mathbb{R}^n \to \mathbb{R}$ pour lesquelles le graphe du sous-différentiel ∂h est un sous-espace vectoriel de $\mathbb{R}^n \times \mathbb{R}^n$

La conjuguée d'une telle fonction est donnée par :

THÉORÈME II.3.3 : *Soit A un opérateur s.s.d.p., H un sous-espace vectoriel de X et h la forme partiellement quadratique convexe définie par :*

$$h(x) = \begin{cases} \frac{1}{2}\langle Ax, x\rangle & si \quad x \in H \\ +\infty & si \quad x \notin H. \end{cases}$$

Alors :

$$h^*(x^*) = \begin{cases} \frac{1}{2}\langle (\pi A \pi)^+ x^*, x^* \rangle & si \quad x^* \in \operatorname{Im} A + H^\perp \\ +\infty \text{ ailleurs}. \end{cases}$$

où π est la projection orthogonale sur H et $H^\perp$ le sous-espace orthogonal à H.

Exemple : Prenons :

$$h(x) = \begin{cases} \frac{1}{2}\| x \|^2 & \text{si} \quad x \in H \\ +\infty & \text{si} \quad x \notin H. \end{cases}$$

On obtient :

$$h^*(x^*) = \frac{1}{2}\langle \pi x^*, x^* \rangle \quad \forall x^* \in X,$$

résultat déjà vu (exemple 1 du théorème II.3.1).

Preuve du théorème : Si f est la forme quadratique associée à A, nous pouvons écrire $h = f + \psi_H$ où ψ_H est la fonction indicatrice de H (qui vaut 0 sur H et $+\infty$ ailleurs).

Les fonctions f et ψ_H sont convexes, l'une partout finie et continue, l'autre finie en 0 et semi-continue inférieurement, donc (voir [6], théorème 6.5.8) :

$$h^* = (f + \psi_H)^* = f^* \nabla \psi_H^* = f^* \nabla \psi_{H^\perp}.$$

En particulier, dom h^* = dom f^* + dom $\psi_H^\perp = \operatorname{Im} A + H^\perp$ (cf. théorème 2.3.1) donc : $\forall x^* \notin \operatorname{Im} A + H^\perp \quad h^*(x^*) = +\infty.$

Considérons maintenant un élément x^* de dom h^* :

$$\begin{aligned} h^*(x^*) &= \operatorname{Inf}\{ f^*(y^*) ; y^* \in \operatorname{Im} A \text{ et } x^* - y^* \in H^\perp \} \\ &= \operatorname{Inf}\left\{ \frac{1}{2}\langle A^+ y^*, y^* \rangle ; y^* \in \operatorname{Im} A \text{ et } x^* - y^* \in H^\perp \right\}. \end{aligned}$$

Premier cas : Supposons $\operatorname{Im} A \subset H$.

Alors $H^{\perp} \subset \operatorname{Ker} A$. Le seul élément y^* qui vérifie à la fois $y^* \in \operatorname{Im} A$ et $x^* - y^* \in H^{\perp}$ est la projection Px^* de x^* sur $\operatorname{Im} A$. Nous en déduisons que :

$$h^*(x^*) = \frac{1}{2}\langle A^+ Px^*, Px^* \rangle = \frac{1}{2}\langle A^+ x^*, x^* \rangle$$

car :

$$x^* - Px^* \in \operatorname{Ker} A^+ = (\operatorname{Im} A^+)^{\perp}.$$

Cas général : Il suffit de se ramener au cas ci-dessus en remarquant que l'on peut écrire :

$$h(x) = \begin{cases} \frac{1}{2}\langle \pi A \pi x, x \rangle & \text{si} \quad x \in H \\ +\infty & \text{si} \quad x \notin H \end{cases}$$

où maintenant l'opérateur $\pi A \pi$ est s.s.d.p. et vérifie $\operatorname{Im}(\pi A \pi) \subset H$. Le résultat précédent prouve que :

$$\underline{\forall x^* \in \operatorname{Im} A + H^{\perp} \quad h^*(x^*) = \frac{1}{2}\langle (\pi A \pi)^+ x^*, x^* \rangle}$$

formule qui englobe le cas $\operatorname{Im} A \subset H$, car alors $\pi A \pi = A$. □

III. ADDITION PARALLÈLE D'OPÉRATEURS S.S.D.P. : INTERPRÉTATION DU POINT DE VUE ANALYSE CONVEXE

Soient A et B deux opérateurs s.s.d.p. de X :

$$f(x) = \frac{1}{2}\langle Ax, x \rangle \quad \text{et} \quad g(x) = \frac{1}{2}\langle Bx, x \rangle$$

les formes quadratiques associées.

