Institutional Repository - Research Portal
Dépôt Institutionnel - Portail de la Recherche
researchportal.unamur.be
University of Namur

THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Les ensembles H dans l'approximation linéaire au sens de Tchebycheff

Maquet, Anne; Bouillon, André-Marie

Award date:
1978

Awarding institution:
Universite de Namur

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 02. May. 2026

FACULTES UNIVERSITAIRES NOTRE-DAME DE LA PAIX

NAMUR

Année académique 1977-1978

LES ENSEMBLES-H DANS L'APPROXIMATION LINEAIRE AU SENS DE TCHEBYCHEFF

Mémoire présenté pour l'obtention
du grade de Licencié en Sciences
Mathématiques par

Anne MAQUET
et
André-Marie BOUILLON

Promoteur :
J.P. THIRAN

FMB1/1978/11

Nous remercions de tout coeur, Jean-Pierre THIRAN de nous avoir guidés dans ce mémoire, ainsi que Philippe DEFERT pour son aide et toutes les personnes qui nous ont permis de mener à bien ce travail.

Introduction :

Dans [3] et [4], COLLATZ a introduit le concept d'ensemble-H pour déterminer des bornes de l'erreur dans un problème d'approximation au sens de TCHEBYCHEFF. Ces ensembles-H ont ensuite été analysés de manière approfondie par divers auteurs, notamment par C.DIERIECK dans sa thèse [6].

Le présent travail est destiné à présenter une étude des ensembles-H dans le cas d'une approximation linéaire, principalement celle relative aux polynomes multivariables.

Dans un premier chapitre, on énonce la définition d'ensemble-H et le théorème d'inclusion qui s'y rapporte. Une forme duale de cette définition fournit des bornes d'erreur plus précises et conduit à un algorithme qui permet de vérifier si un ensemble de points du domaine d'approximation constitue un ensemble-H.

Les principes d'invariance décrits dans la première partie du chapitre 2, réduisent fortement le nombre d'ensembles-H à étudier. Le chapitre se termine par une description des différentes méthodes de construction des ensembles-H.

Au chapitre 3, un algorithme de construction, sans répétition, de tous les ensembles-H minimaux relatifs à une approximation par des polynomes multivariables de degré 1 est déduit de l'étude des propriétés du sous-espace linéaire de ces polynomes.

Deux méthodes rendent possible une construction par récurrence, détaillée au chapitre 4, d'ensembles-H relatifs à une approximation polynomiale de plus haut degré. En fin de chapitre, ces différentes méthodes sont appliquées à la recherche des ensembles-H relatifs à l'espace des polynomes approximants, engendré par l'ensemble $\{1, x, y, xy\}$ des fonctions de base.

Le dernier chapitre est consacré à la description de l'algorithme de Remes et de deux nouveaux algorithmes de recherche d'une meilleure approximation.

Ces derniers utilisent les configurations géométriques des ensembles-H relatifs à l'approximation polynomiale bivariable de degré 1. La présentation et la comparaison des résultats obtenus suivent cette description.

L'annexe 1 regroupe les programmes réalisés et l'annexe 2, les figures.

Ce travail a été réalisé par deux étudiants, Anne MAQUET s'attachant plus spécialement aux chapitres 1 et 3 et André-Marie BOUILLON aux chapitres 2 et 4. Quant au 5ème chapitre, il est le fruit d'une recherche commune.

CHAPITRE I :

Introduction à la théorie des ensembles-H

1.1. Position du problème d'approximation linéaire au sens de TCHEBYCHEFF.

On considère l'espace de Banach $C[B, R]$ des fonctions réelles f, continues, définies sur un domaine compact B de R pour lesquelles la norme de TCHEBYCHEFF est finie. Cette norme est donnée par :

$$||f|| = \max_{x \in B} |f(x)| \cdot p(x)$$

où p est une fonction de poids continue et strictement positive. On prend souvent $p \equiv 1$.
Soit le sous-espace linéaire engendré par p fonctions de base, indépendantes $g_1, g_2, \ldots, g_p \in C[B, R]$.
On notera ce sous-espace

$$W = \text{eng}\{g_1, g_2, \ldots, g_p\} \tag{1.1}$$

Le problème consiste à approximer une fonction objectif $f \in C[B, R] \setminus W$ par un élément de W. En d'autres mots, il faut déterminer $a^\star = (a_1^\star, a_2^\star, \ldots, a_p^\star) \in R^p$ de sorte que :

$$||f - \sum_{i=1}^{p} a_i^\star g_i|| \leqslant ||f - \sum_{i=1}^{p} a_i g_i||$$

pour tout $a = (a_1, a_2, \ldots, a_p) \in R^p$.
Par la définition de la norme de TCHEBYCHEFF, le problème d'approximation revient alors à calculer :

$$\min_{x \in R^p} \max_{x \in B} |f(x) - \sum_{i=1}^{p} a_i g_i(x)| \, p(x).$$

1.2. Définition des ensembles-H et théorème d'inclusion.

1.2.1. Définitions.

Dans [4], COLLATZ introduit la définition d'ensemble-H par le biais de l'hypothèse-H. Celle-ci s'énonce : il existe deux ensembles M_1 et M_2 de B qui vérifient la propriété suivante : aucune paire $w_1, w_2 \in W$ (1.1.) ne satisfait à

$$w_1 - w_2 > 0 \quad \text{sur } M_1$$
$$w_1 - w_2 < 0 \quad \text{sur } M_2 \qquad (1.2)$$

Définition d'ensemble :

L'ensemble $M = M_1 \cup M_2$ s'appelle ensemble-H relatif à W ssi les ensembles M_1 et M_2 satisfont à l'hypothèse-H.

Définition d'ensemble-H minimal :

L'ensemble-H M est dit minimal ssi pour tout sous-ensemble propre N de M, N n'est pas un ensemble-H.

Remarques :

1. Quand on écrira ensemble-H, cela signifiera ensemble-H relatif à W. S'il n'en est pas ainsi, mention explicite en est faite.
2. L'hypothèse-H définie ci-dessus est générale et reste valable pour le cas non linéaire. Au paragraphe1.2.3., on particularisera cette définition pour le cas linéaire.

1.2.2. Théorème d'inclusion.

COLLATZ établit aussi en [4], un théorème d'inclusion qui définit une borne supérieure et une borne inférieure à la norme de l'erreur.

Théorème 1.1 :

On considère un élément w_o *de* W (1.1)
On définit l'erreur ε *par :* $\varepsilon = w_o - f$
On suppose que les ensembles M_1 *et* M_2 *de* B *satisfont à l'hypothèse-H.*
Si en plus, $\varepsilon(.)$ *vérifie : soit* $\varepsilon(x) > 0$ (<0) *pour tout* $x \in M_1$
$\varepsilon(x) < 0$ (>0) *pour tout* $x \in M_2$

alors

$$\inf_{M_1 \cup M_2} p|\varepsilon| \leqslant p_o \leqslant ||\varepsilon||$$

où $p_o = \inf_{w \in W} ||w - f||$

Démonstration :

1. L'inégalité $p_o \leqslant ||\varepsilon||$ est immédiate par les définitions de p_o et ε.

2. On note $\sigma = \inf_{M_1 \cup M_2} p|\varepsilon|$

La seconde inégalité s'exprime alors par $\sigma \leqslant p_o$.

On raisonne par l'absurde et on suppose $\sigma > p_o$. Dès lors, il existe un nombre réel τ tel que :

$$p_o < \tau < \sigma$$

Par la définition de p_o, il existe $w \in W$ tel que :

$$||w - f|| \leqslant \tau$$

Or dans $M_1 \cup M_2$, on a $\sigma \leqslant p\,|\varepsilon|$. En regroupant ces relations on obtient dans $M_1 \cup M_2$:

$$p|\varepsilon| \geqslant \sigma > \tau \geqslant p|w - f|.$$

Comme p est strictement positif, on a :

$$|\varepsilon| - |w - f| > 0 \qquad \text{pour tout } x \in M_1 \cup M_2 \tag{1.3}$$

Pour un élément arbitraire $w \in W$, on peut écrire :

$$\begin{aligned} w_o - w &= w_o - f - (w - f) \\ &= \varepsilon - (w - f) \\ \varepsilon(w_o - w) &= \varepsilon^2 - \varepsilon(w - f) \\ \varepsilon(w_o - w) &\geqslant |\varepsilon|\,(|\varepsilon| - |w - f|) \end{aligned}$$

Or par hypothèse

$$|\varepsilon(x)| > 0 \qquad \text{pour tout } x \in M_1 \cup M_2$$

Et en appliquant le résultat obtenu en (1.3), on obtient :

$$\varepsilon(w_o - w) > 0 \quad \text{sur } M = M_1 \cup M_2 \tag{1.4}$$

D'autre part, M_1 et M_2 satisfont à l'hypothèse-H (1.2).

En remplaçant w_1 par w_o et w_2 par q, (1.2) devient :
Il n'existe aucun $q \in W$ vérifiant :

$$(w_o - q) > 0 \qquad \text{pour tout } x \in M_1$$

$$(w_o - q) < 0 \qquad \text{pour tout } x \in M_2$$

ou encore par les hypothèses faites sur les signes de ε :
Il n'existe aucun $q \in W$ vérifiant :

$$\varepsilon(w_o - q) > 0 \qquad \text{pour tout } x \in M_1 \cup M_2$$

ce qui est en contradiction avec (1.4).
On peut donc écrire $\sigma \leqslant p_o$. ▪
On en déduit le théorème d'oscillation.

Théorème 1.2 :

On considère un élément w_o fixé de W (1.1)
Si l'erreur $\varepsilon = w_o - f$ atteint sa valeur maximale sur un ensemble-H
alors w_o est solution du problème d'approximation au sens de TCHEBYCHEFF.

Remarquons que le théorème d'inclusion, appliqué au cas de l'approximation à une seule variable réelle donne les mêmes conclusions que le théorème de LA VALLEE POUSSIN [2].

Dans la fin de ce paragraphe, en se basant sur l'article [4], on va modifier pour le cas linéaire la formulation de la définition d'ensemble-H de COLLATZ.

1.2.3. Modification de la formulation de la définition d'ensemble-H.

On considère un élément w de W. Par la définition (1.1) de W, on peut écrire

$$w(.) = \sum_{j=1}^{p} a_j \, g_j(.) \qquad \text{où } (a_1, \ldots, a_p) \in R^p$$

En particulier, pour un élément $w^\star = w_1 - w_2$ de W, on obtient :

$$w^* = w_1 - w_2 = \sum_{j=1}^{p} c_j g_j \quad \text{où } (c_1, \ldots, c_p) \in \mathbf{R}^p$$

Avec cette écriture, l'hypothèse-H (1.2) s'exprime :
Il n'existe aucun p-uplet $(c_1, \ldots, c_p) \in \mathbf{R}^p$ qui satisfait à :

$$w^*(P_i) = \sum_{j=1}^{p} c_j g_j(P_i) > 0 \qquad \text{pour } P_i \in M_1$$

$$w^*(P_i) = \sum_{j=1}^{p} c_j g_j(P_i) < 0 \qquad \text{pour } P_i \in M_2$$

Si on adopte les notations suivantes :
$\varepsilon_i = +1$ pour $P_i \in M_1$
$\varepsilon_i = -1$ pour $P_i \in M_2$
$s = \text{card } M_1 \cup M_2$
$g_{ij} = g_j(P_i)$ pour $i = 1, \ldots, s$; pour $j = 1, \ldots, p$,
alors l'hypothèse-H s'écrit :
Il n'existe aucun p-uplet $(c_1, \ldots, c_p) \in \mathbf{R}^p$ qui satisfait :

$$\sum_{j=1}^{p} c_j \varepsilon_i g_{ij} > 0 \qquad \text{pour } i = 1, \ldots, s.$$

La définition d'ensemble-H s'exprime alors par :

> $M = M_1 \cup M_2$ est un ensemble-H ssi le système suivant d'inéquations
>
> $$\sum_{j=1}^{p} c_j \varepsilon_i g_{ij} > 0 \qquad \text{pour } i = 1, \ldots, s \qquad (1.5)$$
>
> est incompatible.

Pour vérifier si un ensemble donné M est un ensemble-H, on est amené à résoudre un système d'inéquations. Dans [5], COLLATZ propose comme méthode de résolution l'élimination de Gauss à facteurs positifs.

1.2.4. Elimination de Gauss à facteurs positifs.

On note le système (1.5) par $Ax > 0$. L'élimination de Gauss à facteurs positifs consiste à passer de ce système initial à un autre système d'inéquations,

noté $Bx > 0$, en multipliant des lignes de la matrice A par des termes positifs et en les additionnant, c'est-à-dire :

$$B = \begin{pmatrix} \sum_i \lambda_i A_i \\ \vdots \\ \sum_i \gamma_i A_i \end{pmatrix}$$

où $\lambda_i, \gamma_i > 0$; $\lambda_i, \gamma_i \in \mathbb{R}$,
où A_i désigne la $i^{ème}$ ligne de la matrice A.

De plus, on choisit les facteurs positifs pour générer une ligne de la matrice B composée uniquement de termes nuls. La contradiction est alors immédiate. Pour illustrer ce qui précède, on donne un exemple numérique.

Exemple : Soit la matrice

$$A = \begin{pmatrix} 2 & 3 \\ -1 & 1 \\ 0 & -2 \\ 1 & 1 \end{pmatrix}$$

On construit la matrice B en remplaçant :

$$A_3 = A_3 + A_2 + A_4.$$

Le système défini par B s'écrit :

$$Bx = \begin{pmatrix} 2 & 3 \\ -1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} > 0$$

On obtient immédiatement la contradiction $0 > 0$.

Cependant, on peut remarquer que cette méthode mène parfois à des conclusions erronées. En effet, il est aisé de montrer que le système transformé $Bx > 0$ peut admettre une solution sans que le système $Ax > 0$ n'en admette une. Pour ce, on raisonne sur le cas particulier où la matrice A est de la forme :

$$A = \begin{pmatrix} a \\ b \end{pmatrix}$$

Le système initial est donc :

$$Bx = \begin{pmatrix} a \\ b \end{pmatrix} [x] > 0$$

En notant $ax = \alpha$ et $bx = \beta$, on obtient les inégalités suivantes :

$$\alpha > 0 \quad \text{et} \quad \beta > 0$$

Le système défini par B est donc de la forme :

$$\lambda_1 \alpha + \beta_2 \beta > 0 \quad \text{où } \lambda_1, \lambda_2 > 0$$
$$\beta > 0$$

Prenons comme valeurs particulières $\lambda_1 = \lambda_2 = 1$. Si le système $Ax > 0$ admet une solution, alors le système $Bx > 0$ admet également une solution.
Cependant $\alpha + \beta > 0$ n'entraîne pas $\alpha > 0$
$\beta > 0$ $\beta > 0$

Ce qui veut dire : le système $Bx > 0$ peut admettre une solution sans que le système $Ax > 0$ n'en admette une.
De plus, si le système $Bx > 0$ contient une contradiction, alors le système $Ax > 0$ en contient également une.

Exemple : Soit la matrice

$$A = \begin{pmatrix} 1 & 0 \\ -1 & -1 \\ 1 & 2 \\ -1 & -3 \end{pmatrix}$$

On considère les transformations suivantes :

$$A_2 = A_2 + A_1$$
$$A_4 = A_4 + A_1$$

on obtient :

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 1 & 2 \\ 0 & -3 \end{pmatrix}$$

Si on remplace les lignes 2 et 4 par :

$$A_2 = A_2 + \frac{1}{2} A_3$$

$$A_4 = A_4 + \frac{3}{2} A_3$$

la matrice ainsi obtenue ne possède plus d'élément négatif

$$B = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 0 \\ 1 & 2 \\ \frac{3}{2} & 0 \end{pmatrix}$$

Le système $Bx > 0$ est compatible, il admet une solution. Cependant, si on permute deux lignes de la matrice A, on obtient A^*.

$$A^* = \begin{pmatrix} -1 & -1 \\ 1 & 2 \\ 1 & 0 \\ -1 & -3 \end{pmatrix}$$

Le système ainsi obtenu est équivalent au système initial $Ax > 0$. En considérant les transformations suivantes :

$$A_1^* = A_1^* + A_2^*,$$
$$A_4^* = A_4^* + A_2^*,$$

on obtient :

$$\begin{pmatrix} 0 & 1 \\ 1 & 2 \\ 1 & 0 \\ 0 & -1 \end{pmatrix}$$

La contradiction apparaît en remplaçant A_4^* par $A_4^* + A_1^*$,

$$\begin{pmatrix} 0 & 1 \\ 1 & 2 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Le système transformé contient une contradiction, donc le système initial en contient une.

1.3. Formulation duale de la définition de ensemble-H.

On sait maintenant que pour déterminer si un ensemble est un ensemble-H, on est amené à résoudre un système (1.5) sans contrainte de signe sur la solution. En reprenant la notation $Ax > 0$, le dual de ce problème s'écrit :

$$\begin{aligned} A'y &= 0 \\ y &\geqslant 0 \end{aligned} \qquad \text{où } A' \text{ : transposée de la matrice } A$$

Dans [11], le théorème de GORDAN éclaire sur l'existence des solutions simultanées au système initial et à son dual. La démonstration de ce théorème est extraite de [6].

1.3.1. Théorème de GORDAN 1.3 :

Soit le système (I) : $Ax > 0$ *où* $A \in \mathbb{R}^{p \times s}$

$x \gneqq 0$

le système dual (II) : $A'y = 0$

$y \geqslant 0$

ou bien (I) admet une solution,

ou bien (II) admet une solution,

mais les deux systèmes (I) et (II) ne peuvent admettre simultanément une solution.

Démonstration :

En notant $(\overline{I})$: le système (I) n'a pas de solution,

(I) : le système (I) admet une solution,

le théorème peut s'exprimer comme suit : (I) ssi $(\overline{II})$ ou $(\overline{I})$ ssi (II).

Il suffit donc de démontrer : $(\overline{I})$ ssi (II).

Or $(\bar{I})$ est équivalent à : pour tout $(x_1, \ldots, x_s) \in R^s$, le système suivant

$$\sum_{i=1}^{s} x_i \, g_{ik} > 0 \qquad \text{pour } k = 1, \ldots, p$$

est incompatible.

(II) est équivalent à : il existe $(y_1, \ldots, y_p) \in R^p$, non tous nuls,

$$y_i \geqslant 0 \qquad \text{pour } i = 1, \ldots, p \text{ et tel que :}$$

$$\sum_{k=1}^{p} y_k \, g_{jk} = 0 \qquad \text{pour } j = 1, \ldots, s.$$

1. Condition suffisante : (II) $\Rightarrow$ $(\bar{I})$.

Comme (II) est équivalent à

$$\sum_{k=1}^{p} y_k \, g_{jk} = 0 \qquad \text{pour } j = 1, \ldots, s,$$

on obtient en multipliant chaque terme par un x_j :

$$\sum_{j=1}^{s} x_j \sum_{k=1}^{p} y_k \, g_{jk} = 0$$

En permutant les signes de sommation, on a :

$$\sum_{k=1}^{p} y_k \sum_{j=1}^{s} x_j \, g_{jk} = 0.$$

Or on a supposé qu'il existe un indice k_o tel que $y_{k_o} > 0$, donc les termes $\sum_{j=1}^{s} x_j \, g_{jk}$ ne peuvent être tous strictement positifs. Ce qui entraîne $(\bar{I})$.

2. Condition nécessaire : $(\bar{I})$ $\Rightarrow$ (II).

Soit un compact C de R^p donné par :

$$C = \{y \in R^p \text{ tel que } y \geqslant 0 \text{ et } \sum_{i=1}^{p} y_i = 1\}$$

On considère la fonction suivante v définie sur ce compact C

$$v(y_1, \ldots, y_p) = \sum_{j=1}^{s} \Big| \sum_{k=1}^{p} y_k \, g_{jk} \Big|^2$$

Comme v est une fonction continue sur un compact, elle atteint sa valeur minimale en un point de ce compact. On note ce point :

$$y^o = (y_1^o, \ldots, y_p^o)$$

On remarque que (II) est alors équivalent à :

$$v(y_1^o, \ldots, y_p^o) = 0$$

En raisonnant par l'absurde, on suppose :

$$(\bar{I}) \text{ et } v(y_1^o, \ldots, y_p^o) > 0$$

On note : $c_j = \sum_{k=1}^{p} y_k^o \, g_{jk}$ pour $j = 1, \ldots, s$.

L'hypothèse $(\bar{I})$ nous permet d'écrire :

$$\sum_{j=1}^{s} c_j \, g_{jp} \leqslant 0$$

En adoptant les notations suivantes :

$$\alpha = \sum_{j=1}^{s} |g_{jp}|^2$$

$$\lambda = \frac{\alpha}{\alpha + v(y_1^o, \ldots, y_p^o)}$$

$$y_k^1 = \lambda \, y_k^o \qquad \text{pour } k = 1, \ldots, p-1$$

$$y_p^1 = (1 - \lambda) + \lambda \, y_p^o,$$

on peut écrire :

$$\begin{aligned} v(y_1^1, \ldots, y_p^1) &= \sum_{j=1}^{s} \Big| \sum_{k=1}^{p} y_k^1 \, g_{jk} \Big|^2 \\ &= \sum_{j=1}^{s} \Big| (1 - \lambda) g_{jp} + \lambda \sum_{k=1}^{p} y_k^o \, g_{jk} \Big|^2 \\ &= (1 - \lambda)^2 \sum_{j=1}^{s} |g_{jp}|^2 \\ &\quad + 2 \lambda (1 - \lambda) \sum_{j=1}^{s} \sum_{k=1}^{p} y_k^o \, g_{jk} \, g_{jp} \\ &\quad + \lambda^2 \sum_{j=1}^{s} \Big| \sum_{k=1}^{p} y_k^o \, g_{jk} \Big|^2 \\ &\leqslant (1 - \lambda)^2 \, \alpha + \lambda^2 \, v(y_1^o, \ldots, y_p^o) \end{aligned}$$

$$\leqslant \frac{\alpha v(y_1^o, \ldots, y_p^o)}{\alpha + v(y_1^o, \ldots, y_p^o)}$$

$$v(y_1^1, \ldots, y_p^1) \leqslant v(y_1^o, \ldots, y_p^o).$$

Ce qui contredit le fait que $(y_1^o, \ldots, y_p^o)$ est le point où v atteint sa valeur minimale. Donc,

$$v(y_1^o, \ldots, y_p^o) = 0. \blacksquare$$

Par ce théorème on peut modifier l'expression de la définition d'ensemble-H comme suit :

M est un ensemble-H ssi $Ax > 0$ n'admet pas de solution
$x \gneqq 0$
ssi $A'p = 0$ admet une solution
$p \geqslant 0$; $p \neq 0$.

En reprenant la notation (1.5),

M est un ensemble-H ssi $\begin{pmatrix} \varepsilon_1 g_{11} & \cdots & \varepsilon_s g_{s1} \\ \vdots & & \vdots \\ \varepsilon_1 g_{1p} & \cdots & \varepsilon_s g_{sp} \end{pmatrix} \begin{pmatrix} p_1 \\ \vdots \\ p_s \end{pmatrix} = \theta_p$

$p_i \geqslant 0$ pour i = 1, ..., s; non tous nuls.

1.3.2. Formulation duale.

La formulation duale est donc :

M est un ensemble H ssi $\begin{pmatrix} g_{11} & \cdots & g_{s1} \\ \vdots & & \vdots \\ g_{1p} & & g_{sp} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_s \end{pmatrix} = \theta_p$ (1.6)

$y_i = \varepsilon_i p_i$, $p_i \geqslant 0$ pour i = 1, ..., s; non tous nuls.

On appelle configuration de signes associée à M le s-uplet ε donné par :

$$\varepsilon_i = y_i / |y_i| \qquad \text{si } y_i \neq 0,$$
$$\varepsilon_i \text{ arbitraire et tel que } |\varepsilon_i| = 1 \text{ sinon.}$$

Reprenant le problème d'approximation posé au paragraphe 1.1,la formulation duale permet de trouver une borne inférieure pour la norme de l'erreur différente de celle du théorème d'inclusion.

Théorème 1.4 :

On suppose [$M = \{P_i \in B$, pour $i = 1, \ldots, s\}$ et $\varepsilon = (\varepsilon_i : i = 1, \ldots, s)$] *est un ensemble-H.*

On note $p_o = ||f - w_o|| = \inf_{w \in W} ||f - w||$.

Si $\varepsilon_i [f(P_i) - w_o(P_i)] > 0$ *pour* $i = 1, \ldots, s$,

alors $$\frac{\sum_{i=1}^{s} \lambda_i f(P_i)}{\sum_{i=1}^{s} |\lambda_i|} \leqslant p_o$$

où $\lambda = (\lambda_1, \ldots, \lambda_s)$ *est solution du système définissant M comme ensemble-H.*

Démonstration :

Par hypothèse on a :

$$\sum_{i=1}^{s} \lambda_i g_k(P_i) = 0 \qquad \text{pour } k = 1, \ldots, s$$

Or tout élément w de W s'écrit comme :

$$w(.) = \sum_{k=1}^{p} \mu_k g_k(.),$$

donc $$\sum_{i=1}^{s} \lambda_i \sum_{k=1}^{p} \mu_k g_k(P_i) = \sum_{i=1}^{s} \lambda_i w(P_i) = 0 \qquad \text{pour tout } w \in W.$$

En particulier, pour w_o on obtient :

$$\sum_{i=1}^{s} \lambda_i w_o(P_i) = 0,$$

donc,

$$\frac{\sum_{i=1}^{s} \lambda_i f(P_i)}{\sum_{i=1}^{s} |\lambda_i|} = \frac{\sum_{i=1}^{s} \lambda_i (f(P_i) - w_o(P_i))}{\sum_{i=1}^{s} |\lambda_i|}$$

Or, par la définition de ε_i, on obtient :

$$\frac{\sum_{i=1}^{s} \lambda_i f(P_i)}{\sum_{i=1}^{s} |\lambda_i|} = \frac{\sum_{i=1}^{s} |\lambda_i| \varepsilon_i (f(P_i) - w_o(P_i))}{\sum_{i=1}^{s} |\lambda_i|}$$

Comme nous avons supposé $\varepsilon_i(f(P_i) - w_o(P_i)) > 0$, on peut écrire :

$$\frac{\sum_{i=1}^{s} \lambda_i f(P_i)}{\sum_{i=1}^{s} |\lambda_i|} = \frac{\sum_{i=1}^{s} |\lambda_i| \, |f(P_i) - w_o(P_i))|}{\sum_{i=1}^{s} |\lambda_i|}$$

$$\leqslant \max_{i=1,..,s} |f(P_i) - w_o(P_i)|$$

$$\leqslant \max_{P \in B} |f(P) - w_o(P)|.$$

Par la définition de la norme de TCHEBYCHEFF, on conclut :

$$\frac{\sum_{i=1}^{s} \lambda_i f(P_i)}{\sum_{i=1}^{s} |\lambda_i|} \leqslant P_o. \quad \blacksquare$$

Remarque : Il est aisé de vérifier que cette borne inférieure est meilleure que celle donnée par le théorème d'inclusion. En effet, dans la démonstration précédente, on obtenait :

$$\frac{\sum_{i=1}^{s} \lambda_i f(P_i)}{\sum_{i=1}^{s} |\lambda_i|} = \frac{\sum_{i=1}^{s} |\lambda_i| \varepsilon_i (f(P_i) - w_o(P_i))}{\sum_{i=1}^{s} |\lambda_i|}$$

En reprenant les notations du théorème d'inclusion, on peut écrire :

$$\frac{\sum_{i=1}^{s} \lambda_i \, f(P_i)}{\sum_{i=1}^{s} |\lambda_i|} = \frac{\sum_{i=1}^{s} |\lambda_i| \, \varepsilon_i \, \varepsilon(P_i)}{\sum_{i=1}^{s} |\lambda_i|}$$

Or les hypothèses de ce théorème sont :

Soit l'ensemble $M = M_1 \cup M_2$ et soit la fonction erreur $\varepsilon(.)$ vérifiant :

$$\varepsilon(P_i) > 0 \qquad \text{pour } P_i \in M_1$$

$$\varepsilon(P_i) < 0 \qquad \text{pour } P_i \in M_2.$$

Donc $\qquad \varepsilon_i \, \varepsilon(P_i) > 0 \qquad$ pour $i = 1, \ldots, s.$

On conclut

$$\frac{\sum_{i=1}^{s} \lambda_i \, f(P_i)}{\sum_{i=1}^{s} |\lambda_i|} = \frac{\sum_{i=1}^{s} |\lambda_i| \, |\varepsilon(P_i)|}{\sum_{i=1}^{s} |\lambda_i|}$$

$$\geqslant \inf_{i=1,..,s} |\varepsilon(P_i)|$$

Les intérêts de la formulation duale de la définition d'ensemble-H sont, d'une part, la découverte d'une borne inférieure de la norme de l'erreur plus précise, d'autre part, la possibilité de déterminer entièrement la configuration ε des signes associée à un ensemble-H par la résolution d'un système d'équations homogènes. Contrairement au système de COLLATZ, on peut toujours dans ce cas tirer une conclusion. La résolution de ce système est facilement programmable. Le programme établi le résout par diagonalisation.

En notant M l'ensemble des points introduits, les résultats sont :

ou bien 1° l'ensemble M n'est pas un ensemble-H,

ou bien 2° la configuration de signes associée à M
si l'ensemble M est un ensemble-H tel que card $n \leqslant p+1$
où p est le nombre de fonctions de base définissant W,

ou bien 3° une base de solutions
si l'ensemble M est un ensemble-H tel que card $n > p+1$
où p est le nombre de fonctions de base définissant W.

De plus amples explications sont données au chapitre traitant de la programmation.

Pour terminer ce paragraphe, on remarque [6] qu'au système (1.6) on peut associer une transformation linéaire U_M définie par :

$$U_M \in \mathrm{Hom}\,(\Lambda^s, \Lambda^p)$$

$$U_M(y_1, \ldots, y_s) = \left(\sum_{i=1}^{s} y_i\, g_1(P_i), \ldots, \sum_{i=1}^{s} y_i\, g_p(P_i)\right) \qquad (1.7)$$

Grâce à cet homomorphisme, la définition d'ensemble-H peut s'exprimer comme suit :

M est un ensemble-H ssi le noyau de l'homomorphisme U_M, noté $\mathrm{Ker}\ U_M$ est différent de θ_s,
ssi le rang de l'homomorphisme U_M, noté rg U_M, est tel que rg $U_M \leqslant$ card M-1.

De même, la définition d'ensemble-H minimal devient :

M est un ensemble-H ssi les conditions suivantes sont vérifiées :
(1) rg $U_M \leqslant$ card M-1,
(2) pour tout sous-ensemble propre M' de M, l'homomorphisme $U_{M'}$, associé à M' satisfait à : rg $U_{M'}$ = card M'.

On montre que cette définition est équivalente à la définition donnée en début de chapitre. Par ce qui précède, la condition (1) signifie que M est un ensemble-H. Il faut encore démontrer que la condition (2) est équivalente à dire : tout sous-ensemble propre M' de M n'est pas un ensemble-H.
Pour ce faire, on remarque que comme (card M') est le nombre de colonnes de la matrice représentant l'homomorphisme $U_{M'}$, on obtient :

$$\text{rg } U_{M'} \leqslant \text{card } M'.$$

En raisonnant par l'absurde, on suppose :

$$\text{rg } U_{M'} < \text{card } M',$$

c'est-à-dire,

$$\text{rg } U_{M'} \leqslant (\text{card } M') - 1.$$

Ce qui s'interprète par : M' est un ensemble-H. Or par hypothèse, $M' \subset M$, on obtient une contradiction avec la définition initiale d'ensemble-H minimal. Donc

$$\text{rg } U_{M'} \geqslant \text{card } M'.$$

On conclut : $$\text{rg } U_{M'} = \text{card } M'.$$

Il est également intéressant de remarquer la propriété suivante.

Propriété 1.1 :

Si M est un ensemble-H minimal, alors $\text{rg } U_M = \text{card } M-1$.

Démonstration :

1. Par la définition d'ensemble-H, on écrit :

$$\text{rg } U_M \leqslant \text{card } M-1.$$

2. Soit $M' = M \setminus \{P_i\}$ où P_i est un élément quelconque de M. Comme par hypothèse, M est minimal, on obtient : $\text{rg } U_{M'} = \text{card } M'$, c'est-à-dire

$$\text{rg } U_{M'} = \text{card } M-1.$$

Par construction, $M' \subset M$, donc $\text{rg } U_{M'} \leqslant \text{rg } U_M$. Ce qui entraîne

$$\text{card } M-1 \leqslant \text{rg } U_M.$$

Les conclusions des étapes 1 et 2 de la démonstration donnent immédiatement la thèse. ▪

1.4. Définitions apparentées.

Dans la littérature traitant du problème posé au premier paragraphe de ce chapitre, le terme de signature extrémale se rencontre aussi fréquemment que celui d'ensemble-H. Dans ce paragraphe, on donne les définitions de signature extrémale proposées par NEWMAN et SHAPIRO dans [12] et par BROSOWSKI dans [1].

1.4.1. Définition de NEWMAN et SHAPIRO.

On considère un ensemble discret X de points de B :

$$X = \{x_1, \ldots, x_n\}$$

et soit n vecteurs unités :

$$\varepsilon_1, \ldots, \varepsilon_n.$$

NEWMAN et SHAPIRO définissent :

1. Une signature Σ est l'ensemble des couples (x_i, ε_i).
2. Une signature Σ relative à W(1.1) est dite extrémale ssi il existe n nombres positifs $p_1, \ldots, p_n$ vérifiant :

$$\sum_{i=1}^{n} p_i \varepsilon_i g(x_i) = 0 \qquad \text{pour tout } g \in W.$$

On remarque immédiatement que la formulation duale (1.6) d'ensemble-H et la définition de signature extrémale donnée ci-dessus sont des définitions équivalentes.

1.4.2. Définition de BROSOWSKI.

Ces définitions peuvent encore s'exprimer par :

> Une signature Σ est dite extrémale relative à W ssi l'enveloppe convexe des vecteurs $\varepsilon_i(g_1(x_i), \ldots, g_p(x_i))$ où $(x_i, \varepsilon_i) \in \Sigma$ contient le vecteur nul.

Ceci est la formulation de signature extrémale proposée par BROSOWSKI.

1.4.3. Définition générale de signature extrémale.

Cette formulation est extraite de l'article de GEARHART [7].

1. Une signature σ est une fonction définie sur B, à support fini et dont les valeurs non nulles sont +1 ou -1.
2. Une signature σ est extrémale par rapport à W(1.1) ssi il existe une mesure positive non nulle μ telle que :

$$\int u(x)\, \sigma(x)\, d\mu(x) = 0 \qquad \text{pour tout } u \in W.$$

Il est de convention d'écrire des sommations finies pour les intégrales par rapport à des mesures discrètes. Dès lors, cette formulation générale dans le cas d'une mesure discrète s'identifie à la définition donnée par NEWMAN et SHAPIRO.

1.5. Propriétés des ensembles-H.

Les propriétés suivantes sont extraites de [6] :

1.5.1. Propriétés des ensembles-H minimaux.

Dans cette partie, on décrit deux propriétés spécifiques des ensembles-H minimaux. L'une concerne le cardinal d'un ensemble-H minimal, l'autre, les éléments du noyau Ker U_M où U_M est défini par (1.7)

Propriété 1.2 :

Si $M = \{P_1, \ldots, P_m\}$ *est un ensemble-H minimal, alors* card $M \leqslant p+1$.

Démonstration :

En raisonnant par l'absurde, on suppose que M est un ensemble-H minimal de cardinal strictement supérieure à p+1. Par la formulation duale (1.6), on a :
Le système suivant admet une solution non nulle

$$\begin{bmatrix} g_1(P_1) & \cdots & g_1(P_p) & g_1(P_{p+1}) & \cdots & g_1(P_m) \\ \vdots & & \vdots & \vdots & & \vdots \\ g_p(P_1) & \cdots & g_p(P_p) & g_p(P_{p+1}) & \cdots & g_p(P_m) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix} = \theta_p$$

où m = card M.
Ce qui veut dire : le déterminant de la matrice formée par les p premières colonnes est non nul. Donc, le système suivant :

$$\begin{bmatrix} g_1(P_1) & \cdots & g_1(P_p) & g_1(P_{p+1}) \\ \vdots & & & \\ g_p(P_p) & \cdots & g_p(P_p) & g_p(P_{p+1}) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \\ \lambda_{p+1} \end{bmatrix} = \theta_p$$

admet une solution non nulle.
Par (1.6), on conclut que l'ensemble $D = \{P_1, \ldots, P_{p+1}\}$ est un ensemble-H.
Comme $D \subset M$, on obtient une contradiction avec l'hypothèse : M est un ensemble-H minimal. Donc card $M \leqslant p+1$. ∎

Remarque :

Dans la théorie d'approximation, on définit un vecteur caractéristique relatif à W(1.1) comme étant un vecteur du type :

$$\phi(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_p(x) \end{pmatrix}$$

La condition de HAAR sur un ensemble M s'exprime alors par : pour tout $\{x_1, \ldots, x_p\} \subseteq M$ tel que $x_i \neq x_j$ pour $i \neq j$, les vecteurs caractéristiques correspondant à ces abscisses sont linéairement indépendants.
On peut donc écrire : si M est un ensemble-H tel que card M > p, alors les fonctions de base définissant W satisfont à la condition de HAAR sur M.

Propriété 1.3 :

Si $M = \{P_1, \ldots, P_m\}$ *est un ensemble-H minimal, alors* $\lambda_i \neq 0$ *pour tout* $i = 1, \ldots, m$ *où* $(\lambda_1, \ldots, \lambda_m) \in \operatorname{Ker} U_M$ *défini en (1.7),* $m = \operatorname{card} M$.

Démonstration :

En raisonnant par l'absurde, on suppose qu'il existe un indice $i \in \{1, \ldots, m\}$ tel que $\lambda_i = 0$. Le système :

$$\begin{pmatrix} g_1(P_1) & \cdots & g_1(P_m) \\ \vdots & & \vdots \\ g_{i-1}(P_1) & \cdots & g_{i-1}(P_m) \\ g_i(P_1) & \cdots & g_i(P_m) \\ g_{i+1}(P_1) & \cdots & g_{i+1}(P_m) \\ \vdots & & \vdots \\ g_p(P_1) & \cdots & g_p(P_m) \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{i-1} \\ 0 \\ \lambda_{i+1} \\ \vdots \\ \lambda_m \end{pmatrix} = \theta_p$$

s'écrit également :

$$\lambda_1 g_k(P_1) + \ldots + \lambda_{i-1} g_k(P_{i-1}) + 0 + \lambda_{i+1} g_k(P_{i+1}) + \ldots + \lambda_m g_k(P_m) = 0$$

pour $k = 1, \ldots, p$.

Il apparaît clairement que dans la résolution de ce système d'équations, on ne tient pas compte du point P_i. Donc $M \setminus \{P_i\}$ est encore un ensemble-H. Comme $M \setminus \{P_i\} \subset M$, on obtient une contradiction avec l'hypothèse. On conclut : pour tout $i \in \{1, \ldots, m\}$, $\lambda_i \neq 0$. ■

Remarque :

Lorsque l'on teste sur ordinateur un ensemble de points, on peut immédiatement voir si cet ensemble est un ensemble-H non minimal. En effet, si la solution du système homogène contient un élément nul, l'ensemble-H est non minimal.

1.5.2. Propriétés des ensembles-H relatifs à des espaces différents.

Propriété 1.4 :

Si $M = \{P_1, \ldots, P_m\}$ *est un ensemble-H relatif à* W *(1.1), alors* M *est un ensemble-H relatif à tous sous-espace linéaire* V *de* W.

Démonstration :

Par définition, M est un ensemble-H relatif à W ssi il existe $\lambda = (\lambda_1, \ldots, \lambda_m) \neq \theta_m$ tel que :

$$\sum_{i=1}^{m} \lambda_i \, g_k(P_i) = 0 \qquad \text{pour } k = 1, \ldots, p \qquad (1.8)$$

Comme $V \subseteq W$, on définit $V = \text{eng}\{h_1, \ldots, h_q\}$ où les fonctions h_i sont données par

$$h_i(.) = \sum_{i=1}^{p} a_{ij} \, g_j(.) \quad \text{pour } i = 1, \ldots, q$$

où $a_{ij} \in \mathbb{R}$ pour $i = 1, \ldots, q$; pour $j = 1, \ldots, p$. On obtient alors :

$$\sum_{i=1}^{m} \lambda_i \, h_k(P_i) = \sum_{i=1}^{m} \lambda_i \sum_{j=1}^{p} a_{kj} \, g_j(P_i) \quad \text{pour } k = 1, \ldots, q.$$

En permutant les signes de sommation, on écrit :

$$\sum_{i=1}^{m} \lambda_i \, h_k(P_i) = \sum_{j=1}^{p} a_{kj} \sum_{i=1}^{m} \lambda_i \, g_j(P_i) \quad \text{pour } k = 1, \ldots, q \qquad (1.9)$$

Par (1.8), (1.9) s'annule pour $k = 1, \ldots, q$. Par la formulation duale (1.6), on conclut que M est un ensemble-H relatif à V. ■

Propriété 1.5 :

Si $M = \{P_1, \ldots, P_m\}$ *est un ensemble-H relatif à W (1.1) tel que* card $M < p+1$, *alors M est un ensemble-H relatif à un sous-espace propre de W.*

Démonstration :

Par définition, M est un ensemble-H relatif à W ssi il existe $\lambda = (\lambda_1, \ldots, \lambda_m) \neq \theta_m$ tel que :

$$\begin{bmatrix} g_1(P_1) & \cdots & g_1(P_m) \\ \vdots & & \vdots \\ g_p(P_1) & \cdots & g_p(P_m) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix} = \theta_p$$

Ou bien, ce système est carré. Alors le déterminant de la matrice est nul, car le système admet une solution non nulle. Ce qui veut dire qu'une ligne est combinaison linéaire d'autres. Il y a donc une équation redondante.

Ou bien, ce système est rectangulaire - il y a plus d'équations que d'inconnues. Comme le système admet une solution non nulle, il contient des équations redondantes.