Étudions l'inf-convolution de f et g en un point x de X :

$$f \nabla g(x) = \operatorname*{Inf}_{\substack{x_1, x_2 \in X \\ x_1 + x_2 = x}} \{ f(x_1) + g(x_2) \} = \operatorname*{Inf}_{x_1 \in X} \{ f(x_1) + g(x - x_1) \}.$$

Une condition suffisante pour avoir l'exactitude de cette inf-convolution en x (c'est-à-dire pour que l'Inf ci-dessus soit atteint) est l'existence d'un point x_1 de X tel que $\partial f(x_1) \cap \partial g(x - x_1) \neq \varnothing$ (voir [6] proposition 6.6.3). Or, ici, f et g sont différentiables de gradient A et B resp. (car A et B sont symétriques), en sorte que $\partial f(x_1) = \{ Ax_1 \}$ et $\partial g(x - x_1) = \{ B(x - x_1) \}$. Donc la condition $\partial f(x_1) \cap \partial g(x - x_1) \neq \varnothing$ sera réalisée si et seulement si :

$$Ax_1 = B(x - x_1) \quad \text{ou encore} \quad \underline{(A + B)\, x_1 = Bx (*)}\,.$$

S'il existe un élément x_1 vérifiant (*), nous pourrons alors écrire :

$$\begin{aligned} f \nabla g(x) &= f(x_1) + g(x - x_1) \\ &= \frac{1}{2} \langle Ax_1, x_1 \rangle + \frac{1}{2} \langle B(x - x_1), x - x_1 \rangle \\ &= \frac{1}{2} \langle (A + B)\, x_1 - Bx, x_1 \rangle + \frac{1}{2} \langle B(x - x_1), x \rangle \end{aligned}$$

c'est-à-dire :

$$f \nabla g(x) = \frac{1}{2} \langle Ax_1, x \rangle .$$

Or $\operatorname{Im} B \subset \operatorname{Im} A + \operatorname{Im} B$. D'autre part, A et B étant s.s.d.p.,

$$\operatorname{Im}(A + B) = \operatorname{Im} A + \operatorname{Im} B .$$

Donc l'élément Bx est dans $\operatorname{Im}(A + B)$, ce qui implique (cf. II.2) :

$$(A + B)(A + B)^+ Bx = Bx .$$

Par conséquent, l'élément $\underline{x_1 = (A + B)^+ Bx}$ vérifie la relation (*), donc réalise l'exactitude de l'inf-convolution en x, et le calcul qui précède prouve que :

$$\boxed{f \nabla g(x) = \frac{1}{2} \langle A(A + B)^+ Bx, x \rangle \quad \forall x \in X} .$$

L'inf-convolution $f \nabla g$ est donc la forme quadratique associée à l'opérateur $\underline{A(A + B)^+ B}$ que nous appellerons « *somme parallèle de A et B* » et noterons « $A /\!\!/ B$ ».

Remarques :

1) L'opérateur $A /\!\!/ B$ généralise $(A^{-1} + B^{-1})^{-1}$ puisque, dans le cas où A et B sont inversibles. $A /\!\!/ B = A(A + B)^{-1} B = (A^{-1} + B^{-1})^{-1}$.

2) Il est clair que le principe de Maxwell de minimisation de la puissance (cf. I) est une manifestation de l'inf-convolution des formes quadratiques.

En effet, si nous reprenons les notations utilisées pour un réseau électrique d'ordre n, la puissance est représentée par le produit scalaire $\langle V, I \rangle$ c'est-à-dire par la forme quadratique $\langle AI, I \rangle$.

Les propriétés qui vont suivre sont celles de l'addition parallèle obtenues en utilisant, d'une manière systématique, les propriétés de l'inf-convolution rappelées en II.1. Elles ont été obtenues d'une manière algébrique par Anderson, Duffin et Trapp (voir [1], [2] et [4] par ex.).

THÉORÈME III.1 : *Si A et B sont deux opérateurs s.s.d.p., alors $A /\!/ B$ est un opérateur s.s.d.p. De plus :*

(i) *pour tout x de X, et tous x_1, x_2 tels que $x_1 + x_2 = x$:*

$$x_1 + x_2 = x \langle A /\!/ Bx, x \rangle \leqslant \langle Ax_1, x_1 \rangle + \langle Bx_2, x_2 \rangle$$

(ii) *pour tout x de X, il existe toujours x_1, x_2 dans X tels que :*

$$x_1 + x_2 = x \quad \text{et} \quad \langle A /\!/ Bx, x \rangle = \langle Ax_1, x_1 \rangle + \langle Bx_2, x_2 \rangle$$

(iii) *pour tous x_1, x_2 de X, l'égalité :*

$$\langle A /\!/ B(x_1 + x_2), x_1 + x_2 \rangle = \langle Ax_1, x_1 \rangle + \langle Bx_2, x_2 \rangle$$

est réalisée si et seulement si $Ax_1 = Bx_2$ auquel cas :

$$Ax_1 = Bx_2 = A /\!/ B(x_1 + x_2)\,.$$

Preuve : Le point $x_1 = (A + B)^+ Bx$ réalisant l'exactitude de l'inf-convolution en x, on sait (voir [6], théorème 6.6.5) que $\partial f \nabla g(x) = \partial f(x_1) \cap \partial g(x - x_1)$ c'est-à-dire $\partial f \nabla g(x) = \{ Ax_1 \}$. Donc, puisque Ax_1 est par définition égal à $A /\!/ Bx$, $\partial f \nabla g(x) = \{ A /\!/ Bx \}$.

Par conséquent, $x \to A /\!/ Bx$ étant le gradient d'une forme quadratique, l'opérateur $A /\!/ B$ est nécessairement symétrique : c'est l'unique opérateur symétrique définissant la forme quadratique $f \nabla g$.

Le fait que $A /\!/ B$ soit semi-défini positif résulte de la convexité de $f \nabla g$ (cf. II).

D'autre part, le (i) résulte de la définition de l'inf-convolution de f et g, le (ii) traduit le fait que cette inf-convolution est exacte en tout point x. Enfin, l'égalité $\langle A /\!/ B(x_1 + x_2), x_1 + x_2 \rangle = \langle Ax_1, x_1 \rangle + \langle Bx_2, x_2 \rangle$ du (iii) signifie que le couple (x_1, x_2) réalise l'exactitude de l'inf-convolution. Cette égalité sera donc vérifiée si et seulement si $\partial f(x_1) \cap \partial g(x_2) \neq \varnothing$ c'est-à-dire $Ax_1 = Bx_2$. Dans ce cas, on aura $\partial f \nabla g(x_1 + x_2) = \partial f(x_1) \cap \partial g(x_2)$. $A /\!/ B$ étant symétrique $\partial f \nabla g(x_1 + x_2) = \{ A /\!/ B(x_1 + x_2) \}$, donc la relation $\partial f \nabla g(x_1 + x_2) = \partial f(x_1) \cap \partial g(x_2)$ s'écrit : $A /\!/ B(x_1 + x_2) = Ax_1 = Bx_2$. □

Remarque : Lorsque deux réseaux de résistance généralisée A et B sont montés en parallèle, le courant I se partage en $I = I_1 + I_2$, avec $AI_1 = BI_2 = V$ (voir I). Le théorème précédent prouve que :

1) Le partage se fait effectivement de façon à minimiser la puissance.

2) $A/\!\!/B$ est la résistance généralisée équivalente aux deux réseaux en parallèle puisque $AI_1 = BI_2 = A/\!\!/BI$.

Conséquences immédiates du théorème III.1.

Rappelons que l'ensemble des opérateurs s.s.d.p. de X est muni de la relation d'ordre $A \geqslant B$ si et seulement si, par définition, $A - B$ est s.d.p., c'est-à-dire $\langle Ax, x \rangle \geqslant \langle Bx, x \rangle$ pour tout x. Le théorème III.1 permet d'obtenir de façon immédiate les inégalités :

1) $\underline{A/\!\!/B \leqslant A}$ avec égalité si et seulement si $A = 0$.

2) $\underline{A/\!\!/B \leqslant 1/4(A + B)}$ avec égalité si et seulement si $A = B$.

En particulier $A/\!\!/A = 1/2\,A$. Plus généralement on peut montrer par récurrence que :

$$\overset{n}{/\!\!/} A = \underbrace{A/\!\!/A/\!\!/ \ldots /\!\!/A}_{n \text{ fois}} = \frac{1}{n} A .$$

COROLLAIRE III.2 : *Soient $A/\!\!/B$ deux opérateurs s.s.d.p. et x un point de X. Si $Ax = ax$ et $Bx = bx$ $(a, b \geqslant 0)$ alors* :

$$A/\!\!/Bx = (a/\!\!/b)\,x .$$

Preuve : Lorsque $a = b = 0$, le (i) du théorème III.1 montre que

$$\langle A/\!\!/Bx, x \rangle = 0$$

donc, $A/\!\!/B$ étant s.s.d.p., $A/\!\!/Bx = 0 = (a/\!\!/b)\,x$.