Si dans le système, on supprime les lignes correspondant aux équations redondantes, on engendre le sous-espace V de W en éliminant les fonctions de base correspondant à ces équations. Le nouveau système ainsi obtenu admet encore une solution non nulle. Donc, M est un ensemble-H relatif au sous-espace propre V de W. ■

Remarques :

1. De ce qui précède, on peut déduire le résultat suivant :
   si M est un ensemble-H relatif à W,
   si pour tout sous-espace propre V de W, M n'est pas un ensemble-H relatif à V, alors card $M = p+1$.
2. On ne peut déterminer une borne inférieure pour le cardinal d'un ensemble-H minimal. En effet, lorsque l'on considère les ensembles-H relatifs à W, il faut tenir compte des ensembles-H relatifs aux sous-espaces propres V de W.

Exemple : soit W = eng {1, x, y}. Les ensembles-H minimaux relatifs à W sont représentés en fig. (1.1) et (1.2), ce dernier (fig. 1.2) étant plus précisément un ensemble-H minimal relatif à V = eng {1, x}.

1.5.3. Propriétés concernant la configuration de signes d'un ensemble-H.

Propriété 1.6 :

Si M = $\{P_1, \ldots, P_m\}$ *est un ensemble-H relatif à W (1.1), si* $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)$ *est la configuration de signes associée à* M, *alors pour tout* $Q \subseteq B$ *tel que* card Q = q, M ∪ Q *est un ensemble-H relatif à W. De plus, la configuration de signes associée à* M ∪ Q *est donnée par :*

$$(\varepsilon_1, \ldots, \varepsilon_m, \alpha_1, \ldots, \alpha_q)$$

où $\alpha_i \in R$ *et* $|\alpha_i| = 1$, α_i *arbitraire.*

Démonstration :

Cette propriété devient évidente lorsque l'on se rappelle la définition d'ensemble-H donnée par COLLATZ. En effet M = $\{P_1, \ldots, P_m\}$ est un ensemble-H ssi le système suivant d'inéquations :

$$\begin{pmatrix} \varepsilon_1 g_1(P_1) & \cdots & \varepsilon_1 g_p(P_1) \\ \vdots & & \vdots \\ \varepsilon_m g_1(P_m) & \cdots & \varepsilon_m g_p(P_m) \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} > \theta_m$$

contient une contradiction. On observe que ajouter les points de Q à l'ensemble-H M, revient à ajouter q inéquations au système. Le système ainsi obtenu contient toujours une contradiction.
Donc M ∪ Q est un ensemble-H, et ce, quels que soient les signes attribués aux points de Q. Donc les α_i peuvent être pris arbitrairement dans les réels pourvu que $|\alpha_i| = 1$. ■

Propriété 1.7 :

Si M *est un ensemble-H non minimal de* (p+k) *points, alors on obtient* 2^k *configurations possibles de signes.*

Démonstration :

1. Si $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)$ est une configuration de signes associée à un ensemble-H M, alors ε' défini par :

$$\varepsilon' = (-\varepsilon_1, \ldots, -\varepsilon_m)$$

est aussi une configuration de signes associée à M.

2. On remarque qu'un ensemble-H minimal n'a que deux configurations possibles de signes. En effet, par la propriété 1.2, la configuration de signes d'un ensemble-H minimal est entièrement déterminée. Par l'étape 1 de la démonstration, on obtient immédiatement le résultat.

3. Par ce qui précède, si M_1 est tel que card $M_1 \leq p+1$, on a 2 configurations de signes. On montre que si M_2 est tel que card $M_2 = p+2$, on a 2^2 configurations de signes. On suppose que $\varepsilon^1 = (\varepsilon_1, \ldots, \varepsilon_{p+1})$ est une configuration associée à M_1 et $M_2 = M_1 \cup \{q\}$. Par la propriété 1.5, $\varepsilon^2 = (\varepsilon_1, \ldots, \varepsilon_{p+1}, \alpha)$ où $|\alpha| = 1$ sont deux configurations associées à M_2.

   De même, en considérant la deuxième configuration associée à M_1, on obtient deux autres configurations associées à M_2. Au total, on compte $2^2 = 4$ configurations associées à M_2.

   Par récurrence, on suppose que la propriété soit vraie pour M' tel que card M' = p+k-1. Soit l'ensemble-H M tel que card M = p+k. A chaque configuration de M' sont associées deux configurations pour M. On conclut qu'il existe 2^k configurations possibles de signes associées à M. ∎

Chapitre 2 :
Invariance et méthodes de construction.

2.1. Invariance d'ensembles-H.

Cette théorie a été étudiée par C.DIERIECK dans [6].

2.1.1. Définition.

Soit $D = \{P_1, P_2, \ldots, P_m\} \subset B$, un ensemble-H relatif au sous-espace W engendré par $\{g_1, \ldots, g_n\}$. En d'autres mots, il existe au moins un ensemble de m scalaires λ_i $(1 \leqslant i \leqslant m)$ non tous nuls tels que :

$$\sum_{i=1}^{m} \lambda_i \, g_j(P_i) = 0 \qquad 1 \leqslant j \leqslant n \qquad (2.1)$$

la configuration des signes ε_i attachés aux points P_i étant donnée par $\varepsilon_i = \lambda_i / |\lambda_i|$ pour $\lambda_i \neq 0$ et $|\varepsilon_I| = 1$ pour $\lambda_I = 0$.

Considérons une transformation biunivoque T de B dans B. On dit que l'ensemble-H : D *est laissé invariant par la transformation* T si $T(D) = \{T(P_1), \ldots, T(P_m)\}$ est également un ensemble-H relatif à W ayant au moins la configuration de signes $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m)$.

2.1.2. Critère d'invariance.

Théorème 2.1 :

Les ensembles-H relatifs au sous-espace W engendré par $\{g_1, \ldots, g_n\}$ *sont laissés invariants par une transformation biunivoque* T *de* B *dans* B *si pour tout* i = 1, ..., n, *on a*

$g_i \circ T$ *est un élément de* W.

Démonstration :

Puisque $g_j \circ T \in W$, on peut l'exprimer sous forme de combinaison linéaire des éléments de la base $(g_1, \ldots, g_n)$

$$g_j \circ T = \sum_{k=1}^{n} a_{kj} \, g_k \qquad (1 \leqslant j \leqslant n)$$

de sorte que

$$\sum_{i=1}^{m} \lambda_i \, g_j(T(P_i)) = \sum_{i=1}^{m} \lambda_i \sum_{k=1}^{n} a_{kj} \, g_k(P_i)$$

$$= \sum_{k=1}^{n} a_{kj} \sum_{i=1}^{m} \lambda_i \, g_k(P_i)$$

est nul par (2.1) :

$T(D) = \{T(P_1), \ldots, T(P_m)\}$ est donc un ensemble-H relatif à W avec la configuration de signes

$$\{\lambda_i \; / \; |\lambda_i| ; \; 1 \leqslant i \leqslant m\} \text{ de } D. \; \blacksquare$$

On note par S, l'ensemble des transformations biunivoques T de B dans B qui satisfont à ce critère. Comme la composition est une loi interne sur S et associative, S muni de cette loi, est un semi-groupe et la transformation T qui à tout $q \in B$ associe son image $T(q) = q$, appelée identité, est l'élément neutre de S.

Dans la plupart des cas pratiques $S, \circ$ est un groupe car il contient l'élément symétrique de tout $T \in S$. Dans ces cas, on a la propriété :

Propriété 2.1 :

Soit un groupe $G, \circ$ *de transformations biunivoques de* B *dans* B.
Si pour tout $T \in G$, *il existe un élément unique* τ, *endomorphisme de* W *et tel que* $v \circ T = \tau(v)$ *pour tout* $v \in W$.
Alors τ *est un automorphisme.*

Démonstration :

Pour tout élément T du groupe $G, \circ$, il existe un symétrique nécessairement unique T^{-1} et donc, par hypothèse, il existe un endomorphisme τ de W tel que $v \circ T = \tau(v)$ pour tout $v \in W$ et un endomorphisme σ de W tel que $v \circ T^{-1} = \sigma(v)$ pour tout $v \in W$. Par conséquent, à l'identité $I_G = T \circ T^{-1} = T^{-1} \circ T$ correspond l'endomorphisme de W : $1_W = \sigma \circ \tau = \tau \circ \sigma$.
Il suit que $\sigma = \tau^{-1}$ et donc τ est un élément du groupe linéaire GL(W). ■

Découlant du critère, on a qu'un ensemble-H relatif à W est laissé invariant par les transformations du groupe $G, \circ$ décrit dans la propriété.

2.1.3. Groupes particuliers de transformations.

Les transformations biunivoques de $\mathbf{R}^s$ dans $\mathbf{R}^s$ appartenant aux groupes suivants laissent les ensembles-H, relatifs à W, invariants moyennant certaines conditions sur W.

1. Le groupe des translations et des dilatations non nulles : H(s, R).

$T \in H(s, \mathbf{R})$ est une application qui à x fait correspondre $T(x) = ax + b$ où $a \in \mathbf{R} \setminus \{0\}$ et $b \in \mathbf{R}^s$.

On prend comme exemple (fig. 2.1), la transformation de $\mathbf{R}^2$ dans $\mathbf{R}^2$ qui associe à x son image $T(x) = \frac{2}{3} x + b$ où x et b sont des points de $\mathbf{R}^2$, de coordonnées respectivement (x_1, x_2) et $(-2, -\frac{4}{3})$.

Les translations permettent de ramener un ensemble-H quelconque sur un ensemble-H comprenant l'origine et les dilatations permettent de réduire ou d'augmenter dans une même proportion les distances entre tous les points.

2. Le groupe des transformations biunivoques linéaires : GL(s, R).

$T \in GL(s, \mathbf{R})$ est une application qui donne pour image à $x \in \mathbf{R}^s$: $T(x) = Ax$ avec A, une matrice s x s de déterminant non nul.

La transformation T qui envoie x sur Ax, $x \in \mathbf{R}^2$ et A désignant la matrice $\begin{pmatrix} \frac{3}{5}, & -\frac{2}{5} \\ \frac{1}{5} & \frac{1}{5} \end{pmatrix}$ est une transformation du groupe GL(2, R) (fig. 2.2).

Les transformations de GL(s, R) permettent de ramener la plupart des ensembles-H comprenant l'origine sur un ensemble-H comprenant l'origine et les s vecteurs de base de $\mathbf{R}^s$.

Remarque 2.1 : Cette propriété n'est pas valable pour tous les ensembles-H. Il suffit pour s'en convaincre de prendre dans $\mathbf{R}^s$, un ensemble contenu dans un hyperplan de $\mathbf{R}^s$. La propriété d'appartenance à un hyperplan est une propriété invariante par les transformations décrites dans cette section.

Exemple : Trois points sur une droite de $\mathbf{R}^2$, ne peuvent être envoyés sur les points (0,0), (0,1) et (1,0).

3. Le groupe des transformations biunivoques affines : GA(s, R).

$T \in GA(s, R)$ est une application de $\mathbf{R}^s$ dans $\mathbf{R}^s$ telle que $T(x) = Ax + b$, où x est un point de $\mathbf{R}^s$, A : une matrice s x s de déterminant non nul et b : un vecteur de $\mathbf{R}^s$. On remarque qu'une transformation affine peut être obtenue par composition d'une transformation linéaire et d'une translation.

On reprend pour A la matrice $\begin{pmatrix} \frac{3}{5} & -\frac{2}{5} \\ \frac{1}{5} & \frac{1}{5} \end{pmatrix}$ et on prend pour b, le vecteur de coordonnée (-1, -1). La fig. 2.3 illustre cet exemple.

On peut donc par ces transformations ramener la plupart des ensembles-H sur un ensemble-H comprenant l'origine et les s vecteurs de base de $\mathbf{R}^s$. La remarque 2.1 faite précédemment s'applique aussi à ce cas.

4. Le groupe de toutes les collinéations : PGL(s+1, R).

Une collinéation est un isomorphisme de l'espace vectoriel de dimension s+1 : $\mathbf{R}^{s+1}$ dans $\mathbf{R}^{s+1}$ qui envoie un sous-espace de dimension 1 sur un sous-espace de dimension 1 de $\mathbf{R}^{s+1}$ et donc, une collinéation détermine une transformation d'un espace projectif P de dimension s sur lui-même.

Afin d'éclairer cette notion, on rappelle quelques concepts :
Sur l'espace vectoriel de dimension n+1, V^{n+1}, on définit, la relation F : "deux vecteurs non nuls sont en relation F si et seulement si ils sont linéairement dépendants".
C'est une relation d'équivalence qui engendre les classes d'équivalence constituées des sous-espaces de dimension 1 de $V^{n+1} \setminus \{0\}$.

L'espace projectif de dimension n, P_n, est l'ensemble de ces classes d'équivalence, par conséquent, un point de P_n est une droite de V^{n+1}.

Coordonnées homogènes.

Soit P, un point de P_n, de coordonnées $x_1, x_2, \ldots, x_n$. On considère tous les n+1-uplets de nombres réels $(\xi_1, \ldots, \xi_n, \xi_o)$ pour lesquels $\xi_o \neq 0$: ils déterminent un point de V_{n+1}. On met en correspondance ces ensembles de n+1 nombres réels et les points de P_n en utilisant la règle suivante : $P = (x_1, x_2, \ldots, x_n)$ et un n+1-uplet $(\xi_1, \ldots, \xi_n, \xi_o)$ avec $\xi_o \neq 0$ sont mis en correspondance si et seulement si

$$x_1 = \frac{\xi_1}{\xi_o}, \; x_2 = \frac{\xi_2}{\xi_o}, \; \ldots, \; x_n = \frac{\xi_n}{\xi_o}$$

On remarque de suite qu'alors $(\lambda \xi_1, \ldots, \lambda \xi_n, \lambda \xi_o)$ avec $\lambda \neq 0$ correspond aussi à $(x_1, x_2, \ldots, x_n)$.
En particulier, tous les (n+1)-uplets associés au point $P = (x_1, x_2, \ldots, x_n)$ sont de la forme $(\lambda x_1, \lambda x_2, \ldots, \lambda x_n, \lambda)$ avec $\lambda \neq 0$.
On définit une correspondance biunivoque en imposant $\xi_o = 1$. On appelle alors $x_1, x_2, \ldots, x_n, 1$ les *coordonnées homogènes* du point $P = (x_1, x_2, \ldots, x_n)$.

On suppose maintenant que $P = (x_1, x_2, \ldots, x_n)$ est un point fixé de P_n, distinct de l'origine (fig. 2.4).
Si on pose $Q = (\lambda x_1, \ldots, \lambda x_n)$ et si on fait varier λ depuis +1 jusqu'à $+\infty$, alors Q court le long de la droite g déterminée par les points O et P (fig. 2.4) depuis P vers l'infini.

On peut prendre $\xi_o = \frac{1}{\lambda}$, $\xi_1 = x_1, \ldots, \xi_n = x_n$ comme coordonnées homogènes généralisées pour Q. Alors comme $\lambda \to \infty$, on a $\xi_o \to 0$, $\xi_i \to x_i$, $i = 1, \ldots, n$ et Q devient point à l'infini. On considère donc $x_1, x_2, \ldots, x_n, 0$ comme les coordonnées homogènes d'un point à l'infini.

Remarques: - On définit aussi un point à l'infini comme étant le "point d'intersection" de deux droites parallèles.
- On définit la droite à l'infini comme étant la totalité des points à l'infini. Son équation en coordonnées homogènes est $\xi_o = 0$.

Une collinéation de PGL(s+1, R) est donc une transformation T qui à un point x de P_s fait correspondre un point T(x) de P_s.

Cette transformation est définie comme suit : à $x = (x_1, x_2, \ldots, x_s)$, on fait correspondre ses coordonnées homogènes $(\lambda x_1, \ldots, \lambda x_s, \lambda)$ qui déterminent un point de $\mathbf{R}^{s+1}$: ξ auquel on fait correspondre, par la transformation τ de $GL(s+1, \mathbf{R})$, le point $\tau(\xi) = A\xi$, de coordonnées $\tau(\xi) = (\xi_1, \ldots, \xi_s, \xi_o)$. Ce s+1-uplet détermine un point de P_s écrit en coordonnées homogènes. On pose

$$T(x) = \left(\frac{\xi_1}{\xi_o}, \ldots, \frac{\xi_s}{\xi_o}\right) \in P_s$$

La transformation T définie par la matrice de τ

$$A = \begin{pmatrix} \frac{2}{5} & -\frac{3}{5} & 0 \\ \frac{1}{5} & \frac{1}{5} & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

ramène l'ensemble-H : D' sur l'ensemble-H comprenant l'origine (fig. 2.5).

Les collinéations permettent de ramener un point en l'origine et s vecteurs linéairement indépendants sur les s vecteurs de base de $\mathbf{R}^s$, tout comme les transformations affines de $\mathbf{R}^s$ mais en plus, elles ont la propriété fort importante de pouvoir ramener un point à l'infini en un point fini.

On reprend l'exemple précédent (fig. 2.5) avec comme quatrième point pour D', non plus celui de coordonnée (6,4,0) mais un point à l'infini sur la droite passant par l'origine et par (3,2,1):ce point à l'infini a donc comme coordonnées homogènes, le triplet (3, 2, 0).
La transformation de PGL(3, R) définie par la matrice

$$A = \begin{pmatrix} \frac{2}{5} & -\frac{3}{5} & 0 \\ \frac{1}{10} & \frac{1}{10} & -\frac{5}{10} \\ \frac{1}{10} & \frac{1}{10} & \frac{5}{10} \end{pmatrix}$$

envoie D' sur D (fig. 2.6).

Il est à remarquer que si ces transformations sont prises en considération pour l'invariance, on est conduit à chercher des propriétés affines et projectives comme principes de construction des ensembles-H.

En termes plus géométriques, on va travailler avec des propriétés de co - linéarité de points et de droites concourantes.

2.1.4. Conditions d'invariance imposées à W pour ces groupes.

Dans cette recherche, on s'intéresse aux sous-espaces W engendrés par les fonctions de base $g_1 = x^i y^j \ldots z^k$ avec $(i, j, \ldots, k) \in K$ où K est un ensemble de s-uplets (si on travaille dans $\mathbf{R}^s$) composés d'entiers positifs ou nuls. On notera ces sous-espaces :

$$V_s = \text{eng } \{ x^i y^j \ldots z^k \mid (i, j, \ldots, k) \in K \subset \mathbf{N}^s \} \qquad (2.2)$$

Il serait intéressant de déterminer les conditions que doit satisfaire cet ensemble K pour que les ensembles-H relatifs à V_s restent invariants par les transformations des quatre groupes.

Quand K sera déterminé à partir du degré de l'approximation : n

$$K = \{(i, j, \ldots, k) \mid i, j, \ldots, k \geqslant 0 \text{ et } i+j+ \ldots + k \leqslant n\} \qquad (2.3)$$

on notera le sous-espace vectoriel par V_s^n. On démontrera les propositions donnant ces conditions dans $\mathbf{R}^2$, les démonstrations étant faciles à généraliser dans $\mathbf{R}^s$.

1. Conditions sur V_s quand le groupe de transformations est H(s, R).

Théorème 2.2 :

Les ensembles-H relatifs à V_s (2.2) *restent invariants sous les transformations de* H(s, R) *si pour tout* (a, b, ..., c) ∈ K, *les s-uplets du produit cartésien* (0, 1, ..., a) x (0, 1, ..., b) x ... x (0, 1, ..., c) *sont des éléments de* K.

Démonstration :

Soient donc V_2 et T une transformation de H(2, R) qui à un point quelconque P de coordonnées (ξ, η) fait correspondre le point T(P) de coordonnées (ξ', η') avec

$$\xi' = \lambda \xi + t_1 \quad \text{et} \quad \eta' = \lambda \eta + t_2 \qquad \lambda \neq 0$$

Pour tout (a, b) ∈ K, on a

$$[x^a y^b] \circ T(P) = \xi'^a \xi'^b$$

$$= (\lambda \xi + t_1)^a (\lambda \eta + t_2)^b$$

En appliquant la formule du binôme de Newton et en réarrangeant les termes, on peut écrire en posant :

$$t_{ij} = \binom{a}{i}\binom{b}{j} \lambda^{i+j} t_1^{a-i} t_2^{b-j}$$

$$[x^a y^b] \circ T(P) = \sum_{i=0}^{a} \sum_{j=0}^{b} t_{ij} [x^i y^j] (P)$$

Le critère d'invariance est dès lors satisfait si $(i, j) \in K$, $i = 0, 1, \ldots, a$ et $j = 0, 1, \ldots, b$. ∎

Exemple : l'ensemble K_1 dont les points sont représentés à la fig. (2.7) respecte la condition imposée.

2. Condition sur V_s quand le groupe de transformations est $GL(s, \mathbb{R})$.

Théorème 2.3 :

Les ensembles-H relatifs à V_s (2.2) *restent invariants sous des transformations linéaires de* $GL(s, \mathbb{R})$ *si pour tout* $(a, b, \ldots, c) \in K$, *l'ensemble des s-uplets*

$$\{(i, j, \ldots, k) \mid i + j + \ldots + k = a + b + \ldots + c;$$
$$i, j, \ldots, k \in \{0, 1, \ldots, a + b + \ldots + c\}\}$$

est un sous-ensemble de K.

Démonstration :

Soit $(a, b) \in K$ et une transformation $T \in GL(2, \mathbb{R})$

$$[x^a y^b] \circ T(P) = (\alpha_{11} \xi + \alpha_{12} \eta)^a (\alpha_{21} \xi + \alpha_{22} \eta)^b$$

Par la formule du binôme de Newton et après réarrangement des termes, on a

$$[x^a y^b] \circ T(P) = \sum_{i=0}^{a} \sum_{j=0}^{b} t_{ij} \xi^{i+j} \eta^{a+b-i-j}$$

$$= \sum_{i=0}^{a} \sum_{j=0}^{b} t_{ij} [x^{i+j} y^{a+b-(i+j)}] (P)$$

Par conséquent, si les fonctions $x^c \cdot y^d$ vérifiant $c + d = a + b$ sont des fonctions de base de V_2 avec $c, d \in \{0, 1, \ldots, a+b\}$, alors le critère d'invariance est satisfait. ∎

Exemple : l'ensemble K_2 comprenant les points représentés à la fig. 2.7 satisfait à cette condition.

3. Conditions sur V_s quand le groupe des transformations est GA(s, R).

Théorème 2.4 :

Les ensembles-H relatifs à V_s (2.2) *sont laissés invariants par les transformations affines de* GA(s, R) *si pour tout* (a, b, ..., c) $\in K$, *l'ensemble*

$$\{(i, j, \ldots, k) \mid i + j + \ldots + k \leqslant a + b + \ldots + c;$$
$$i, j, \ldots, k \in \{0, 1, \ldots, a + b + \ldots + c\}\}$$

est une partie de K.

Il suffit de combiner le théorème 2.2 et le théorème 2.3.

Exemple : l'ensemble K_3 de la fig. 2.7 vérifie cette condition.

4. Conditions sur V_s quand le groupe de transformations est PGL(s+1, R).

- Homogénéisation d'un polynome à deux variables :

Soit $p(x, y) \in V_2^k$: $p(x, y) = \sum_{i=0}^{k} \sum_{j=0}^{k-i} a_{ij}\, x^i\, y^j$

Ce polynome s'écrit en coordonnées homogènes :

$$P(X, Y, Z) = Z^k \sum_{i=0}^{k} \sum_{j=0}^{k-i} a_{ij} \left(\frac{X}{Z}\right)^i \left(\frac{Y}{Z}\right)^j$$

$$= \sum_{i=0}^{k} \sum_{j=0}^{k-i} a_{ij}\, X^i\, Y^j\, Z^{k-(i+j)}$$

$$= \sum_{I+J+K=k} A_{IJK}\, X^I\, Y^J\, Z^K$$

avec I = 0, ..., k et J = 0, ..., k-I

L'espace des polynomes homogénéisés de degré k est donc engendré par $\{X^I\, Y^J\, Z^K \mid I + J + K = k,\ I = 0, \ldots, k\ \ J = 0, \ldots, k\ \ K = 0, \ldots, k\}$ ou de manière équivalente, par

$$\{Z^k \left(\frac{X}{Z}\right)^I \left(\frac{Y}{Z}\right)^J \mid I = 0, \ldots, k \text{ et } J = 0, \ldots, k-I\}$$

On note cet espace U_2^k, ou encore, $Z^k \cdot V_2^k$.
L'homogénéisation des polynomes multivariables se fait de la même façon.

Lemme 2.1 :

Les ensembles-H relatifs à U_s^k sont invariants sous des transformations de GL(s+1, R).

Il suffit d'appliquer le théorème 2.3.

Lemme 2.2 :

Soit $D = \{P_1, \ldots, P_m\}$ *un ensemble-H relatif à* W *(1.1) dont la configuration de signes est* $(\varepsilon_1, \ldots, \varepsilon_m)$.

D *avec la configuration de signes* $\left(\varepsilon_i \cdot \frac{g_o(P_i)}{|g_o(P_i)|} \mid i = 1, \ldots, m\right)$ *est un ensemble-H relatif à*

$$W_o = \text{eng}\{g_o \cdot g_1, g_o \cdot g_2, \ldots, g_o \cdot g_n\}$$
$$= g_o \cdot W$$

si $g_o(P_i) \neq 0 \quad i = 1, \ldots, m.$

La démonstration de ce lemme est donnée au début du chapitre 4.

Théorème 2.5 :

Les ensembles-H relatifs à V_s^{2n} (2.3) sont laissés invariants sous des collinéations de PGL(s+1, R).

Démonstration :

1. L'homogénéisation $h : \mathbf{R}^2 \to \mathbf{R}^3$ qui permet de passer de V_2^k à U_2^k laisse les ensembles-H invariants :

$$h(D) = \{h(P_i), i = 1, \ldots, m\}$$
$$= \{(x_i, y_i, 1), i = 1, \ldots, m, P_i = (x_i, y_i)\}$$

Il suffit d'appliquer le lemme 2.2. On a

$$\varepsilon_i \cdot \frac{g_o(P_i)}{|g_o(P_i)|} = \varepsilon_i \frac{Z^k(P_i)}{|Z^k(P_i)|} = \varepsilon_i.$$

2. Le lemme 2.1 affirme l'invariance des ensembles-H relatifs à U_2^k sous une transformation $\tau \in GL(3, \mathbb{R})$

$$\tau(P_i) = A \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = \begin{pmatrix} \xi_i \\ \eta_i \\ \alpha_i \end{pmatrix} \qquad i = 1, \ldots, m$$

où A est la matrice définissant τ.

3. Le lemme 2.2 permet de passer de l'ensemble-H : D'

$$D' = \tau(h(D)) = \{P'_i = (\xi_i, \eta_i, \alpha_i) \mid i = 1, \ldots, m\}$$

relatif à U_2^k à

$$D'' = \{P''_i = (\frac{\xi_i}{\alpha_i}, \frac{\eta_i}{\alpha_i}) \mid i = 1, \ldots, m'\}$$

relatif à V_2^k moyennant un changement de configuration gouverné par la règle

$$\varepsilon''_i = \varepsilon_i \cdot \frac{Z^k(P'_i)}{|Z^k(P'_i)|} = \varepsilon_i \frac{\alpha_i^k}{|\alpha_i^k|}$$

On conclut que la configuration ne change pas si k est pair et que certaines configurations sont modifiées si k est impair. ▪

Exemple :

Soit $D = \{P_1 = (0,0), P_2 = (1,0), P_3 = (0,1), P_4 = (1,1)\}$ un ensemble-H relatif à V_2^1 et $(\varepsilon_1 = +1, \varepsilon_2 = -1, \varepsilon_3 = -1, \varepsilon_4 = +1)$ sa configuration de signes associée (fig.2.8).
On définit une collinéation : T de PGL(3, $\mathbb{R}$) par sa matrice

$$\begin{pmatrix} \alpha & 0 & 0 \\ \beta & \alpha-\beta & 0 \\ \beta & \alpha-1 & 1-\beta \end{pmatrix}$$

qui envoie D sur D"

$$D'' = T(D) = \{(0,0), (\alpha, \beta), (0,1), (1,1)\}$$

On étudie les transformations de D suivant quelques valeurs de α et β .

1. $0 < \alpha < \beta < 1$.

   D" pour ces valeurs de α et β est représenté à la fig. 2.9a. On a un changement de signe dans la configuration. En effet,

$$\tau(h(P_1)) = (0,0,1-\beta) \quad \text{et donc} \quad \varepsilon''_1 = +1 \cdot \frac{1-\beta}{|1-\beta|} = +1$$

$$\tau(h(P_2)) = (\alpha,\beta,1) \quad \text{et donc} \quad \varepsilon''_2 = -1 \cdot \frac{1}{|1|} = -1$$

$$\tau(h(P_3)) = (0,\alpha-\beta,\alpha-\beta) \quad \text{et donc} \quad \varepsilon''_3 = -1 \cdot \frac{\alpha-\beta}{|\alpha-\beta|} = +1$$

$$\tau(h(P_4)) = (\alpha,\alpha,\alpha) \quad \text{et donc} \quad \varepsilon''_4 = +1 \cdot \frac{\alpha}{|\alpha|} = +1$$

2. $0 < \beta < 1$ et $\beta < \alpha$

D" pour $\alpha = 2$ et $\beta = \frac{1}{2}$ est représenté à la fig. 2.9b. La configuration de signes reste (+1, -1, -1, +1) car $\alpha - \beta > 0$.

3. $0 < \alpha < \beta$ et $\beta > 1$

D" pour $(\alpha, \beta) = (1, 2)$ est représenté à la fig. 2.9c.
Deux signes sont changés : $\varepsilon_1'' = -1$ et $\varepsilon_3'' = +1$.

2.2. Méthodes de construction d'ensembles-H.

Dans cette section, on va étudier différentes méthodes de construction d'ensembles-H.

2.2.1. Principes de décomposition des ensembles-H.

On peut trouver ces principes dans [4] de COLLATZ et [6] de C. DIERIECK. L'analyse d'une structure particulière de la matrice, associée à la transformation linéaire U_D (1.7) $\in$ Hom (R^m, R^n) et notée $M(U_D)$ mène à ces principes.

1. Notion d'ensemble-B.

Soit un ensemble $D = \{P_1, \ldots, P_m\}$; on suppose que la matrice $M(U_D)$ a la structure particulière suivante :

$$M(U_D) \quad \begin{pmatrix} M(U_{11}) & \vdots & M(U_{12}) \\ \cdots\cdots & \vdots & \cdots\cdots \\ 0 & \vdots & M(U_{22}) \end{pmatrix} \qquad (2.4)$$

avec $U_{11} \in$ Hom (R^{m_1}, R^{n_1}), $U_{12} \in$ Hom (R^{m_2}, R^{n_1}) et $U_{22} \in$ Hom (R^{m_2}, R^{n_2}) et telles que $m_1 + m_2 = m$ et $n_1 + n_2 = n$.

Remarque : Les résultats auxquels on aboutira pour cette structure particulière de $M(U_D)$ pourront être facilement étendus aux structures $(\alpha \times \alpha)$ diagonales.

On suppose que D est un ensemble-H relatif à W (1.1). En d'autres mots, il existe un vecteur non nul $\lambda \in R^m$ appartenant au noyau de U_D : $U_D(\lambda) = \theta$.
Dès lors, si on considère un vecteur $\lambda_1 \in R^{m_1}$, ayant comme coordonnées les m_1 pre-

mières composantes de λ, et un vecteur $\lambda_2 \in R^{m_2}$, composé des m_2 dernières coordonnées de λ, on peut écrire :

$$U_{11}(\lambda_1) = - U_{12}(\lambda_2) \tag{2.5}$$

et

$$U_{22}(\lambda_2) = \theta \tag{2.6}$$

Puisque par hypothèse, D est un ensemble-H, on a rg $U_D \leqslant m-1$, il suit donc, en excluant le cas où $\lambda_2 = \theta$, que

$$\text{rg } U_{22} \leqslant m_2 - 1 \tag{2.7}$$

Par conséquent, l'ensemble $D_2 = \{P_{m_1+1}, \ldots, P_m\}$ est un ensemble-H relatif au sous-espace linéaire W_2 engendré par $g_{n_1+1}, \ldots, g_n$.

Il est important de remarquer que la configuration de signes $(\varepsilon_i \mid i = 1, \ldots, m_1)$ de $D_1 = \{P_1, \ldots, P_{m_1}\}$ dépend fortement de la configuration de signes de D_2 puisque (2.5). Il s'ensuit que l'on ne peut déterminer d'une manière unique les signes des points de D_1, c'est pourquoi on appellera D_1, *un ensemble-B* relatif à $\{g_1, \ldots, g_{n_1}\}$.
Mais dans beaucoup de cas pratiques, il sera possible de déterminer la configuration de signes de D_1 en utilisant des méthodes particulières et le principe d'invariance.

2. Notion de bloc.

Une forme particulière de $M(U_{12})$ conduira à la méthode de composition de blocs.

$$M(U_{12}) = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ & & & \\ & \bigcirc & & \end{pmatrix} \in R^{n_1 \times m_2} \tag{2.8}$$

Elle est particulièrement intéressante quand une fonction de base est la fonction constante $g \equiv 1$, ce qui se passe dans les problèmes d'approximation polynomiale.

Considérons maintenant l'ensemble $D = D_1 \cup D_2$ supposé être un ensemble-H relatif à W engendré par l'ensemble

$$\{1, \psi_1, \psi_2, \ldots, \psi_{n_1}, \chi_1, \chi_2, \ldots, \chi_{n_2}\}$$

et tel que card $D_1 = m_1$ et card $D_2 = m_2$. La matrice $M(U_D)$ a dès lors la structure suivante :

$$M(U_D) = \left(\begin{array}{c|c} 1 \ldots 1 & 1 \ldots 1 \\ \hline M(U_1) & 0 \\ \hline 0 & M(U_2) \end{array}\right) \qquad (2.9)$$

avec $U_1 \in \text{Hom}\,(\mathbf{R}^{m_1}, \mathbf{R}^{n_1})$ et $U_2 \in \text{Hom}\,(\mathbf{R}^{m_2}, \mathbf{R}^{n_2})$

Par la définition d'ensemble-H, il existe un vecteur $\lambda \in \mathbf{R}^{m_1+m_2}$ qui vérifie $U_D(\lambda) = \theta$ ou encore, si on pose $\lambda = (\lambda_1, \lambda_2)$, $U_1(\lambda_1) = \theta$ et $U_2(\lambda_2) = \theta$, $\lambda_1 \in \mathbf{R}^{m_1}$, $\lambda_2 \in \mathbf{R}^{m_2}$, λ_1 et λ_2 vecteurs non nuls.
Par conséquent D_1 est un ensemble-H relatif à $W_1 = \text{eng}\,\{\psi_1, \ldots, \psi_{n_1}\}$ et D_2 un ensemble-H relatif à $W_2 = \text{eng}\,\{\chi_1, \ldots, \chi_{n_2}\}$.

De plus, la configuration de signes de D_1 est globalement influencée par celle de D_2 puisque l'on a la relation

$$\sum_{i=1}^{m_1} \lambda_{1i} = - \sum_{j=1}^{m_2} \lambda_{2j} \qquad (2.10)$$

qui découle de la multiplication de la première ligne de $M(U_D)$ par le vecteur (λ_1, λ_2).

On appellera D_1 et D_2 des blocs.
Le bloc D_1 sera dit "minimal" si D_1 est un ensemble-H minimal relatif à W_1. Le bloc D_1 avec la configuration $(\lambda_{1i} \,/\, |\lambda_{1i}| \mid i = 1, \ldots, n_1)$ où λ_1 appartient au noyau de U_1 est dit "bloc positif" si et seulement si

$$\sum_{i=1}^{n_1} \lambda_{1i} > 0$$

En conséquence, l'ensemble-H : $D = D_1 \cup D_2$ est composé d'un bloc D_1 positif et d'un bloc négatif : D_2.

Remarques : 1. La décomposition d'un ensemble-H minimal mène à deux blocs minimaux.

On démontre la contraposée.

On suppose que le bloc $D_1 = \{P_1, \ldots, P_{m_1}\}$ relatif à W_1 résultant de la décomposition de l'ensemble-H minimal $D = D_1 \cup D_2$ relatif à W n'est pas minimal.

Autrement dit $rg\ U_{D_1} < m_1 - 1$. Par conséquent, $rg\ (U_D) < m_1 + m_2 - 1$.

L'ensemble-H D n'est donc pas minimal. ▪

2. La composition multiple des blocs ne mène jamais à des ensembles-H minimaux.

Les signes des blocs peuvent être choisis librement à condition toutefois qu'il y ait au moins un bloc négatif et un bloc positif.

Démonstration :

On compose les deux blocs positifs minimaux C et D et un bloc négatif minimal B. Comme $rg\ U_C = card\ C-1$ et $rg\ U_D = card\ D-1$

$$rg\ U_{C\cup D} = card\ C + card\ D-2$$

où $U_{C\cup D}$ est définie par la matrice

$$M(U_{C\cup D}) = \begin{pmatrix} M(U_C) & \vdots & 0 \\ \cdots\cdots & \vdots & \cdots\cdots \\ 0 & \vdots & M(U_D) \end{pmatrix}$$

Par conséquent, le bloc positif $C \cup D$ n'est pas minimal. La thèse suit par la remarque 1. La démonstration se généralise facilement. ▪

3. Propriété.

Théorème 2.6 :

Si D est un sous-ensemble d'un espace affine et si $\psi_i(\theta) = 0$, $i = 1, \ldots, m$, *alors D est un bloc (minimal) relatif à* $\{\psi_1, \ldots, \psi_m\}$ *ssi* $\{\theta\} \cup D$ *est un ensemble-H (minimal) relatif à W engendré par* $\{1, \psi_1, \ldots, \psi_m\}$.

Si $(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_m)$ *est une configuration de signes de* $\{\theta\} \cup D$ *alors le bloc D a la configuration* $(\varepsilon_1, \ldots, \varepsilon_m)$.

Ce théorème est une conséquence directe des définitions d'ensemble-H et de bloc.

Il a de plus une grande importance dans la construction des ensembles-H relatifs à V_s (2.2) car les hypothèses du théorème 2.6 sont vérifiées. On passera

donc facilement à un bloc, à partir des ensembles-H déjà construits, pour élaborer d'autres ensembles-H.

On prend un exemple simple qui illustre cette remarque. La configuration de la fig. 2.10 forme un ensemble-H relatif à V_1^4 (2.3). En fixant l'origine θ, en un de ces points, on forme un bloc positif ou négatif suivant que ε_o est négatif ou positif (fig. 2.11 et 2.12).

2.2.2. Méthode de construction suivant le signe du Jacobien.

Cette méthode, introduite par H. SHAPIRO [14], se base sur le théorème suivant qui se démontre en utilisant une formule d'Euler et Jacobi.

1. Théorème :

Théorème 2.7 :

Soient $P_1(\xi)$, $P_2(\xi)$, ..., $P_k(\xi)$, k *polynômes de degré* $m_1 . m_2 . \ldots . m_k$ *respectivement en* $\xi = (x_1, \ldots, x_k)$ *qui vérifient les deux propriétés :*

(i) Le système d'équations

$$P_1(\xi) = P_2(\xi) = \ldots = P_k(\xi) = 0$$

a exactement $M = m_1 . m_2 . \ldots . m_k$ *racines distinctes* ξ^r *dans l'espace Euclidien réel de dimension* k.

(ii) Les nombres J_r $(r = 1, 2, \ldots, M)$ *sont tous différents de zéro, où* J_r *désigne la valeur du Jacobien* $(\partial P_i / \partial x_j)$ *au point* $\xi = \xi^r$. *Alors, aucun polynôme* $f(\xi)$ *de degré* $\leq m_1 + m_2 + \ldots + m_k - k - 1$ *n'est positif aux points* ξ^r *où* $J_r > 0$ *et n'est négatif aux points* ξ^r *où* $J_r < 0$.

Démonstration :

Grâce à une formule d'Euler et Jacobi (la démonstration se trouve dans [8], p. 135) et sous les hypothèses du théorème 2.7, on a :

$$\sum_{r=1}^{M} \frac{f(\xi^r)}{J_r} = 0 \qquad (2.11)$$

pour tout polynôme $f(\xi)$ de degré $\leq m_1 + m_2 + \ldots + m_k - k - 1$. On suppose que f est non négative où $J_r > 0$ et non positive où $J_r < 0$, il s'ensuit que tous les ter-

mes de la somme (2.11) sont positifs ou nuls. Et donc pour satisfaire l'égalité (2.11), f doit s'annuler en tous les ξ^r, ce qui est absurde. ▪

Il est clair que l'ensemble D des M racines distinctes ξ^r est un ensemble-H relatif à V : sous-espace des polynômes de degré $\leqslant m_1 + m_2 + \ldots + m_k - k - 1$.

2. Application de ce théorème à la construction d'ensembles-H.

On décrit la méthode dans le cas où k = 2 et on écrit

$$x_1 = x \quad \text{et} \quad x_2 = y$$

On choisit pour les $P_i(x, y)$ un produit de m_i polynômes linéaires (i = 1, 2) :

$$P_i(x, y) = \prod_{j=1}^{m_i} (a_{ij}\, x + b_{ij}\, y) \qquad i = 1, 2 \quad a_{ij}, b_{ij} \in \mathbb{R}$$

On obtient donc une figure (fig. 2.13) où m_1 droites rouges (représentées en tirets) coupent m_2 droites noires (représentées en trait plein) en $m_1\, m_2$ points.

Aucun calcul n'est nécessaire pour déterminer le signe du Jacobien en chacun de ces points. Car, considérant la paire de fonctions P_1 et P_2 comme définissant une transformation, on a seulement à déterminer si en chacun des points ξ^r, le sens de rotation est préservé ou renversé par cette transformation.

Dans ce but, on considère les régions délimitées par les droites, on en prend une sur le pourtour et on suppose que P_1 et P_2 sont normalisés de sorte que chacun soit positif dans cette région. On écrit dans cette région (+, +) et on continue à marquer les autres régions avec les quatre couples de signe (+, +), (+, -), (-, +), (-, -) en respectant le schéma suivant :

- si deux régions se touchent le long d'une ligne rouge, leurs couples diffèrent en la première position;
- si elles se touchent le long d'une ligne noire, leurs couples diffèrent en la seconde position.