Supposons a et b non tous deux nuls.

$A(bx) = B(ax) = abx$, donc d'après (iii) : $A/\!\!/B(ax + bx) = abx$ ou encore puisque $a + b \neq 0$, $A/\!\!/Bx = \dfrac{ab}{a + b}\,x = (a/\!\!/b)\,x$. □

COROLLAIRE III.3 : *Soit S_n le cône convexe des opérateurs symétriques définis positifs sur $\mathbb{R}^n$ et $\phi : S_n \times \mathbb{R}^n \to \mathbb{R}$ l'application définie par* :

$$\phi(A, x) = \langle A^{-1} x, x \rangle \quad \textit{pour tout} \quad (A, x) \subset S_n \times \mathbb{R}^n .$$

Alors ϕ est convexe en tant que fonction du couple (A, x).

Preuve : Il est clair que ϕ est positivement homogène, donc il suffit de vérifier sa sous-additivité.

$$\begin{aligned}\phi(A_1 + A_2, x_1 + x_2) &= \langle (A_1 + A_2)^{-1}(x_1 + x_2), x_1 + x_2 \rangle \\ &= \langle A_1^{-1} /\!/ A_2^{-1}(x_1 + x_2), x_1 + x_2 \rangle \\ &\leqslant \langle A_1^{-1} x_1, x_1 \rangle + \langle A_2^{-1} x_2, x_2 \rangle \quad \text{d'après III.1}.\end{aligned}$$

donc :

$$\phi(A_1 + A_2, x_1 + x_2) \leqslant \phi(A_1, x_1) + \phi(A_2, x_2) \qquad \text{C.Q.F.D.}$$

Remarque : On pourrait penser étendre le corollaire III.3 au cône convexe $\tilde{S}_n$ des opérateurs s.s.d.p. de $\mathbb{R}^n$ en remplaçant l'inverse A^{-1} par le pseudo-inverse A^+ ; mais le résultat ne subsiste pas, l'application ainsi obtenue n'étant plus sous-additive.

Observons enfin que, lorsque A et B sont deux opérateurs symétriques (et seulement dans ce cas), nous avons l'équivalence :

$$A = B \Leftrightarrow \forall x \in X \quad \langle Ax, x \rangle = \langle Bx, x \rangle .$$

Donc, la somme parallèle de deux opérateurs s.s.d.p. étant s.s.d.p., toutes les égalités (ou inégalités) montrées sur les opérateurs pourront l'être en raisonnant sur les formes quadratiques associées.

THÉORÈME III.4 : *L'addition parallèle d'opérateurs s.s.d.p. est commutative, associative et compatible avec l'ordre.*

Preuve : Ceci découle de la commutativité, l'associativité et la monotonie de l'inf-convolution (cf. II.1).

En effet si f, g, h sont les formes quadratiques associées à trois opérateurs s.s.d.p., A, B, C respectivement, on sait que $f \nabla g = g \nabla f$ et $(f \nabla g) \nabla h = f \nabla (g \nabla h)$ ce qui se traduit par $\underline{A /\!/ B = B /\!/ A \text{ et } (A /\!/ B) /\!/ C = A /\!/ (B /\!/ C)}$.

D'autre part, si $f \geqslant g$, alors $f \nabla h \geqslant g \nabla h$ donc $\underline{A \geqslant B \text{ implique } A /\!/ C \geqslant B /\!/ C}$.

THÉORÈME III.5 : *Soient A et B deux opérateurs s.s.d.p. Alors* :

$$\operatorname{Im} A /\!/ B = \operatorname{Im} A \cap \operatorname{Im} B .$$

Preuve : Désignons par f et g les formes quadratiques associées. Nous avons vu au (II.3) que :

$$f^*(x^*) = \begin{cases} \frac{1}{2} \langle A^+ x^*, x^* \rangle \; x^* \in \operatorname{Im} A \\ +\infty \quad \text{si} \quad x^* \notin \operatorname{Im} A . \end{cases}$$

En particulier dom $f^* = \operatorname{Im} A$. De même, dom $g^* = \operatorname{Im} B$ et, puisque $A /\!/ B$ est l'opérateur s.s.d.p. associé à la forme quadratique $f \nabla g$,

$$\operatorname{dom} (f \nabla g)^* = \operatorname{Im} A /\!/ B .$$

Or, par ailleurs, $(f \nabla g)^* = f^* + g^*$, ce qui prouve que :

$$\operatorname{dom} (f \nabla g)^* = \operatorname{dom} f^* \cap \operatorname{dom} g^*$$

d'où le résultat annoncé. □

Le Théorème III.5 prouve en particulier que la formule :

$$A /\!/ B = (A^{-1} + B^{-1})^{-1}$$

du cas inversible ne peut être étendue au cas non inversible sous la forme $A /\!/ B = (A^+ + B^+)^+$.