On obtient donc la fig. 2.14.

Quand toutes les régions ont été marquées, on considère un point quelconque où une droite noire rencontre une droite rouge. On dessine un cercle, autour de ce point, qui passe de la région (+, +) à la région (+, -) : (fig. 2.15). Suivant que le sens de rotation du cercle est le même que celui de rotation des aiguilles d'une montre, ou opposé à ce sens, on divise l'ensemble de ces points ξ^r en deux classes.

Les points de ces deux classes peuvent être pris comme étant les points blancs et noirs d'une configuration telle qu'aucun polynôme de degré $\leqslant m_1 + m_2 - 3$ n'est positif sur les points noirs et négatif sur les points blancs (fig. 2.16). On trouvera à la fig. 2.17 un autre exemple, toujours avec $m_1 = 2$, $m_2 = 3$.

Cette construction peut évidemment être utilisée pour des courbes de tout degré, pourvu qu'elles se coupent en le nombre de points demandé et, avec les modifications convenables, à des surfaces, ...

Exemple : soit K une ellipse; la configuration formée par 2n points de K, marqués alternativement blanc et noir est un ensemble-H relatif à V_2^{n-1} (2.3).

En effet, on divise les 2n points en paires, de façon quelconque, et on fait passer une droite par chacune des paires de points (fig. 2.18 et 2.19). Le système de droites, de degré n, a 2n intersections distinctes avec l'ellipse. Par conséquent le théorème 2.7 s'applique.

Remarques : 1. Tous les ensembles-H ne sont pas accessibles par cette méthode. Pour k = 1 ainsi que pour le degré 1 quand k = 2, la méthode donne tous les ensembles-H.

exemples : - k = 1, degré 3 fig. 2.20
- k = 2, degré 1 fig. 2.21 et 2.22

Dans le cas général, ceci n'est plus vrai. En effet, si on se place dans le cas du degré 2 pour une approximation linéaire d'une fonction bivariable : k = 2.
D'après le théorème 2.7, $m_1 + m_2 - 2 - 1 = 2$.
Par conséquent, $m_1 = 4$ et $m_2 = 1$ ou $m_1 = 3$ et $m_2 = 2$ sont les seules possibilités. Autrement dit, seuls des ensembles-H de 4 ou 6 points sont possibles, alors que les ensembles-H minimaux relatifs

à cette approximation contiennent sept points (fig. 2.23).

2. On ne peut pas encore affirmer que toutes les configurations obtenues à partir du théorème 2.7 sont minimales.

2.2.3. Méthode de construction basée sur le produit cartésien.

Cette méthode a été introduite par NEWMAN et SHAPIRO [12].

1. Hypothèses.

Avec les hypothèses suivantes, on pourra construire des ensembles-H relatifs à l'approximation de fonctions continues sur un produit cartésien de K domaines compacts $B_i \subset \mathbf{R}$. On note

$$W_i = \text{eng}\,\{g_1^i, \ldots, g_{m_i}^i\} \qquad g_j^i \in C[B_i, \mathbf{R}] \quad j = 1, \ldots, m_i \quad i = 1, \ldots, K$$

et

$$W = \text{eng}\,\{g_i^1\, g_j^2 \ldots g_l^K\} \qquad \begin{matrix} i = 1, \ldots, m_1 \\ j = 1, \ldots, m_2 \\ \vdots \\ l = 1, \ldots, m_K \end{matrix} \tag{2.12}$$

De plus, on suppose que pour tout i, D_i est un ensemble-H relatif à W_i dont la configuration de signes est $(\varepsilon_1^i, \ldots, \varepsilon_{m_i}^i)$, i = 1, ..., K.

2. Description de la méthode dans le cas où K = 2.

La méthode se généralise facilement pour un K quelconque. Afin de rendre la description plus facile, on prendra B_1 sur l'axe X et B_2 sur l'axe Y et on notera :

$$g_j^1(x_1) = g_j(x) \quad j = 1, \ldots, m_1 \qquad \text{et} \qquad g_k^2(x_2) = h_k(y) \quad k = 1, \ldots, m_2$$

Les configurations de signes de D_1 et D_2 seront notées respectivement $(\varepsilon_1, \ldots, \varepsilon_{n_1})$ et $(\mu_1, \ldots, \mu_{n_2})$.

Par la définition duale d'ensemble-H (1.6), il existe n_1 scalaires p_i positifs et n_2 scalaires q_l positifs qui vérifient

$$\sum_{i=1}^{n_1} p_i \varepsilon_i g_j(x_i) = 0 \qquad j = 1, \ldots, m_1,\ x_i \in D_1 \tag{2.13}$$

$$\sum_{l=1}^{n_2} q_l \mu_l h_k(y_l) = 0 \qquad k = 1, \ldots, m_2,\ y_l \in D_2 \tag{2.14}$$

Décomposant (2.13) et (2.14) en deux sommations, on a en posant

$M_1 = \{1, 2, \ldots, n_1\}$ et I une partie quelconque de M_1

$M_2 = \{1, 2, \ldots, n_2\}$ et L une partie quelconque de M_2

$$\sum_{i\in I} p_i \varepsilon_i g_j(x_i) = - \sum_{i'\in M_1 \setminus I} p_{i'} \varepsilon_{i'} g_j(x_{i'}) \tag{2.15}$$

$$\sum_{l\in L} q_l \mu_l h_k(y_l) = - \sum_{l'\in M_2 \setminus L} q_{l'} \mu_{l'} h_k(y_{l'}) \tag{2.16}$$

Multipliant ces deux égalités membre à membre, la somme (2.17) s'annule.

$$\sum_I \sum_L p_i q_l \varepsilon_i \mu_l g_j(x_i) h_k(y_l) - \sum_{M_1\setminus I} \sum_{M_2\setminus L} p_{i'} q_{l'} \varepsilon_{i'} \mu_{l'} g_j(x_{i'}) h_k(y_{l'}) = 0 \tag{2.17}$$

Après regroupement et moyennant les notations :

$$t_{il} = p_i q_l \qquad (i, l) \in I \times L \text{ ou } (i, l) \in (M_1 \setminus I) \times (M_2 \setminus L)$$

$$\eta_{il} = \varepsilon_i \mu_l \qquad (i, l) \in I \times L$$

$$\eta_{il} = - \varepsilon_i \mu_l \qquad (i, l) \in (M_1 \setminus I) \times (M_2 \setminus L)$$

$$f_{jk}(x_i, y_l) = g_j(x_i) \cdot h_k(y_l)$$

(2.17) s'écrit :

$$\sum_{(i,l)\in (I\times L)\cup((M_1\setminus I)\times(M_2\setminus L))} t_{il} \eta_{il} f_{j,k}(x_i, y_l) = 0 \qquad \begin{matrix} j = 1, \ldots, m_1 \\ k = 1, \ldots, m_2 \end{matrix} \tag{2.18}$$

avec n_{IL} scalaires positifs t_{il} où

$$n_{IL} = (\text{card } I \cdot \text{card } L) + (\text{card } (M_1 \setminus I) \cdot \text{card } (M_2 \setminus L))$$

et tel $|n_{il}| = 1$, les $f_{j,k}(x, y)$ étant les fonctions de base de l'espace vectoriel W (2.12) pour K = 2.

On a donc construit un nouvel ensemble-H composé de n_{IL} points relatif à W :

$$D_{xy} = \{(x_i, y_l) \mid (i, l) \in (I \times L) \cup ((M_1 \setminus I) \times (M_2 \setminus L))\}$$

et dont la configuration de signes est :

$$(n_{il} \mid (i, l) \in I \times L \cup ((M_1 \setminus I) \times (M_2 \setminus L))).$$

3. Exemples.

On prend pour I, l'ensemble des indices des points $x_i \in D_1$ tels que $\varepsilon_i = +1$ et L, l'ensemble des indices des points $y_l \in D_2$ tels que $\mu_l = +1$ où D_1 et D_2 sont des ensembles-H minimaux relatifs à V_1^n (2.3).
Alors $D_{xy} = \{(x_i, y_l) \mid \varepsilon_i = \mu_l\}$ est un ensemble-H dont la configuration de signes est :

$$(n_{il} \mid \begin{array}{l} n_{il} = +1 \text{ si } \varepsilon_i = \mu_j = +1 \\ n_{il} = -1 \text{ si } \varepsilon_i = \mu_j = -1) \end{array}$$

La fig. 2.24 représente D_{xy} dans le cas où n = 2 et 2.25 dans le cas où n = 3.

Remarque : Les points sur les axes ne font pas partie de l'ensemble-H.

D'autres types d'ensembles-H sont représentés aux fig. 2.26 pour n = 2, K = 2.

Remarques :

1. Tous les ensembles-H ne sont pas accessibles par cette méthode.
   Les ensembles-H relatifs aux sous-espaces W qui ne sont pas de la forme (2.12) ne sont pas accessibles; exemple : $V_2^1 = \text{eng}\{1, x, y\}$.
2. Les ensembles-H obtenus par cette méthode à partir d'ensembles-H même minimaux ne sont pas tous minimaux.
   Soit V = eng {1, x, y, xy}. L'ensemble D représenté à la fig. 2.27 est un ensemble-H relatif à V mais n'est pas minimal puisque

$$\text{rg } U_D \leqslant 4 < \text{card } D-1 = 8.$$

2.2.4. Méthode se basant sur le passage aux coordonnées polaires.

GEARHART [7] introduit cette méthode pour construire des ensembles-H relatifs à V_2 (2.2). Il a montré, de plus, comment obtenir des ensembles-H supplémentaires en utilisant la méthode basée sur le produit cartésien.

1. Résultats de base.

Soit V_2^n (2.3) et une rotation du plan d'angle $\frac{2\pi}{m}$. On définit U_m^n : c'est l'ensemble des polynômes de V_2^n invariants par cette rotation notée $R(\frac{2\pi}{m})$

$$U_m^n = \{p \in V_2^n \mid p \circ R(\tfrac{2\pi}{m}) = p\} \qquad (2.19)$$

Les fonctions de base de V_2^n s'écrivent en coordonnées polaires :

$$r^{i+j} \cos^i \phi \sin^j \phi \qquad i+j \leqslant n \quad i \geqslant 0 \text{ et } j \geqslant 0$$

Utilisant les formules trigonométriques permettant le passage des puissances de cosinus et sinus à une somme de cosinus et de sinus, on peut écrire :

$$V_2^n = \text{eng}\{\{(r^2)^l r^p \cos p\phi \mid 2l + p \leqslant n \quad 0 \leqslant l, 0 \leqslant p\}$$
$$\cup \{(r^2)^l r^p \sin p\phi \mid 2l + p \leqslant n \quad 0 \leqslant l, 0 \leqslant p\}\} \qquad (2.20)$$

Par conséquent, U_m^n est le sous-espace engendré par les fonctions de base de V_2^n (2.20) périodiques de période $\frac{2\pi}{m}$.

Autrement dit,

$$U_m^n = \text{eng}\{\{(r^2)^l r^{km} \cos km\phi \mid 0 \leqslant l \leqslant [\tfrac{n-km}{2}], 0 \leqslant k \leqslant [\tfrac{n}{m}]$$
$$\cup \{(r^2)^l r^{km} \sin km\phi \mid 0 \leqslant l \leqslant [\tfrac{n-km}{2}], 0 \leqslant k \leqslant [\tfrac{n}{m}]\}\} \qquad (2.21)$$

Soit S_m, un secteur du plan, d'angle $\frac{2\pi}{m}$ que l'on définit en coordonnées polaires par

$$S_m = \{(r, \phi) \mid r \geqslant 0 \text{ et } 0 \leqslant \phi < \tfrac{2\pi}{m}\} \qquad (2.22)$$

La méthode est basée sur le théorème suivant.

Théorème 2.8 :

Soit $D = \{P_1, P_2, \ldots, P_q\}$ *un ensemble-H relatif à* U_m^n *dont la configuration de signes est* $(\varepsilon_1, \ldots, \varepsilon_q)$ *et tel que* $P_i \in S_m$, $i = 1, \ldots, q$. *Alors*

$$D' = D \cup \left(\bigcup_{s=1}^{m-1} R\left(\frac{2\pi s}{m}\right)(D) \right) \tag{2.23}$$

est un ensemble-H relatif à V_2^n *dont la configuration de signes est un* m.q-*uplet* $(\mu_1, \mu_2, \ldots, \mu_{m.q})$ *dont les coordonnées vérifient*

$$\mu_i = \mu_{q+i} = \mu_{2q+i} = \ldots = \mu_{(m-1)q+i} = \varepsilon_i \qquad 1 \leqslant i \leqslant q \tag{2.24}$$

La démonstration de ce théorème découle du critère d'invariance (théorème 1.1). Par ce critère, $R(\frac{2\pi S}{m})(D)$ $S = 1, \ldots, \infty$ est un ensemble-H relatif à U_m^n et a la même configuration que D.

Il suit que D' (2.23) est un ensemble-H relatif à U_m^n (propriété 1.6). Par la formulation duale (1.6), il existe mq scalaires positifs p_i qui vérifient

$$\begin{aligned} &\sum_{i=1}^{mq} p_i \mu_i \, (r^{2l} \, r^{km} \cos km\phi)(Q_i) = 0 \\ \text{et} \qquad &\sum_{i=1}^{mq} p_i \mu_i \, (r^{2l} \, r^{km} \sin km\phi)(Q_i) = 0 \end{aligned} \tag{2.25}$$

avec $Q_i \in D'$, $0 \leqslant l \leqslant [\frac{n-km}{2}]$, $0 \leqslant k \leqslant [\frac{n}{m}]$

(2.25) peut encore s'écrire en utilisant les hypothèses et en exprimant Q_i en fonction des $P_j = (\rho_j, \theta_j) \in D$:

$$\begin{aligned} &\sum_{S=o}^{m-1} \sum_{j=1}^{q} q_j \varepsilon_j \rho_j^{2l+km} \cos km \left(\theta_j + \frac{2\pi S}{m}\right) = 0 \\ \text{et} \qquad &\sum_{S=o}^{m-1} \sum_{j=1}^{q} q_j \varepsilon_j \rho_j^{2l+km} \sin km \left(\theta_j + \frac{2\pi S}{m}\right) = 0 \end{aligned} \tag{2.26}$$

avec $0 \leqslant l \leqslant [\frac{n-km}{2}]$, $0 \leqslant k \leqslant [\frac{n}{m}]$ et $q_j > 0$, $j = 1, \ldots, q$.

Mais D' (2.23) est aussi un ensemble-H relatif à V_2^n; en effet (1.6) est vérifiée par les $q_j > 0$, $j = 1, \ldots, q$ de (2.26).

(1.6) s'écrit en utilisant les hypothèses :

$$\sum_{S=o}^{m-1} \sum_{j=1}^{q} q_j \varepsilon_j \rho_j^{2l+p} \cos p(\theta_j + \frac{2\pi S}{m}) = 0$$
$$\sum_{S=o}^{m-1} \sum_{j=1}^{q} q_j \varepsilon_j \rho_j^{2l+p} \sin p(\theta_j + \frac{2\pi S}{m}) = 0 \qquad (2.27)$$

$$0 \leqslant l \leqslant [\frac{n-p}{2}] \ , \ 0 \leqslant p \leqslant n$$

En permutant les signes de sommation et en utilisant les formules trigonométriques (2.28) et (2.29)

$$\sum_{k=o}^{n-1} \cos(\alpha + k\delta) = \frac{\cos(\alpha + \frac{n-1}{2}\delta) \sin\frac{n\delta}{2}}{\sin\frac{\delta}{2}} \qquad (2.28)$$

$$\sum_{k=o}^{n-1} \sin(\alpha + k\delta) = \frac{\sin(\alpha + \frac{n-1}{2}\delta) \sin\frac{n\delta}{2}}{\sin\frac{\delta}{2}} \qquad (2.29)$$

on a la thèse. ▪

Par conséquent, on peut construire des ensembles-H relatifs à V_2^n à partir d'ensembles-H sur S_m et relatifs à U_m^n.

2. Cas particulier.

Le cas où $n \leqslant m-1$ est un cas intéressant à étudier. Les polynômes de U_m^n se réduisent à la forme $p(r^2)$ avec p un polynôme à une variable de degré $[n/2]$. Il est facile de voir que tout ensemble-H : D sur S_m relatif à U_m^n doit être d'un des deux types suivants :

1. $D = \{(r, \phi_1), (r, \phi_2) \mid r > 0 \text{ et } \varepsilon_1 \varepsilon_2 = -1\}$

2. $D = \{x_1, \ldots, x_N \mid x_i = (r_i, \phi_i), N = [n/2] + 2$
$0 < r_1 < r_2 < \ldots < r_N$
$\varepsilon_{i-1} \cdot \varepsilon_i = -1\}$

En effet, p est un polynôme à une seule variable, de degré $[n/2]$; les ensembles-H doivent donc satisfaire aux conditions du théorème d'oscillation.

Découlant de la proposition précédente, les ensembles-H relatifs à V_2^n correspondant aux ensembles-H du type 1 consistent en au moins 2n+2 points, de signes alternés, sur la circonférence de rayon r.
Par conséquent, ces ensembles-H sont un cas particulier d'un ensemble-H bien connu (fig. 2.18,2.19, en prenant K = cercle de rayon r) ou (COLLATZ [4], p. 52).
Par contre, la plupart des ensembles-H relatifs à V_2^n correspondant aux ensembles-H du type 2 sont nouveaux.
Exemple : l'ensemble-H représenté à la fig. 2.28 obtenu dans le cas n = 2, m = 3 ne peut pas être réalisé par la méthode de SHAPIRO.

4. D'autres ensembles-H plus généraux.

En général, la méthode basée sur le produit cartésien (2.2.3) permet d'obtenir beaucoup d'ensembles-H pour les polynômes. D. NEWMAN et SHAPIRO utilisèrent cette méthode pour des polynômes écrits sous la forme d'un produit cartésien.

Il est clair que cette méthode peut être appliquée aux polynômes de V_2^n écrits en coordonnées polaires. Par exemple, on suppose que $\{0, r_1, \ldots, r_m \mid r_i > 0, i = 1, \ldots, m\}$ et $\{\phi_1, \ldots, \phi_{m'} \mid 0 \leqslant \phi_j < 2\pi\}$ sont des ensembles-H. Autrement dit, il existe des scalaires positifs p_i et q_j, non tous nuls, qui vérifient respectivement

$$p(o) + \sum_{i=1}^{m} p_i \varepsilon_i p(r_i) = 0 \qquad (2.30)$$

pour tout p polynôme de degré n à une variable,

$$\sum_{l=1}^{m'} q_l \mu_l t(\phi_l) = 0 \qquad (2.31)$$

pour tout polynôme trigonométrique de degré n.

Soient $I \subset M = \{1, \ldots, m\}$ arbitraire et $L \subset M' = \{1, \ldots, m'\}$, on décompose en deux sommes chacune des équations (2.30) et (2.31). On multiplie membre à membre et on arrange les termes comme en 2.2.3.2.
On normalise enfin pour que le scalaire correspondant à l'origine soit +1. On obtient l'égalité suivante :

$$\sum_{(i,l)\in(I\times L)\cup((M\setminus I)\times(M'\setminus L))} t_{il}\,\eta_{il}\,p(r_i\cos\phi_l,\ r_i\sin\phi_l) = 0 \qquad (2.32)$$

Et donc l'ensemble contenant les points $(r_i \cos \phi_l, r_i \sin \phi_l)$, $(i, l) \in (I \times L) \cup ((M \setminus I) \times (M' \setminus L))$ est un ensemble-H relatif aux polynômes de U_m^n.

Exemples : deux ensembles-H obtenus par la méthode du produit cartésien sont reproduits aux fig. 2.29 où on a pris $n = 4$ et 2.30 où $n = 5$.

CHAPITRE 3 :
Les ensembles-H pour une approximation multivariable de degré 1.

Dans ce chapitre, on se propose d'étudier plus en détail, les ensembles-H minimaux obtenus par une approximation polynomiale du premier degré. On définit le sous-espace linéaire W par :

$$W = \text{eng}\ \{1, x_1, \ldots, x_n\} \qquad (3.1)$$

Dans un premier temps, on décrit quelques résultats obtenus grâce à la forme particulière de W. TAYLOR en [15] a établi la majorité de ceux-ci.

3.1. Théorèmes de base.

Les résultats suivants sont extraits de [6] et de [15]. On considère un sous-ensemble $M = \{P_1, \ldots, P_m\}$ de l'espace affine $\mathbf{R}^n$ muni du système de coordonnées $0; x_1, \ldots, x_n$. On suppose que M est un ensemble-H et on note $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)$ où $|\varepsilon_i| = 1$ pour $i = 1, \ldots, m$ la configuration de signes associée à M.

On peut décomposer M en deux ensembles M^+ et M^- définis par :

$$M^+ = \{P_i \in \mathbf{R}^n \text{ tel que : } P_i \in M \text{ et } \varepsilon_i = +1\}$$

$$M^- = \{P_i \in \mathbf{R}^n \text{ tel que : } P_i \in M \text{ et } \varepsilon_i = -1\}$$

Cette décomposition de $M = M^+ \cup M^-$ est unique si M est un ensemble-H minimal, car dans ce cas, M ne possède que deux configurations de signes possibles et $M^+ \cap M^- = \emptyset$.

Théorème 3.1 :

Si $M = M^+ \cup M^-$ est un ensemble-H minimal, alors il n'existe pas d'hyperplan de $\mathbf{R}^n$ qui sépare M^+ et M^-.

Démonstration :

Raisonnant par l'absurde, on suppose qu'il existe un hyperplan H défini par une fonctionnelle linéaire et continue L de $\mathbf{R}^n$ dans $\mathbf{R}^n$ qui sépare les ensembles M^+ et M^-. En d'autres mots, on suppose qu'il existe un hyperplan $H = \{x \in \mathbf{R}^n$ tel que $L(x) = 0\}$ qui vérifie :

$$L(x) > 0 \quad \text{pour tout} \quad x \in M^+$$

$$L(x) < 0 \quad \text{pour tout} \quad x \in M^-$$

La forme particulière des fonctionnelles linéaires et continues de R^n dans R^n permet de définir l'hyperplan H comme suit :

$$H = \{x \in R^n \text{ tel que } \sum_{i=1}^{n} a_i x_i = 0\},$$

où $(a_1, \ldots, a_n) = a$ est un n-uplet de réels.

Donc si H sépare M^+ et M^-, on a :

$$\sum_{i=1}^{n} a_i x_i > 0 \qquad \text{pour tout } x = (x_1, \ldots, x_n) \in M^+$$

$$\sum_{i=1}^{n} a_i x_i < 0 \qquad \text{pour tout } x = (x_1, \ldots, x_n) \in M^-$$

On note $g_o = 1$,

$g_i = x_i$: la (i+1)ème fonction de base définissant W.

Les inégalités précédentes s'écrivent :

$$\sum_{i=1}^{n} a_i g_i(x) > 0 \qquad \text{pour tout } x \in M^+$$

$$\sum_{i=1}^{n} a_i g_i(x) < 0 \qquad \text{pour tout } x \in M^-$$

En considérant $a_o = 0$, on conclut : il existe $a = (a_o, a_1, \ldots, a_n) \in R^{n+1}$ satisfaisant à :

$$\sum_{i=o}^{n} a_i g_i(x) > 0 \qquad \text{pour tout } x \in M^+$$

$$\sum_{i=o}^{n} a_i g_i(x) < 0 \qquad \text{pour tout } x \in M^-$$

Par la définition d'ensemble-H donnée par COLLATZ, on obtient une contradiction avec l'hypothèse du théorème. ▪

On adopte la notation Co(D) pour désigner l'enveloppe convexe d'un ensemble D de R^n. Plus précisément,

$$Co(D) = \{x \in R^n \text{ tel que } x = \sum_{i=1}^{m} \lambda_i x_i\}$$

où : - $\sum_{i=1}^{m} \lambda_i = 1$,

- $(\lambda_i > 0;\ x_i \in D)$ pour $i = 1, \ldots, m$.

Une conséquence immédiate du théorème précédent est :

Corollaire 3.1 :

Si $M = M^+ \cup M^-$ *est un ensemble-H minimal où* $\overline{Co}(M)$ *est la fermeture de* $Co(M)$, *alors* $\overline{Co}(M^+) \cap \overline{Co}(M^-) \neq \emptyset$.

Démonstration :

Raisonnant par l'absurde, on suppose que :

$$\overline{Co}(M^+) \cap \overline{Co}(M^-) = \emptyset$$

Les hypothèses du théorème de séparation des convexes [13] sont satisfaites. Par conséquent, il existe un hyperplan de R^n qui sépare $\overline{Co}(M^+)$ et $\overline{Co}(M^-)$. A fortiori, il existe un hyperplan de R^n qui sépare M^+ et M^-. ■

Pour être plus précis, on démontre :

Théorème 3.2 :

On considère un sous-ensemble $M = M^+ \cap M^-$ *tel que* card $M = m$. *Soit l'homomorphisme* U_M *défini en (1.7). Il existe* $\lambda = (\lambda_1, \ldots, \lambda_m)$, $\lambda \in \text{Ker } U_M$ *vérifiant*

$$\sum_{i=1}^{m} |\lambda_i| = 1 \text{ et } \lambda_i \neq 0 \quad \text{pour } i = 1, \ldots, m.$$

ssi il existe un et un seul $\xi \in R^n$ *tel que :*

$$\xi \in Co(M^+) \cap Co(M^-)$$

Démonstration :

1. Condition nécessaire :

On suppose qu'il existe $\lambda = (\lambda_1, \ldots, \lambda_m) \in \text{Ker } U_M$. Par la définition (1.7) de U_M, ceci se traduit par :

$$\sum_{i=1}^{m} \lambda_i\, g_j(p_i) = 0 \qquad \text{pour } j = 0, \ldots, n \qquad (3.2)$$

La forme particulière (3.1) de W permet d'écrire :

$$g_0(x) = 1 \qquad \text{pour tout } x \in R^n$$

La première équation du système (3.2) devient :

$$\sum_{i=1}^{m} \lambda_i = 0$$

On note : card $M^+ = m_1$,

card $M^- = m_2$,

$\lambda_i^+ = \lambda_i$ si $P_i \in M^+$,

$\lambda_i^- = \lambda_i$ si $P_i \in M^-$.

En ordonnant les points de M tel que les (m_1) premiers appartiennent à M^+, la première équation de (3.2) s'exprime par :

$$\sum_{i=1}^{m} \lambda_i = \sum_{i=1}^{m_1} \lambda_i^+ + \sum_{i=1}^{m_2} \lambda_{m_1+i}^- = 0$$

Or, par hypothèse, on a

$$\sum_{i=1}^{m} |\lambda_i| = \sum_{i=1}^{m_1} \lambda_i^+ - \sum_{i=1}^{m_2} \lambda_{m_1+i}^- = 1$$

Par conséquent,

$$\sum_{i=1}^{m_1} \lambda_i^+ = \frac{1}{2} \; ; \quad \sum_{i=1}^{m_2} \lambda_{m_1+i}^- = -\frac{1}{2}$$

On pose : $\alpha_i = 2\,\lambda_i^+$, $i = 1, \ldots, m_1$ et $\beta_i = -2\,\lambda_{m_1+i}^-$, $i = 1, \ldots, m_2$.

Comme $\lambda_i \neq 0$, $i = 1, \ldots, m$, on a $\alpha_i, \beta_j > 0$, $i = 1, \ldots, m_1$, $j = 1, \ldots, m_2$.

Or, les n dernières équations du système (3.2) s'écrivent encore :

$$\sum_{i=1}^{m_1} \lambda_i^+ g_j(P_i) = -\sum_{i=1}^{m_2} \lambda_i^- g_j(P_{m_1+i}) \qquad j = 1, \ldots, n$$

$$2 \sum_{i=1}^{m_1} \lambda_i^+ g_j(P_i) = -2 \sum_{i=1}^{m_2} \lambda_i g_j(P_{m_1+i}) \qquad j = 1, \ldots, n$$

En employant les notations en α_i, β_i, on a :

$$\sum_{i=1}^{m_1} \alpha_i g_j(P_i) = \sum_{i=1}^{m_2} \beta_i g_j(P_{m_1+i}) \qquad j = 1, \ldots, n$$

Comme : $\sum_{i=1}^{m_1} \alpha_i = 2 \sum_{i=1}^{m_1} \lambda_i^+ = 1$ et $\sum_{i=1}^{m_2} \beta_i = -2 \sum_{i=1}^{m_2} \lambda_{m_1+i}^- = 1$,

et par la forme particulière des fonctions de base, l'élément $\xi = (\xi_1, \ldots, \xi_n)$ de R^n défini par :

$$\xi_j = \sum_{i=1}^{m_1} \alpha_i \, g_j(P_i) = \sum_{i=1}^{m_2} \beta_i \, g_j(P_{m_1+i}) \qquad j = 1, \ldots, n$$

satisfait à

$$\xi \in Co(M^+) \cap Co(M^-)$$

De plus, ce point ξ est unique car M est un ensemble-H minimal et donc λ est déterminé de manière unique.

2. Condition suffisante :

On suppose qu'il existe un et un seul $\xi \in \mathbb{R}^n$ vérifiant :

$$\xi \in Co(M^+) \cap Co(M^-)$$

On peut ordonner les points de M de telle manière que :

$$M^+ = \{P_1, \ldots, P_{m_1}\}$$

$$M^- = \{P_{m_1+1}, \ldots, P_{m_1+m_2}\}$$

Par la définition d'enveloppe convexe, on obtient :

$$\xi = \sum_{i=1}^{m_1} \alpha_i \, P_i \qquad \text{où} \quad \sum_{i=1}^{m_1} \alpha_i = 1; \; \alpha_i > 0 \text{ pour } i = 1, \ldots, m_1$$

$$\xi = \sum_{j=1}^{m_2} \beta_j \, P_{m_1+j} \qquad \text{où} \quad \sum_{j=1}^{m_2} \beta_j = 1; \; \beta_j > 0 \text{ pour } j = 1, \ldots, m_2$$

Les fonctions de base $\{g_0, \ldots, g_n\}$ étant de la forme (3.1), on a :

$$\sum_{i=1}^{m_1} \alpha_i \, g_k(P_i) = \sum_{j=1}^{m_2} \beta_j \, g_k(P_{m_1+j}) \qquad k = 1, \ldots, n$$

$$\sum_{i=1}^{m_1} \alpha_i = \sum_{j=1}^{m_2} \beta_j$$

d'où : $\sum_{i=1}^{m_1} \alpha_i \, g_k(P_i) - \sum_{j=1}^{m_2} \beta_j \, g_k(P_{m_1+j}) = 0 \qquad k = 0, \ldots, n$

On pose : $\lambda_i = \frac{\alpha_i}{2} \qquad i = 1, \ldots, m_1$

$$\lambda_{m_1+j} = -\frac{\beta_j}{2} \qquad j = 1, \ldots, m_2$$

Comme par définition $\alpha_i > 0$ pour $i = 1, \ldots, m_1$ et $\beta_j > 0$ pour $j = 1, \ldots, m_2$, on a $\lambda_i \neq 0 \quad i = 1, \ldots, m = m_1 + m_2$, on obtient :

$$\sum_{i=1}^{m_1} \frac{\alpha_i}{2} g_k(P_i) - \sum_{j=1}^{m_2} \frac{\beta_i}{2} g_k(P_{m_1+j}) = 0 \quad \text{pour } k = 0, \ldots, n$$

$$\sum_{i=1}^{m_1} \lambda_i g_k(P_i) + \sum_{j=1}^{m_2} \lambda_{m_1+j} g_k(P_{m_1+j}) = 0 \quad \text{pour } k = 0, \ldots, n$$

Donc $\lambda = (\lambda_1, \ldots, \lambda_m)$ est une solution non nulle du système suivant :

$$\sum_{i=1}^{m} \lambda_i g_k(P_i) = 0 \qquad k = 0, \ldots, n$$

En d'autres mots : $\lambda = (\lambda_1, \ldots, \lambda_m) \in \text{Ker } U_M$.

De plus : $\sum_{i=1}^{m} |\lambda_i| = \sum_{i=1}^{m_1} |\lambda_i| + \sum_{j=1}^{m_2} |\lambda_{m_1+j}|$

$$= \sum_{i=1}^{m_1} \frac{\alpha_i}{2} + \sum_{j=1}^{m_2} \frac{\beta_j}{2}$$

$$= \frac{1}{2} + \frac{1}{2} = 1. \blacksquare$$

On illustre les résultats précédents par quelques exemples simples :

1. $W = \text{eng}\{1, x, y\}$
   $M = M^+ \cup M^- = \{P_1, P_2, P_3\} \cup \{P_4\}$
   M est un ensemble-H minimal, par exemple celui de la fig. 3.1.
   On remarque que $\text{Co}(M^+) \cap \text{Co}(M^-) = \{P_4\}$.

2. $W = \text{eng}\{1, x, y, z\}$
   $M = M^+ \cup M^- = \{P_1, P_2, P_3\} \cup \{P_4, P_5\}$
   M est un ensemble-H minimal, par exemple celui de la fig. 3.2
   On remarque que $\text{Co}(M^+) \cap \text{Co}(M^-) = \{P\}$.

Dans le deux cas, il n'existe pas d'hyperplan séparant $\text{Co}(M^+)$ et $\text{Co}(M^-)$. On remarque facilement sur ces exemples que les enveloppes convexes de M^+ et M^- sont des sous-espaces de dimension $(m_1 - 1)$ et $(m_2 - 1)$.

Théorème 3.3 :

On note : $M^+ = \{P_1, \ldots, P_{m_1}\}$ et $M^- = \{Q_1, \ldots, Q_{m_2}\}$. *Si* $M = M^+ \cup M^-$ *est un ensemble-H minimal relatif à W donné par (3.1), alors :*

1. $(P_2 - P_1), \ldots, (P_{m_1} - P_1)$ *forment une famille de vecteurs linéairement indépendants;*
2. *De même pour* $(Q_2 - Q_1), \ldots, (Q_{m_2} - Q_1)$.

Démonstration :

On ne démontre que le premier résultat; la démonstration du second se faisant d'une manière analogue.

En raisonnant par l'absurde, on suppose que $(P_2 - P_1),\ldots, (P_{m_1} - P_1)$ forment une famille de vecteurs linéairement dépendants. Comme M est un ensemble-H minimal, par le théorème 3.2, on obtient : il existe $\xi \in \mathbb{R}^n$ tel que : $\xi \in Co(M^+) \cap Co(M^-)$. Donc

$$(\xi - P_1) \in Co(M^+ - P_1) \cap Co(M^- - P_1).$$

Or, $M^+ - P_1 = \{0, P_2 - P_1, \ldots, P_{m_1} - P_1\}$. Par hypothèse, $(P_2 - P_1, \ldots, P_{m_1} - P_1)$ sont linéairement dépendants. Ce qui entraîne que $M^+ - P_1$ est inclus dans un sous-espace de $\mathbb{R}^n$ de dimension $m_1 - 2$.

Par le théorème de CARATHEODARY [9], p. 74, on sait que si x est combinaison linéaire convexe de points de $\Gamma \subset \mathbb{R}^n$, alors x est combinaison linéaire de au plus (n+1) points de Γ. Or

$$\xi - P_1 \in Co(M^+ - P_1),$$

donc
$$\xi - P_1 = \sum_{i=1}^{m_1} \lambda_i (P_i - P_1)$$

où $\sum_{i=1}^{m_1} \lambda_i = 1$ avec au moins un λ_i nul. On note i_o l'indice du premier λ_i qui s'annule. On obtient :

$$\xi - P_i = \sum_{\substack{i=o \\ i\neq i_o}}^{m_1} \lambda_i (P_i - P_1) \qquad \text{avec } \lambda_i \geqslant 0 \quad i = 1, \ldots, m_1$$
$$\text{et } \sum_{\substack{i=1 \\ i\neq i_o}}^{m_1} \lambda_i = 1$$

$$\xi - P_1 = \sum_{\substack{i=1 \\ i\neq i_o}}^{m_1} \lambda_i P_i - \sum_{\substack{i=1 \\ i\neq i_o}}^{m_1} \lambda_i P_1$$

$$\xi = \sum_{\substack{i=1 \\ i\neq i_o}}^{m_1} \lambda_i P_i \quad .$$

Ce qui s'exprime, en notant $M^\star = M^+ - \{P_{i_o}\}$, par :

$$\xi \in Co(M^\star) \cap Co(M^-)$$

Par le théorème 3.2, on conclut : $M^* \cup M^-$ est un ensemble-H. Or

$$M^* \cup M^- \subset M.$$

On trouve une contradiction avec l'hypothèse du théorème. ■

Pour terminer ce paragraphe, on donne une propriété des ensembles-H relatifs aux sous-espaces linéaires W et V définis par :

$$W = \text{eng}\ \{1, x_1, \ldots, x_n\}$$

$$V = \text{eng}\ \{1, x_1, \ldots, x_{n+1}\}$$

Propriété 3.1 :

Si $M = \{P_1, P_2, \ldots, P_m\}$ *est un ensemble-H relatif à W (3.1), alors M est un ensemble-H relatif à* V.

Démonstration :

Par définition, M est un ensemble-H relatif à W ssi il existe $\lambda = (\lambda_1, \ldots, \lambda_m) \neq \theta_m$ tel que

$$\begin{pmatrix} 1 & , \ldots, & 1 \\ x_1(P_1) & , \ldots, & x_1(P_m) \\ \vdots & & \vdots \\ x_n(P_1) & , \ldots, & x_n(P_m) \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = \theta_{n+1}$$

Or, passer du sous-espace W au sous-espace V revient, en quelque sorte à considérer les points de M comme éléments de R^{n+1}. En ce faisant, la dernière composante de ces points est nulle. C'est-à-dire : $x_{n+1}(P_i) = 0$ pour i = 1, ..., m.

Donc l'équation ajoutée est de la forme : $0\ \lambda_1 + 0\ \lambda_2 + \ldots + 0\ \lambda_m = 0$. Le nouveau système admet encore une solution λ non nulle. M est donc un ensemble-H relatif à V. ■

Remarques :

1. Le résultat reste encore correct pour un degré quelconque fixé, c'est-à-dire : les ensembles-H pour l'approximation polynomiale à n variables, de degré k, sont encore des ensembles-H pour l'approximation polynomiale à (n+1) variables, de degré k.

2. Le résultat n'est plus correct si au lieu d'augmenter le nombre de variables, on augmente le degré. En d'autres termes, si on passe de l'approximation polynomiale à n variables de degré k, à l'approximation polynomiale à n variables de degré (k+1). Pour illustrer cette remarque, on donne un exemple numérique : soit un sous-espace linéaire W = eng {1, x},
soit un sous-ensemble M de $\mathbb{R}$: M = {0, 1, 2}.

Le système suivant

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \theta_2$$

admet une infinité de solutions : $\lambda_3(1, -2, 1)$. En posant $\lambda_3 = 1$, on obtient une solution $\lambda = (1, -2, 1)$ non nulle. Comme par définition : $\varepsilon_i = \lambda_i / |\lambda_i|$ si $\lambda_i \neq 0$, la configuration de signes est $\varepsilon = (+1, -1, +1)$. Donc, M est un ensemble-H relatif à W.

On considère V = eng $\{1, x, x^2\}$ et on garde le même ensemble M = {0, 1, 2}. Le système suivant

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \theta_3$$

admet la solution triviale, car le déterminant de la matrice vaut : 4 - 2 = 2 $\neq 0$. Donc M n'est pas un ensemble-H relatif à V.

3.2. Choix d'une méthode de construction.

Les différents résultats obtenus au paragraphe précédent permettent de ne considérer, par la suite, que les ensembles-H minimaux d'un type particulier. Le raisonnement suivant est extrait de [6].

1. Par le théorème 3.2, on sait que : $M = M^+ \cup M^-$ est un ensemble-H minimal ssi il existe un et un seul $\xi \in Co(M^+) \cap Co(M^-)$.
Or on a montré que les ensembles-H minimaux sont invariants par rapport aux transformations affines de $GA(n, \mathbb{R})$. En particulier, on a invariance par rap-

port aux translations. La condition nécessaire du théorème 3.2 s'énonce donc :

$$\theta \in Co(M^+) \cap Co(M^-)$$

2. Il est aisé de voir qu'il suffit d'étudier les ensembles-H minimaux de $\mathbf{R}^n$ vérifiant : $\quad$ card $M = n + 2$

En effet, la propriété 1.2 appliquée à W (3.1) donne : card $M \leqslant n + 2$.

D'autre part, on sait que les ensembles-H de $\mathbf{R}^{n-1}$ sont des ensembles-H de $\mathbf{R}^n$.

Il suffit donc de considérer les ensembles-H M minimaux de $\mathbf{R}$ vérifiant : card $M = n + 2$.