La « bonne » extension est la suivante :

THÉORÈME III.6 : *Si A et B sont deux opérateurs s.s.d.p., alors* :

$$A /\!/ B = (\pi(A^+ + B^+)\,\pi)^+$$

où π est la projection orthogonale sur $\operatorname{Im} A \cap \operatorname{Im} B$.

Preuve : $f \nabla g$ étant une fonction convexe continue partout finie est égale à sa biconjuguée $(f \nabla g)^{**} = (f^* + g^*)^*$.

Or :

$$(f^* + g^*)(x^*) = \begin{cases} \frac{1}{2}\langle (A^+ + B^+)\,x^*, x^* \rangle & \text{si} \quad x^* \in \operatorname{Im} A \cap \operatorname{Im} B \\ +\infty \quad \text{si} \quad x^* \notin \operatorname{Im} A \cap \operatorname{Im} B . \end{cases}$$

$f \nabla g$ est donc la conjuguée de la forme partiellement quadratique convexe ci-dessus. Il suffit alors d'appliquer le théorème II.3.3. □

THÉORÈME III.7 : *Soit $(H_i)_{1 \leqslant i \leqslant r}$ une famille finie de sous-espaces vectoriels de X et $(P_i)_{1 \leqslant i \leqslant r}$ les projections orthogonales correspondantes.*

La projection P sur $\bigcap_{i=1}^{i=r} H_i$ est donnée par :

$$P = r \mathop{/\!/}_{i=1}^{i=r} P_i = r(P_1 /\!/ P_2 \dots /\!/ P_r) .$$

Preuve : Notons f_i (resp. f) la forme quadratique associée à P_i (resp. P). Nous savons (cf. II.3) que :

$$f_i^*(x^*) = \begin{cases} \frac{1}{2} \| x^* \|^2 & \text{si} \quad x^* \in H_i \\ +\infty & \text{si} \quad x^* \notin H_i \end{cases}$$

et :

$$f^*(x^*) = \begin{cases} \frac{1}{2} \| x^* \|^2 & \text{si} \quad x^* \in \bigcap_{i=1}^{i=r} H_i \\ +\infty & \text{si} \quad x^* \notin \bigcap_{i=1}^{i=r} H_i . \end{cases}$$

L'égalité annoncée est équivalente à l'égalité des formes quadratiques $\frac{1}{r} f$ et $\mathop{\nabla}\limits_{i=1}^{i=r} f_i$, ou encore, chacune de ces fonctions coincidant avec sa biconjuguée, à l'égalité de leurs conjuguées.

Or, pour tout x^* :

$$\left(\mathop{\nabla}\limits_{i=1}^{i=r} f_i \right)^* (x^*) = \sum_{i=1}^{i=r} f_i^*(x^*) = \begin{cases} \frac{r}{2} \| x^* \|^2 & \text{si} \quad x^* \in \bigcap_{i=1}^{i=r} H_i \\ +\infty & \text{si} \quad x^* \notin \bigcap_{i=1}^{i=r} H_i \end{cases}$$

$$= r f^*(x^*) = \frac{1}{r} f^*(r x^*) = \left(\frac{1}{r} f \right)^* (x^*) \qquad \text{C.Q.F.D.}$$

THÉORÈME III.8 : *Soient A, B, C trois opérateurs s.s.d.p., tels que $AC = CA$ et $BC = CB$. Alors* :

$$AC /\!\!/ BC = (A /\!\!/ B)\, C .$$

Preuve : Remarquons d'abord qu'un opérateur quelconque qui commute avec un opérateur symétrique, commute aussi avec son pseudo-inverse. En conséquence, C commutant avec A et B, commute avec $(A + B)^+$, donc aussi avec $A /\!\!/ B$.

D'autre part, C étant s.s.d.p., il existe un opérateur s.s.d.p., noté $C^{1/2}$, tel que $C^{1/2} C^{1/2} = C$ et $C^{1/2}$ commute avec tout opérateur commutant avec C, en particulier avec A, B et $A /\!\!/ B$.