3. Par le théorème 3.3, on obtient :

$$Co(M^+) \subseteq \mathbf{R}^{m_1} \qquad \text{où } m_1 = \text{card } M^+$$

$$Co(M^-) \subseteq \mathbf{R}^{m_2} \qquad \text{où } m_2 = \text{card } M^-$$

En résumé, il suffit d'étudier les ensembles-H minimaux de $\mathbf{R}^n$ de la forme : $M = M^+ \cup M^-$ tel que

$$\begin{aligned} &(1)\ \theta \in Co(M^+) \cap Co(M^-) \\ &(2)\ Co(M^+) \subseteq \mathbf{R}^{m_1} \quad \text{où } m_1 = \text{card } M^+ \\ &\qquad Co(M^-) \subseteq \mathbf{R}^{m_2} \quad \text{où } m_2 = \text{card } M^- \\ &(3)\ \mathbf{R}^n = R^{m_1-1} \oplus R^{m_2-1} \end{aligned} \tag{3.3}$$

L'analyse des ensembles-H minimaux de $\mathbf{R}^n$ de la forme (3.3) guide le choix d'une méthode de construction. On ordonne les points de M tel que :

$$M^+ = \{P_1, \ldots, P_{m_1}\}$$

$$M^- = \{P_{m_1}, \ldots, P_{m_1+m_2}\}$$

L'homomorphisme U_M associé à un ensemble-H M minimal est défini par :

$$U_M(\lambda_1, \ldots, \lambda_{m_1+m_2}) = \left(\sum_{i=1}^{m_1+m_2} \lambda_i\, g_o(P_i), \ldots, \sum_{i=1}^{m_1+m_2} \lambda_i\, g_n(P_i) \right)$$

Par (3.1), on a : $g_o(x) = 1$ pour tout $x \in M$; donc

$$\sum_{i=1}^{m_1+m_2} \lambda_i\, g_o(P_i) = \sum_{i=1}^{m_1+m_2} \lambda_i$$

$$\sum_{i=1}^{m_1+m_2} \lambda_i \, g_k(P_i) = \sum_{i=1}^{m_1+m_2} \lambda_i \, x_k(P_i) \qquad 1 \leqslant k \leqslant n$$

Comme $Co(M^+) \subseteq R^{m_1}$ et $R^n = R^{m_1-1} \oplus R^{m_2-1}$, on obtient :

$$x_k(P_i) = 0 \qquad \begin{array}{l} \text{pour } i = m_1+1, \ldots, m_2 \\ \text{pour } K = 1, \ldots, m_1-1 \end{array}$$

De manière analogue, on a :

$$x_k(P_i) = 0 \qquad \begin{array}{l} \text{pour } i = 1, \ldots, m_1 \\ \text{pour } k = m_1, \ldots, m_1+m_2-2 = n+1 \end{array}$$

La représentation matricielle $M(U_M)$ de l'homomorphisme U_M est de la forme :

$$M(U_M) = \left(\begin{array}{c|c} 1 \ldots 1 & 1 \ldots 1 \\ \hline M(u) & 0 \\ \hline 0 & M(u) \end{array}\right) \qquad (3.4)$$

où $u_1 \in \text{Hom}\,(R^{m_1}, R^n)$ et $u_2 \in \text{Hom}\,(R^{m_2}, R^n)$.

Par le chapitre précédent, on conclut que M^+ et M^- sont deux blocs, l'un positif, l'autre négatif. De plus, par le paragraphe 2.2.1.2, le bloc minimal positif M^+, noté M_1, est obtenu de l'ensemble-H minimal $\{\theta\} \cup M_1$ relatif à $w_1 = \text{eng}\,\{1, x_1, \ldots, x_{m_1-1}\}$, en omettant un point, en l'occurrence θ. De même, le bloc minimal négatif M^-, noté M_2, est obtenu de l'ensemble-H minimal $\{\theta\} \cup M_2$ relatif à $w_2 = \text{eng}\,\{1, x_{m_1}, \ldots, x_n\}$.

Réciproquement, on a la propriété suivante :

Propriété 3.2 :

Si M_1 est un bloc positif minimal obtenu d'un ensemble-H minimal relatif à $W_1 = \text{eng}\,\{1, x_1, \ldots, x_{m_1-1}\}$

si M_2 est un bloc négatif minimal obtenu d'un ensemble-H minimal relatif à $W_2 = \text{eng}\,\{1, x_{m_1}, \ldots, x_n\}$

alors $M = M_1 \cup M_2$ est un ensemble-H relatif à W (3.1).

Démonstration :

On note : $M_1 = \{P_1, \ldots, P_{m_1}\}$. Par hypothèse, $\{\theta\} \cup M_1$ est un ensemble-H minimal relatif à W_1. Donc, il existe $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{m_1}) \in R^{m_1+1}$, $\alpha_i \neq 0$, $i = 0, .., m_1$

vérifiant :

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ 0 & x_1(P_1) & \ldots & x_1(P_{m_1}) \\ \vdots & \vdots & & \vdots \\ 0 & x_{m_1-1}(P_1) & \ldots & x_{m_1-1}(P_{m_1}) \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{m_1} \end{pmatrix} = \theta_{m_1}$$

ce qui est équivalent à : il existe $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{m_1}) \in \mathbb{R}^{m_1+1}$, $\alpha_i \neq 0$, $i = 0, \ldots, m_1$ tel que :

$$\sum_{i=0}^{m_1} \alpha_i = 0$$

$$\sum_{i=0}^{m_1} \alpha_i \, x_k(P_i) = 0 \qquad \text{pour } k = 1, \ldots, m_1-1$$

Donc :

$$-\alpha_0 = \sum_{i=1}^{m_1} \alpha_1$$

On normalise cette somme en divisant par $-\alpha_0$, en notant $\alpha_i' = -\frac{\alpha_i}{\alpha_0}$ pour $1 = 1, \ldots, m_1$. On obtient

$$\sum_{i=1}^{m_1} \alpha_i' = 1$$

$$\sum_{i=1}^{m_1} \alpha_i' \, x_k(P_i) = 0$$

où par définition du bloc positif,

$$\alpha_i' > 0 \qquad \text{pour } i = 1, \ldots, m_1.$$

Par la définition des fonctions de base, on conclut :

$$\theta \in Co(M_1).$$

De manière analogue on obtient $\theta \in Co(M_2)$. Par conséquent : $\theta \in Co(M_1) \cap Co(M_2)$. De plus comme M_1 et M_2 sont des blocs minimaux, θ est l'unique point appartenant aux enveloppes convexes de M_1 et M_2.

Par le théorème 3.2, en notant $m = \text{card } M = \text{card}(M_1 \cup M_2)$, on conclut immédiatement, il existe $\lambda = (\lambda_1, \ldots, \lambda_m) \in \text{Ker } U_M$, où U_M est de la forme (3.4) tel que $\lambda_i \neq 0$ pour $i = 1, \ldots, m$, ce qui entraîne que M est un ensemble-H minimal relatif à W (3.1). ■

En conclusion, on choisit comme méthode de construction la méthode des blocs. Pour terminer ce paragraphe, on donne quelques exemples de construction d'ensembles-H minimaux relatifs à W (3.1).

Exemples :

Soit W = eng $\{1, x_1, x_2, x_3\}$. On désigne les points de signes positifs par o et les points de signes négatifs par •. On note :

B_i^+ , le bloc B_i formé de points positifs,

B_i^- , le bloc B_i formé de points négatifs.

On considère les blocs minimaux de la fig. 3.3.
Ces blocs sont obtenus d'ensembles-H minimaux relatifs à des sous-espaces de W. En composant ces différents blocs, on obtient les ensembles-H minimaux suivants (fig. 3.4) :

H_1 obtenu par composition de B_1^+ et de B_2^-

H_2 obtenu par composition de B_2^+ et de B_2^-

H_3 obtenu par composition de B_1^+ et de B_3^-

H_4 obtenu par composition de B_2^+ et de B_3^-

H_5 obtenu par composition de B_1^+ et de B_4^-

On remarque que pour trouver certains blocs - par exemple B_4 - il faut déjà connaître l'ensemble-H correspondant H_5. Dans ce qui suit, on fait un autre choix de blocs et on essaie d'obtenir les mêmes ensembles-H minimaux. Plus précisément, on choisit des blocs trouvés à partir des ensembles-H minimaux relatifs à un sous-espace propre de W.

C'est-à-dire : de H_1 : on trouve B_2^+ qui permet de construire H_2 en composant B_2^+ et B_2^-;

de H_1 : on trouve B_5 (fig. 3.5) qui permet par composition avec B_2^+ de construire H_3 (fig. 3.4);

de H_3 : on trouve B_3^- qui avec B_2^+ permet de construire H_4 (fig. 3.4). De plus B_3^- et B_5 permettent de construire H_5.

On a donc construit H_5 sans faire intervenir B_4^-, ce qui est intéressant lorsque l'on désire construire les ensembles-H minimaux relatifs à des sous-espaces linéai-

res W de dimensions de plus en plus grandes en se basant sur les ensembles-H minimaux déjà connus, c'est-à-dire, les ensembles-H minimaux relatifs aux sous-espaces linéaires V contenus strictement dans W.

Il est amusant de remarquer que l'on peut retrouver tous les ensembles-H minimaux relatifs à W = eng $\{1, x_1, x_2, x_3\}$ à partir de l'un d'entre eux et d'une droite. En effet, on considère un ensemble-H minimal M = $\{P_1, P_2, P_3, P_4, P_5\}$ relatif à W (fig. 3.6). Et soit une droite D_5 passant par P, on suppose que P_1, P_2, P_3, P_4 sont 4 points non coplanaires. Faisant glisser P_5 sur la droite D, on dessine aux fig. 3.6 les différents ensembles-H minimaux relatifs à W que l'on obtient.

Il est intéressant de remarquer qu'un point P_i change de signe lorsque le cardinal de l'ensemble-H minimal diminue d'une unité, ce qui se traduit par un point P_i qui change de signe lorsque λ_i s'annule où $\lambda = (\lambda_1, \ldots, \lambda_5) \in \text{Ker } U_M$. Ceci se produit lorsque l'on passe de (fig. 3.6.1.) à (fig. 3.6.3), de (fig. 3.6.3) à (fig. 3.6.5) et de (fig. 3.6.5) à (fig. 3.6.7).

3.3. Nombre des ensembles-H dans $\mathbf{R}^n$.

Dans l'article [15], TAYLOR énonce une règle qui donne le nombre exact d'ensembles-H minimaux relatifs à W défini en (3.1) et pour un nombre n quelconque. Pour pouvoir démontrer cette règle, on a besoin du théorème suivant :

Théorème 3.4 :

S'il existe J *ensembles-H minimaux dans* $\mathbf{R}^{n-1}$, *alors il existe* $J + [\frac{n}{2}] + 1$ *ensembles-H minimaux dans* $\mathbf{R}^n$ *où* [a] *désigne la valeur entière de* a, *c'est-à-dire, le plus grand entier inférieur ou égal à* a.

Démonstration :

1. La propriété 3.1 assure que les ensembles-H minimaux de $\mathbf{R}^{n-1}$ sont des ensembles-H minimaux de $\mathbf{R}^n$. Par hypothèse, on en a J.
2. Par (1.6), les ensembles-H M minimaux de $\mathbf{R}^n$ qui ne sont pas ensembles-H minimaux de $\mathbf{R}^{n-1}$ vérifient :

$$\text{card } M = n + 2.$$

On décompose M en deux ensembles M^+ et M^- de telle manière que :

$$M = M^+ \cup M^-$$
$$\text{card } M^+ = m_1 \quad \text{et} \quad \text{card } M^- = m_2$$

on obtient : $$\text{card } M = m_1 + m_2 = n + 2 \qquad (3.5)$$

On peut supposer : $m_1 \geqslant m_2$, car les ensembles-H minimaux n'ont que deux configurations possibles de signes. Si $m_2 > m_1$, il suffit de prendre la configuration des signes opposés pour obtenir $m_2 \leqslant m_1$. On établit donc les correspondances suivantes :

si $m_2 = 1$ alors $m_1 = n + 1$
si $m_2 = 2$ alors $m_1 = n$
⋮ ⋮

Si n est paire, $n = 2k$, comme $m_1 \geqslant m_2$, la plus grande valeur possible de m_2 est $m_1 = \frac{n}{2} + 1$. Par conséquent, la plus petite valeur de m_1 est $m_1 = \frac{n}{2} + 1$.
De même, si n est impaire, $n = 2k + 1$, on obtient :
la plus grande valeur possible de m_2 est $m_2 = k + 1$ et la plus petite valeur possible de m_1 est $m_1 = k + 2$, c'est-à-dire :

la valeur maximale de m_2 est : $m_2 = [\frac{n}{2}] + 1$

la valeur minimale de m_1 est : $m_1 = [\frac{n}{2}] + 2$

Par conséquent pour m_2 et pour m_1, on obtient exactement $[\frac{n}{2}] + 1$ valeurs différentes. En d'autres mots, on a exactement $[\frac{n}{2}] + 1$ ensembles-H minimaux de $\mathbf{R}^n$ vérifiant (3.5).

3. En comptant les ensembles-H minimaux obtenus aux étapes 1 et 2, on conclut la thèse.

On peut maintenant énoncer la règle de TAYLOR.

Théorème 3.5 :

Dans $\mathbf{R}^n$, *il y a exactement* $h(n)$ *ensembles-H minimaux, où*
$h(n) = k^2 + 2k$ *si* $n = 2k$
$= k^2 + 3k + 1$ *si* $n = 2k+1$

Démonstration :

Par le théorème précédent, on obtient immédiatement :

$$h(n) = \sum_{j=1}^{n} ([\frac{j}{2}] + 1)$$

si n = 2k, cette somme devient :

$$h(n) = \sum_{j=1}^{2k} [\frac{J}{2}] + 1 = 1 + ([\frac{2}{2}] + 1) + ([\frac{3}{2}] + 1) + \ldots + (k+1)$$

$$= 1 + 2(2) + 2(3) + \ldots + 2(k) + (k+1)$$

$$= 2(1+2+\ldots+k) + k$$

$$= k(k+1) + k$$

$$= k^2 + 2\,k.$$

Si n = 2k+1, la somme devient :

$$h(n) = \sum_{j=1}^{2k+1} ([\frac{J}{2}] + 1) = 1 + 2(2) + 2(3) + \ldots + 2(k) + 2(k+1)$$

$$= 1 + 2(1+2+\ldots+k) + 2k$$

$$= 1 + k(k+1) + 2k$$

$$= k^2 + 3k + 1$$

Exemple :

Soit w = eng {1, x, y}. Dans $\mathbf{R}^2$, il existe 3 ensembles-H minimaux H_1, H_2, H_3 représentés à la fig. 3.4, on remarque qu'il existe un ensemble-H dont le cardinal est égal à 3 et qu'il existe deux ensembles-H dont le cardinal est égal à 4. En effet pour **R**, la règle assure l'existence d'un ensemble-H M tel que card M = 3 et par le théorème 3.4 appliqué à $\mathbf{R}^2$, on conclut qu'il existe $[\frac{2}{2}] + 1 = 2$ ensembles-H de R^2 satisfaisant à card M = 4.

La méthode des blocs nous permet de construire des ensembles-H minimaux, mais il est intéressant de pouvoir construire des ensembles-H sans répétition. Le but du paragraphe suivant est de proposer une telle méthode.

3.4. Construction systématique des ensembles-H minimaux.

A partir des résultats obtenus aux paragraphes précédents, on essaye de guider la construction des ensembles-H minimaux de $\mathbf{R}^n$ de telle manière qu'il n'y ait pas de répétition. On emploie toujours la méthode des blocs et on ne considère que les blocs minimaux obtenus des ensembles-H minimaux relatifs aux sous-

espaces propres de $\mathbf{R}^n$. La règle de TAYLOR donne le nombre exact d'ensembles-H minimaux de $\mathbf{R}^n$. De plus, cette règle a été établie en distinguant les ensembles-H $M = M^+ \cup M^-$ suivant le cardinal de M^+. Comme la représentation graphique des ensembles-H est encore possible dans $\mathbf{R}^3$, on amorce le raisonnement dans cet espace.

3.4.1. Construction systématique dans $\mathbf{R}^3$.

Les ensembles-H minimaux M de $\mathbf{R}^3$ vérifiant : card $M < n+2 = 5$ sont des ensembles-H minimaux de $\mathbf{R}^2$. Ils sont représentés à la fig. 3.4 : H_1, H_2, H_3.

Par la règle de TAYLOR, on sait qu'il existe : $[\frac{n}{2}] + 1 = 2$ ensembles-H minimaux notés H_4 et H_5, de cardinal égal à $n + 2 = 5$. Ils proviennent de la composition d'un bloc de 3 points et d'un autre de 2 points. Pour construire des ensembles-H minimaux, on tient le même raisonnement que TAYLOR, c'est-à-dire, qu'on distingue les ensembles-H suivant le nombre de points positifs qu'ils possèdent. Il en résulte :

1. card $H_5^+ = 1$.
   Pour obtenir H_5 à partir des blocs obtenus de H_1, H_2, H_3, il n'y a qu'une seule composition possible; celle de B_1 avec B_2 où B_1 et B_2 sont données par la fig. 3.7.
2. card $H_4^+ = [\frac{n}{2}] + 1 = 2$.
   Pour obtenir H_4 à partir des blocs obtenus de H_1, H_2, H_3, une composition possible est celle de B_3 avec B_4 où B_3 et B_4 sont données à la fig. 3.7.
   On obtient H_4 (fig. 3.4)

On remarque que dans les deux cas, on compare un bloc obtenu de l'ensemble-H de $\mathbf{R}^2$ qui ne contient qu'un seul point positif avec un bloc obtenu de H_1.

3.4.2. Construction systématique dans $\mathbf{R}^4$.

Par la règle de TAYLOR, on sait qu'il existe $[\frac{n}{2}] + 1 = 3$ ensembles-H minimaux de cardinal égal à 6. On les note : H_6, H_7, H_8. On tient le même raisonnement que précédemment :

1. card $H_6^+ = 1$.
   Pour construire H_6 à partir des blocs obtenus des ensembles-H déjà connus, il

n'y a qu'une seule composition possible, celle de B_1 avec B_2 où B_1 et B_2 sont donnés à la fig. 3.8.
On obtient H_6.

2. card $H_7^+ = 2$.
Pour construire H_7, on peut composer les blocs B_3 et B_4 (fig. 3.8) ; on obtient H_7.

3. card $H_8^+ = 3$.
On compose les blocs B_5 et B_6 (fig. 3.8) ; on obtient H_8.

> Il est à souligner que l'on ne compose que des blocs obtenus des ensembles-H minimaux ne contenant qu'un seul point positif ou qu'un seul point négatif.

3.4.3. Construction systématique dans R^n.

De manière générale, on note :

B_n^+ : le bloc obtenu de l'ensemble-H minimal de $\mathbf{R}^n$ qui ne contient qu'un seul point positif, en omettant ce point; de plus : card $B_n^+ = n+1$.

$(B_n')^+$: est construit de la même façon mais en omettant un autre point que l'unique point positif.

B_n^- et $(B_n')^-$ sont des blocs de signes opposés aux blocs B_n^+ et $(B_n')^+$.

Par la règle de TAYLOR, on sait qu'il existe $[\frac{n}{2}] + 1$ ensembles-H minimaux M de cardinal n + 2 :

1. card $M^+ = 1$
   on compose : B_1^+ et $(B'_{n-1})^+$
2. card $M^+ = 2$
   on compose : B_1^- et B_{n-1}^+
3. card $M^+ = 3$
   on compose : B_2^- et B_{n-2}^+

et ainsi de suite.

Si n est paire : $n = 2k$, le dernier ensemble-H que l'on peut construire est tel que : $\text{card } M^+ = k + 1$

on compose : B_k^- et B_k^+

Si n est impaire : $n = 2k + 1$, le dernier ensemble-H que l'on peut construire est tel que : $\text{card } M^+ = k + 1$

on compose : B_k^- et B_{k+1}^+

Ainsi, on est parvenu à guider la construction des ensembles-H minimaux de telle manière qu'il n'y ait aucune répétition. On remarque facilement que pour cette construction, il suffit de retenir les ensembles-H minimaux M tels que :

$$\text{card } M^+ = 1.$$

3.5. Classification de BROSOWSKY.

Dans [1], BROSOWSKY propose une classification des ensembles-H minimaux relatifs à W (3.1), semblable à celle de TAYLOR. Pour ce, il introduit une nouvelle notion, celle de classe dont voici la définition.

Définition : On dit que l'ensemble-H $M = M^+ \cup M^-$ est de classe (k, j) si et seulement si : 1. $\text{card } M = k + 2$,

2. $j = \min \{\text{card } M^+, \text{card } M^-\}$

De plus, si W est défini par (3.1), on obtient immédiatement :

1. $1 \leqslant k \leqslant n$
2. $1 \leqslant j \leqslant [\frac{k}{2}] + 1$

On montre dans le théorème suivant que les classes (k, j) sont non vides.

Théorème 3.6 :

Pour tout k, $1 \leqslant k \leqslant n$ *et pour tout* $j = 1, 2, \ldots, [\frac{k}{2}] + 1$, *il existe un ensemble-H minimal de classe* (k, j).

Démonstration :

Par le théorème 3.4, on sait que pour tout $k \in \mathbb{N}$, il existe $[\frac{k}{2}] + 1$ ensembles-H minimaux de cardinal égal à $k + 2$. De plus, ces ensembles-H $M = M^+ \cup M^-$ sont construits de sorte que : min (card M^+, card M^-) prend les valeurs successives $1, 2, 3, \ldots, [\frac{k}{2}] + 1$. Donc, il existe un ensemble-H minimal de classe (k, j) pour $k = 1, \ldots, n$ et $j = 1, 2, \ldots, [\frac{k}{2}] + 1$. ■

Des exemples sont donnés :

1. H_1 (fig. 3.4) est de classe (1.1)

2. H_3 (fig. 3.4) est de classe (2.1)

3. H_2 (fig. 3.4) est de classe (2.2)

4. H_4 (fig. 3.4) est de classe (3.2)

Remarque :

En utilisant cette classification, la méthode de construction des ensembles-H minimaux proposée au paragraphe 4 s'exprime par : on compose les blocs B_i obtenus des ensembles-H de classe (i, 1) en omettant le point qui appartient au sous-ensemble de cardinal égale 1, c'est-à-dire, on compose les blocs B_i et B_{n-i} pour $i = 1, \ldots, k$.