Soit x un point de X. Le théorème III.1 nous donne l'existence de points x_1, x_2 tels que :

$$\langle AC /\!\!/ BCx, x \rangle = \langle ACx_1, x_1 \rangle + \langle BCx_2, x_2 \rangle . \tag{1}$$

Ces points vérifient :

$$\begin{cases} x_1 + x_2 = x \\ ACx_1 = BCx_2 . \end{cases}$$

L'égalité $ACx_1 = BCx_2$ peut s'écrire $C^{1/2}[C^{1/2} Ax_1] = C^{1/2}[C^{1/2} Bx_2]$. La restriction de $C^{1/2}$ à son image étant injective, ceci implique :

$$C^{1/2} Ax_1 = C^{1/2} Bx_2 .$$

Nous avons par conséquent :

$$\begin{cases} AC^{1/2} x_1 = BC^{1/2} x_2 \\ C^{1/2} x_1 + C^{1/2} x_2 = C^{1/2} x \end{cases}$$

ce qui prouve, toujours grâce au théorème III.1, que :

$$\langle A /\!/ B(C^{1/2} x), C^{1/2} x \rangle = \langle A(C^{1/2} x_1), C^{1/2} x_1 \rangle + \langle B(C^{1/2} x_2), C^{1/2} x_2 \rangle$$

$C^{1/2}$ commutant avec A, B et $A /\!/ B$, cette égalité devient :

$$\langle (A /\!/ B)\, Cx, x \rangle = \langle ACx_1, x_1 \rangle + \langle BCx_2, x_2 \rangle . \tag{2}$$

Il suffit alors de comparer (1) et (2) pour conclure. □

Remarque : L'hypothèse $AC = CA$ est indispensable pour que l'opérateur AC soit symétrique. Plus généralement on a le résultat suivant :

THEOREME III.9 : *Soient A et B deux opérateurs s.s.d.p. et Z un opérateur quelconque. Alors :*

$${}^TZ(A /\!/ B)\, Z \leqslant ({}^TZAZ) /\!/ ({}^TZBZ)$$

avec égalité lorsque Z est inversible.

Preuve : Pour A par exemple, on a :

$$\frac{1}{2} \langle {}^TZAZx, x \rangle = \frac{1}{2} \langle AZx, Zx \rangle = f(Zx) .$$

Il s'agit donc de comparer :

$$f \nabla g(Zx) = \operatorname*{Inf}_{y_1 + y_2 = Zx} \{ f(y_1) + g(y_2) \} \quad \text{et} \quad \operatorname*{Inf}_{x_2 + x_2 = x} \{ f(Zx_1) + g(Zx_2) \} .$$

Or il est clair que :

$$\{ f(y_1) + g(y_2)/y_1 + y_2 = Zx \} \quad \text{contient} \quad \{ f(Zx_1) + g(Zx_2)/x_1 + x_2 = x \}$$

avec égalité dès que Z est inversible, d'où le résultat. □

THEOREME III.10 : *Soit* $(A_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant q}}$ *une famille d'opérateurs s.s.d.p.*

Alors :

$$\mathop{/\!\!/}_{i=1}^{i=p} \sum_{j=1}^{j=q} A_{ij} \geqslant \sum_{j=1}^{j=q} \mathop{/\!\!/}_{i=1}^{i=p} A_{ij} .$$

Preuve : Nous avons à monter la relation suivante :

$$\left(\sum_{j=1}^{j=q} A_{1j}\right) /\!\!/ \left(\sum_{j=1}^{j=q} A_{2j}\right) /\!\!/ \cdots /\!\!/ \left(\sum_{j=1}^{j=q} A_{pj}\right) \geqslant \sum_{j=1}^{j=q} (A_{1j} /\!\!/ A_{2j} /\!\!/ \cdots /\!\!/ A_{pj})$$

ou, de façon équivalente et avec des notations évidentes :

$$\left(\sum_{j=1}^{j=q} f_{1j}\right) \nabla \left(\sum_{j=1}^{j=q} f_{2j}\right) \nabla \dots \nabla \left(\sum_{j=1}^{j=q} f_{pj}\right) > \sum_{j=1}^{j=q} (f_{1j} \nabla f_{2j} \dots \nabla f_{pj}) .$$

Cette inégalité est une propriété de l'inf-convolution qui s'obtient sans difficulté en remarquant que si $x = x_1 + \cdots + x_p$, alors :

$$\sum_{j=1}^{j=q} f_{1j}(x_1) + \sum_{j=1}^{j=q} f_{2j}(x_2) + \cdots + \sum_{j=1}^{j=q} f_{pj}(x_p) =$$
$$= \sum_{j=1}^{j=q} [f_{1j}(x_1) + f_{2j}(x_2) + \cdots + f_{pj}(x_p)] \geqslant \sum_{j=1}^{j=q} (f_{1j} \nabla f_{2j} \nabla \dots \nabla f_{pj})(x) . \quad □$$

COROLLAIRE III.11 : *Si* $A_1, A_2, \dots, A_r$ *sont s.s.d.p., alors* :

$$A_1 + A_2 + \cdots + A_r \geqslant r^2 (A_1 /\!\!/ A_2 /\!\!/ \cdots /\!\!/ A_r) .$$

Preuve :

$$\frac{1}{r}(A_1 + \cdots + A_r) = /\!\!/ (A_1 + \cdots + A_r)$$
$$= (A_1 + A_2 + \cdots + A_r) /\!\!/ (A_2 + A_3 + \cdots + A_1) /\!\!/ \cdots /\!\!/ (A_r + A_1 + \cdots + A_r)$$
$$\geqslant r(A_1 /\!\!/ A_2 /\!\!/ \cdots /\!\!/ A_r) \quad \text{d'après le théorème III.10} . \; □$$

COROLLAIRE III.12 : *L'addition parallèle est une application concave de* $\tilde{S}_n \times \tilde{S}_n$ *dans* $\tilde{S}_n$.

Preuve : Soit $\psi : \tilde{S}_n \times \tilde{S}_n \to \tilde{S}_n$ définie par $\psi(A, B) = A /\!\!/ B$.
ψ est positivement homogène car l'inf-convolution l'est. Donc pour montrer la concavité de ψ, il suffit de vérifier qu'elle est suradditive, c'est-à-dire que, pour tous A_1, A_2, B_1, B_2 de $\tilde{S}_n$:

$$\psi(A_1 + A_2, B_1 + B_2) \geqslant \psi(A_1, B_1) + \psi(A_2, B_2)$$

ou encore $(A_1 + A_2) /\!\!/ (B_1 + B_2) \geqslant A_1 /\!\!/ B_1 + A_2 /\!\!/ B_2$ ce qui résulte du théorème II.7. □

Interprétons ce résultat d'un point de vue physique (voir [5]).

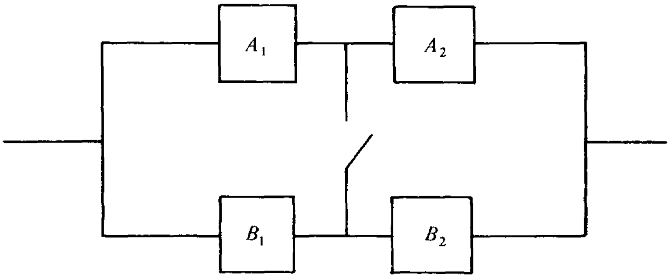


Figure 1.

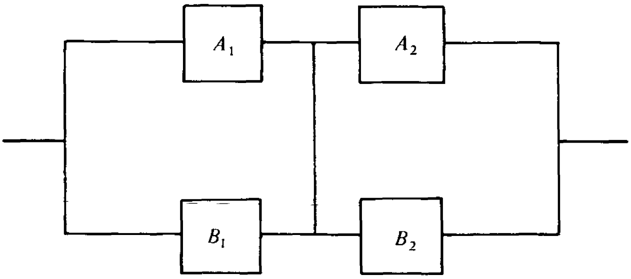


Figure 2.

La matrice résistance équivalente au système est $(A_1 + A_2) /\!\!/ (B_1 + B_2)$ dans le cas de la figure 1, et $A_1 /\!\!/ B_1 + A_2 /\!\!/ B_2$ dans le cas de la figure 2.

La sous-additivité de la fonction ψ nous dit que la puissance est moindre dans la deuxième configuration.

THEOREME III.13. — *Si A et B sont s.s.d.p., alors $\| A /\!\!/ B \| \leqslant \| A \| /\!\!/ \| B \|$.*

Lemme : Pour tout opérateur s.s.d.p. non nul A, $\| A \| = \operatorname{Sup}_{Ax \neq 0} \dfrac{\| Ax \|^2}{\langle Ax, x \rangle}$.

$\| A \| = \operatorname{Sup}_{Ax \neq 0} \cdot \dfrac{\| Ax \|}{\| x \|}$. Nous savons que $\langle Ax, x \rangle = 0$ si et seulement si $Ax = 0$ et que $\langle Ax, x \rangle \leqslant \| Ax \| \| x \|$. Donc :

$$\operatorname{Sup}_{Ax \neq 0} \cdot \frac{\| Ax \|^2}{\langle Ax, x \rangle} \geqslant \operatorname{Sup}_{Ax \neq 0} \cdot \frac{\| Ax \|^2}{\| Ax \| \| x \|} = \operatorname{Sup}_{Ax \neq 0} \cdot \frac{\| Ax \|}{\| x \|} = \| A \| .$$

Pour obtenir l'inégalité en sens inverse, remarquons que pour $Ax \neq 0$:

$$\frac{\| Ax \|^2}{\langle Ax, x \rangle} = \frac{\langle AA^{1/2} x, A^{1/2} x \rangle}{\langle A^{1/2} x, A^{1/2} x \rangle} \leqslant \frac{\| A \| \| A^{1/2} x \|^2}{\| A^{1/2} x \|^2} = \| A \| .$$

Preuve du théorème III.13.

L'inégalité $\| A /\!\!/ B \| \leqslant \| A \| /\!\!/ \| B \|$ est trivalement vérifiée lorsque $\| A /\!\!/ B \| = 0$. Plaçons-nous dans l'hypothèse où $\| A /\!\!/ B \| \neq 0$ (auquel cas $\| A \| \neq 0$ et $\| B \| \neq 0$).

Fixons un ε tel que $0 < \varepsilon < \| A /\!\!/ B \|$. D'après le lemme précédent on peut trouver un x_ε tel que :

$$\frac{\| A /\!\!/ B x_\varepsilon \|^2}{\langle A /\!\!/ B x_\varepsilon, x_\varepsilon \rangle} \geqslant \| A /\!\!/ B \| - \varepsilon .$$

Or, (cf. théorème III.1) $\langle A /\!\!/ B x_\varepsilon, x_\varepsilon \rangle = \langle Ax_1, x_1 \rangle + \langle Bx_2, x_2 \rangle$ avec $Ax_1 = Bx_2 = A /\!\!/ B x_\varepsilon$.

Donc :

$$\frac{\langle A /\!\!/ B x_\varepsilon, x_\varepsilon \rangle}{\| A /\!\!/ B x_\varepsilon \|^2} = \frac{Ax_1, x_1 \rangle}{\| Ax_1 \|^2} + \frac{\langle Bx_2, x_2 \rangle}{\| Bx_2 \|^2} \geqslant \frac{1}{\| A \|} + \frac{1}{\| B \|}$$

ce qui implique

$$\| A /\!\!/ B \| - \varepsilon \leqslant (\| A \|^{-1} + \| B \|^{-1}) = \| A \| /\!\!/ \| B \| .$$

L'inégalité cherchée est obtenue en faisant tendre ε vers 0. □

IV. CONCLUSION

En interprétant le principe variationnel de Maxwell comme une inf-convolution sur les formes quadratiques représentant les puissances, nous avons retrouvé les propriétés de ce qu'il est convenu d'appeler l'addition parallèle d'opérateurs s.s.d.p. et que plusieurs auteurs avaient obtenues par une suite de calcul sur les matrices. L'analyse convexe s'est montrée ici un outil efficace pour exhiber ces propriétés d'une manière élégante.

Les résultats énoncés ont été démontrés en dimension finie. D'une façon plus générale, il est possible de se placer dans un cadre hilbertien (c'est le bon cadre d'étude des formes quadratiques) mais l'interprétation physique n'est pas claire. Cependant, les résultats subsistent, avec toutefois des hypothèses techniques supplémentaires, du type fermeture de l'image des opérateurs.

Enfin, nous signalons qu'il existe une autre opération associée à un opérateur s.s.d.p. et à un sous-espace vectoriel — qualifiée de « shorted operator » — Cette opération est reliée, dans une certaine mesure, à l'addition parallèle d'opérateurs et peut aussi s'interpréter *via* l'Analyse convexe. Ce sera l'objet d'une présentation ultérieure.

REFERENCES

[1] W. N. ANDERSON Jr., *Shorted Operators,* SIAM J. Appl. Math., 20 (1971) pp. 520-525.

[2] W. N. ANDERSON Jr. and R. J. DUFFIN, *Series and parallel addition of matrices,* J. Math. Anal. Appl., 26 (1969) pp. 576-594.

[3] W. N. ANDERSON Jr. and M. SHREIBER, *On the infimum of two projections,* Acta Sci Math. (Szeged), 33 (1972) pp. 165-168.

[4] W. N. ANDERSON Jr. and G. E. TRAPP, *Inequalities for the parallel connection of resistive n. port networks,* J. Franklin Inst., 5 (1975) pp. 305-313.

[5] W. N. ANDERSON Jr. and G. TRAPP, *Shorted Operators,* SIAM J. Appl. Math., 28 (1975) pp. 60-71.

[6] P. J. LAURENT, *Approximation et Optimisation,* Hermann (1972).

[7] J. J. MOREAU, *Fonctionnelles convexes,* séminaire sur les équations aux dérivées partielles II, Collège de France (1966-1967).

[8] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press (1970).