CHAPITRE 4 :

Ensembles-H minimaux relatifs à une approximation par des polynômes à plusieurs variables de degré supérieur ou égal à deux.

4.1. Préliminaires.

Soient D, un ensemble de m points : $\{P_1, P_2, \ldots, P_m\}$ et la famille génératrice de W_o : $\{g_o \cdot g_i \mid i = 1, \ldots, p\}$ (4.1)

On suppose que $g_o \in C[B, \mathbf{R}]$ et $g_o(P_i) \neq 0$, $i = 1, \ldots, m$.

Proposition 4.1 :

Si D avec la configuration de signes $(\varepsilon_i \mid i = 1, \ldots, m)$ *est un ensemble-H (minimal) relatif à W (1.1), alors D muni de la configuration de signes*

$$\left(\varepsilon_i \frac{g_o(P_i)}{|g_o(P_i)|} \mid i = 1, \ldots, m\right)$$

est un ensemble-H (minimal) relatif à W_o *(4.1).*

Démonstration :

Par hypothèse, il existe m scalaires positifs, non tous nuls ρ_i, $i = 1, \ldots, m$ qui vérifient

$$\sum_{i=1}^{m} \rho_i \varepsilon_i g_j(P_i) = 0 \qquad j = 1, \ldots, p$$

ou encore

$$\sum_{i=1}^{m} \frac{\rho_i}{|g_o(P_i)|} \frac{\varepsilon_i |g_o(P_i)|}{g_o(P_i)} g_o(P_i) \cdot g_j(P_i) = 0 \qquad j = 1, \ldots, p \quad \blacksquare$$

Exemple :

Soit $D = \{P_1, P_2, P_3, P_4\}$, l'ensemble-H minimal relatif à V_2^1, représenté à la fig. 4.1, les ensembles D' et D" des fig. 4.2 obtenus de D par des transformations invariantes sont des ensembles-H minimaux relatifs à

$$\text{eng}\{x^a y^b, x^{a+1} y^b, x^a y^{b+1}\}$$

où a est un entier impair et b un entier quelconque.

Rappel 4.1 :

On retrouve dans l'étude de l'approximation dans $\mathbf{R}$, le résultat suivant : un ensemble-H minimal relatif à V_1^n (2.11) est constitué de n+2 points P_i, $i = 1, \ldots, n+2$ sur l'axe réel. Ces points vérifient la propriété d'alternance des signes de l'er-

reur. Autrement dit, les P_i, i = 1, ..., n+2 étant ordonnés sur l'axe, $\varepsilon(P_{i+1}) \cdot \varepsilon(P_i) < 0$ où $\varepsilon(x)$ désigne l'erreur commise en x (cfr. théorème 1.1).

La fig. 4.3 montre un ensemble-H minimal relatif à V_1^4.

4.2. Construction d'ensembles-H à l'aide des ensembles-B.

4.2.1. Méthode.

Soit $W = \text{eng } \{h_1, \ldots, h_{n_1}, g_o \cdot g_1, \ldots, g_o \cdot g_{n_2}\}$ (4.2)

Proposition 4.2 :

Si on suppose que $D_2 = \{P_{m_1+1}, \ldots, P_{m_1+m_2}\}$ *est un ensemble-H minimal relatif à* $W_2 = \text{eng } \{g_o \cdot g_1, \ldots, g_o \cdot g_{n_2}\}$ *tel que* $g_o(P_j) \neq 0$, $j = m_1+1, \ldots, m_1+m_2$, *si de plus,* $D_1 = \{P_1, \ldots, P_m\}$ *est obtenu d'un ensemble-H minimal relatif à* $W_1 = \text{eng } \{h_1, \ldots, h_{n_1}\}$ *en omettant un point et est tel que* $g_o(P_i) = 0$, $i = 1, \ldots, m_1$, *alors* D_1 *est un ensemble-B relatif à* $\{h_1, \ldots, h_{n_1}\}$ *et dépendant de* D_2 *et* $D = D_1 \cup D_2$ *est soit un ensemble-H minimal relatif à W (4.2), soit un ensemble-H relatif à un sous-espace propre de W.*

Démonstration :

1. D_1 contient le nombre de points nécessaire pour que card $D_1 \cup D_2 = n_1 + n_2 + 1$ et les signes des points de D_1 dépendent globalement de ceux des points de D_2.
2. La matrice $M(U_D)$ a la structure (2.4). Par hypothèse, rg $U_{22} = m_2 - 1 \leqslant n_2$ et rg $U_{11} = m_1 = n_1$

   Par conséquent, rg $U_D \leqslant m-1 \leqslant n$. ∎

Critère de minimalité :

D ainsi construit est un ensemble-H minimal relatif à W (4.2) ssi toute partie propre de D : $D' \subset D$ vérifie rg $U_{D'}$ = card D'.

Exemple :

Dans l'exemple suivant, ce critère n'est pas vérifié. L'ensemble-H : D représenté à la fig. 4.4, construit à partir de l'ensemble-B {(0, 0), (1, 0), (2, 0)} relatif à $\{1, x, x^2\}$ et de l'ensemble-H minimal {(0, 1), (0, 2), (1, 1), (1, 3/2)} relatif

à eng $\{y, xy, y^2\}$ n'est pas minimal. Par contre, l'ensemble D' = $\{(0, 0), (1, 0), (0, 1), (0, 2), (1, 1), (1, 3/2)\}$ est un ensemble-H minimal relatif à

$$\text{eng } \{1, x, y, y^2, xy\}$$

En effet $M(U_D) =$

$$\left[\begin{array}{ccc|cccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 0 & 0 & 1 & 1 \\ 0 & 1 & 4 & 0 & 0 & 1 & 1 \\ \hline 0 & 0 & 0 & 1 & 2 & 1 & 3/2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3/2 \\ 0 & 0 & 0 & 1 & 4 & 1 & 9/4 \end{array}\right] \begin{array}{l} \leftarrow \phi_1 = 1 \\ \leftarrow \phi_2 = x \\ \leftarrow \phi_3 = x^2 \\ \\ \leftarrow \phi_4 = y \\ \leftarrow \phi_5 = xy \\ \leftarrow \phi_6 = y^2 \end{array}$$

On voit de suite que $M(U_{D'})$ obtenue de $M(U_D)$ en enlevant la 3ème colonne a 2 lignes égales et donc rg $U_{D'}$ < card D'.

Détermination des signes de D_2 :

Les signes des points de l'ensemble-B, D_2 peuvent être déterminés dans certains cas particuliers comprenant entre autres l'approximation polynomiale.

La méthode permettant de déterminer ces signes est décrite dans les sections suivantes.

4.2.2. Construction par récurrence dans R^2.

1. Récurrence.

V_2^{n+1} étant bien de la forme (4.2), on peut écrire :

$$V_2^{n+1} = \text{eng } \{1, x, \ldots, x^{n+1}\} \oplus \text{eng } \{y \cdot x^i y^j : i+j \leqslant n \quad i, j \geqslant 0\}$$

et constater que des ensembles-H minimaux relatifs à V_2^{n+1} : D peuvent être construits par la méthode (cfr. 4.2.1) précédente à partir des ensembles-H minimaux relatifs à V_2^n.

On ajoute,à ces derniers,n+2 points colinéaires formant un ensemble-B.

2. Détermination des signes de D.

On peut déterminer les signes de ces n+2 points si D contient au moins trois sous-ensembles de n+2 points colinéaires. Sur trois de ces sous-ensembles, on fixe tour à tour l'axe X par des transformations invariantes. On garde toujours la décomposition en un ensemble-B : D_1 (sur l'axe) et un ensemble-H minimal : D_2, mais, avec

certains signes déterminés (invariance), pour D_1 et d'autres indéterminés pour D_2. Ces derniers peuvent être facilement trouvés puisque l'on connaît la configuration des signes des ensembles-H minimaux relatifs à V_2^n (fig. 4.5, 4.6 et 4.7).

Ayant agi de la sorte pour chacune des trois positions de l'axe, on aura déterminé la configuration des signes de D.

3. Exemple.

Les ensembles-H de départ de la récurrence ainsi définie sont ceux relatifs à V_2^1 (fig. 1.1). On ajoute à ces ensembles, trois points P_5, P_6, P_7 sur l'axe X de façon à former au moins trois sous-ensembles de trois points colinéaires $S_1 = \{P_5, P_6, P_7\}$, $S_2 = \{P_1, P_2, P_5\}$ et $S_3 = \{P_1, P_3, P_7\}$.

L'axe X fixé sur S_1, on détermine les signes de P_1, P_2, P_3, P_4 (fig. 4.5).

Par une transformation affine appropriée, on amène les points de S_2 sur l'axe X, et on détermine les signes de P_6 et P_7 de façon à ce que $\{P_3, P_4, P_6, P_7\}$ soit un ensemble-H minimal (fig. 4.6).
On agit de même pour déterminer le dernier signe (fig. 4.7).

Plusieurs autres configurations sont données également pour n = 2 (fig. 4.8) et pour n = 3 (fig. 4.9).

4.2.3. Construction par récurrence dans R^3.

1. Récurrence.

V_3^{n+1} peut s'écrire :

$$\text{eng}\{1, x, x^2, \ldots, x^{n+1}\} \oplus \text{eng}\{y \cdot x^i y^j : i+j \leqslant n,\ i, j \geqslant 0\} \oplus$$
$$\text{eng}\{z \cdot x^i y^j z^k : i+j+k \leqslant n,\ i, j, k \geqslant 0\}$$

Utilisant la méthode décrite en 4.2.1 et adaptée à R^3, on se base sur les ensembles-H minimaux relatifs à V_3^n, à V_2^n et à V_1^{n+1} pour construire des ensembles-H minimaux relatifs à V_3^{n+1} : D.
Les ensembles-B seront des ensembles de $\frac{(n+2)(n+3)}{2}$ points coplanaires.

2. Détermination des signes associés à D.

Si D contient au moins quatre sous-ensembles de $\frac{(n+2)(n+3)}{2}$ points coplanaires, on

peut déterminer la configuration qui lui est associée. Il suffit pour cela, d'amener par une transformation invariante, le plan XY sur quatre de ces sous-ensembles et de déterminer les signes de façon à ce que les points hors de ce plan forment un ensemble-H relatif à V_3^n.

3. Exemple :

Pour n = 2, l'ensemble représenté à la fig. 4.10 est un ensemble-H minimal.

4.3. Construction d'ensembles-H à l'aide des blocs.

4.3.1. Méthode.

Proposition 4.3 :

Soit l'ensemble D_1, *composé de* $n_1 + 1$ *points, un bloc négatif minimal relatif à* $\{h_1, h_2, \ldots, h_{n_1}\}$ *et* D_2, *un bloc positif minimal relatif à* $\{g_1, g_2, \ldots, g_{n_2}\}$.
Si de plus, aucun h_i, $i = 1, \ldots, n_1$, *aucun* g_j, $j = 1, \ldots, n_2$ *ne sont constants et si* $h_i(P) = 0$ *sur* D_2 *et* $i = 1, \ldots, n_1$ *et* $g_j(P) = 0$ *pour tout* $P \in D_1$, $j = 1, \ldots, n_2$,
alors l'ensemble $D = D_1 \cup D_2$ *est un ensemble-H relatif à* $W = \text{eng}\{1, h_1, h_{n_1}, g_1, \ldots, g_{n_2}\}$.

Démonstration :

La matrice $M(U_D)$ ayant $n_2 + n_1 + 2$ colonnes et $n_2 + n_1 + 1$ lignes, on a

$$\text{rg } U_D \leq n_2 + n_1 + 1 = \text{card } D - 1. \blacksquare$$

D sera un ensemble-H minimal relatif à W ssi le critère de minimalité est vérifié.

4.3.2. Construction par récurrence.

La récurrence que l'on peut déduire de la méthode des blocs n'est pas du même type que celles définies par la méthode des ensembles-B. Celles-ci étaient des récurrences sur le degré des polynomes approximants; ici la récurrence se fait sur le nombre de variables des polynomes.

1. Récurrence.

La construction des ensembles-H relatifs à l'approximation par des polynomes de degré k dans R^{n+1} se base sur les ensembles-H relatifs à l'approximation par des polynomes de degré k dans R^i : i = 1, ..., n.
En effet, $V^k_{n+1} = \text{eng}\{x_1^i\, x_2^j \ \ldots\ x_{n+1}^l : i+j+\ldots+l \leqslant k \quad i, j, \ldots, l \geqslant 0\}$ peut être obtenu par des transformations affines à partir de l'espace engendré par les ensembles $\{1, x_1, x_1^2, \ldots, x_1^k\}$ et $\{x_2^i\, x_3^j \ \ldots\ x_{n+1}^l : i+j+\ldots+l \leqslant k$ $i, j, \ldots, l \geqslant 0\}$

On illustre ceci dans R^2 : V_2^2 peut être obtenu à partir des ensembles $\{1, x, x^2\}$ et $\{1, y, y^2\}$ par une transformation linéaire.

$$P = (\xi, \eta) \in R^2$$

$$[x^2] \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = (a\xi + b\eta)^2$$

$$= a^2 [x^2] \begin{pmatrix} \xi \\ \eta \end{pmatrix} + 2ab [xy] \begin{pmatrix} \xi \\ \eta \end{pmatrix} + b^2 [y^2] \begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

On engendre donc par cette transformation, la fonction de base xy.

Connaissant donc les blocs minimaux relatifs à $\{x_1, x_1^2, \ldots, x_1^k\}$ obtenus directement des ensembles-H minimaux relatifs à V_1^k et les blocs minimaux relatifs à $\{x_2^i\, x_3^j \ \ldots\ x_{n+1}^l : 0 < i+j+\ldots+l \leqslant k \quad i, j, \ldots, l \geqslant 0\}$ obtenus des ensembles-H minimaux relatifs à V_n^k (le théorème 2.6 est applicable) et en respectant la condition de composition d'un bloc négatif et minimal avec un bloc positif et minimal, on construit donc un ensemble-H (minimal) relatif à V^k_{n+1}.

2. Exemples :

- Un ensemble-H minimal relatif à V_2^4 est représenté à la fig. 4.11b.
  A la fig. 4.11a, on a un ensemble-H, obtenu par composition d'un bloc négatif minimal et d'un bloc positif minimal, relatif à $\text{eng}\{1, x, x^2, x^3, x^4, y, y^2, y^3, y^4\}$. Par une transformation affine, on obtient l'ensemble-H minimal relatif à V_2^4 (fig. 4.11b).
- Un ensemble-H minimal (fig. 4.12b) relatif à V_3^3 est obtenu à partir d'un ensemble-H minimal relatif à V_2^3 (fig. 4.12a) et d'un ensemble-H minimal relatif à V_1^3.

Il faut noter que l'on peut aussi composer des blocs minimaux obtenus à partir d'ensembles-H minimaux relatifs à V_i^k et des blocs minimaux obtenus à partir d'ensembles-H relatifs à V_j^k en respectant les conditions de composition (1 bloc positif minimal composé avec 1 bloc négatif minimal) pour obtenir des ensembles-H minimaux relatifs à V_{i+j}^k.

4.4. Premières conclusions.

Ayant en notre possession les méthodes permettant de construire tous les ensembles-H relatifs à une approximation de degré 1 dans $\mathbf{R}^n$ avec n quelconque et connaissant la façon de passer du degré k au degré k+1 pour une approximation polynomiale dans $\mathbf{R}$, $\mathbf{R}^2$, $\mathbf{R}^3$ et de passer d'une approximation dans $\mathbf{R}^n$ à une approximation dans $\mathbf{R}^{n+1}$, on peut construire des ensembles-H relatifs à une approximation par des polynomes de degré quelconque et dans n'importe quel espace $\mathbf{R}^n$.
Cependant tous les ensembles-H ne peuvent être obtenus uniquement par ces méthodes.

4.5. Ensembles-H minimaux relatifs à V = eng. {1, x, y, xy}.

4.5.1. Groupes d'invariance pour V.

Certaines restrictions sont à apporter aux groupes décrits (en 2.1.3), afin de respecter le critère d'invariance (théorème 2.1).

Le groupe des translations et des dilatations H(2, $\mathbf{R}$) est un groupe d'invariance pour V.

Le groupe des transformations linéaires de GL(2, $\mathbf{R}$) qui ont pour matrice, une matrice diagonale :

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} \quad \text{avec } a.b \neq 0 \tag{4.4}$$

est aussi un groupe d'invariance.
Font partie de ce groupe :
- les rotations d'angle $\frac{k\pi}{2}$, $k \in \mathbb{Z}$
- les symétries par rapport aux axes et bissectrices,
- les dilatations de grandeur différente sur chaque axe.

Laissent aussi V invariant, les transformations affines de GA(2, R) ayant une matrice du type (4.4) et les transformations projectives qui ont pour matrice

$$\begin{pmatrix} a & 0 & c \\ 0 & b & d \\ 0 & 0 & f \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} 0 & a & c \\ b & 0 & d \\ 0 & 0 & f \end{pmatrix} \quad \text{avec } a.b \neq 0, \; f \neq 0$$

Ces restrictions étant faites, on peut observer les effets qu'elles ont sur la détermination de tous les ensembles-H minimaux relatifs à V.

c1. Le parallélisme est conservé après toute transformation d'un de ces quatre groupes. Ceci est dû au fait que les transformations projectives, avec les restrictions faites ci-avant, ne permettent plus d'envoyer un point à l'infini sur un point fini et vice-versa.

c2. Aucune droite passant par l'origine et non parallèle aux axes ne peut être ramenée sur un des axes, car les rotations d'angle quelconque ne laissent pas les ensembles-H invariants.

c3. Il est toutefois possible de ramener tout ensemble de points dans le premier quadrant par translations ou par des symétries appropriées.

c4. Sous certaines conditions, on peut rendre deux droites perpendiculaires. En effet, soient $\vec{u}$ et $\vec{v}$ les vecteurs directeurs unitaires des deux droites D_1 et D_2 faisant entre elles un angle α.

$$\cos \alpha = u_1 v_1 + u_2 v_2$$

Il nous faut rendre ce cosinus nul. On applique une transformation linéaire $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ à D_1 et D_2. On obtient

$$\cos \beta = \frac{a^2 u_1 v_1 + b^2 u_2 v_2}{\sqrt{a^2 u_1^2 + b^2 u_2^2}\ \sqrt{a^2 v_1^2 + b^2 v_2^2}}$$

La condition de perpendicularité est : $\cos \beta = 0$, autrement dit $a^2 u_1 v_1 = -b^2 u_2 v_2$. On ne peut donc trouver a et b et donc une transformation linéaire que si

$$u_1 v_1 \cdot u_2 v_2 < 0 \qquad (4.5)$$

c5. Les transformations linéaires, affines et projectives décrites en 4.5.1 permettent d'envoyer deux points différents sur deux points, sans oublier toutefois

les restrictions des conséquences c1 et c2. En effet :

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} ax_1 \\ by_1 \end{pmatrix} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

Connaissant (x_1, y_1) et (x_2, y_2), on peut déterminer a et b tels que $a.b \neq 0$ si x_1, y_1, x_2 et y_2 sont différents de 0. L'origine (2ème point) est envoyée sur elle-même.

$$\begin{pmatrix} a & 0 & c \\ 0 & b & d \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = \begin{pmatrix} ax_1 + c \\ by_1 + d \\ 1 \end{pmatrix} = \begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix}$$

et

$$\begin{pmatrix} a & 0 & c \\ 0 & b & d \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_3 \\ y_3 \\ 1 \end{pmatrix} = \begin{pmatrix} ax_3 + c \\ by_3 + d \\ 1 \end{pmatrix} = \begin{pmatrix} x_4 \\ y_4 \\ 1 \end{pmatrix}$$

Il est nécessaire de connaître l'image de deux points pour déterminer la transformation projective ou la transformation affine.

4.5.2. Ensembles-H minimaux relatifs à V.

Un ensemble-H minimal relatif à V est un ensemble de 5 points en ne considérant pas les ensembles-H minimaux relatifs aux sous-espaces de V.

1. Etude des dégénérescences.

Soient les 5 points $P_i = (x_i, y_i)$, $i = 1, \ldots, 5$ $P_i \neq P_j$ $i \neq j$; la matrice associée à U_D s'écrit :

$$M(U_D) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 & x_5 \\ y_1 & y_2 & y_3 & y_4 & y_5 \\ x_1 y_1 & x_2 y_2 & x_3 y_3 & x_4 y_4 & x_5 y_5 \end{pmatrix}$$

L'étude du rang de cette matrice va permettre de déterminer tous les ensembles-H relatifs à V mais dégénérés. Autrement dit, les ensembles-H relatifs à des sous-espaces propres de V.

1. Si $x_i = a$ ou $y_i = b$, $i = 1, \ldots, 5$, le rang de U_D est égal à 2.

On conclut :

Trois points de signe alterné sur une droite parallèle à un des axes forment un ensemble-H dégénéré relatif à V.

2. Si $a + bx_i + cy_i = 0$, $b + c \neq 0$, $i = 1, \ldots, 5$, le rang de U_D est égal à 3.

Quatre points de signe alterné sur une droite non parallèle aux axes forment un ensemble-H dégénéré relatif à V.

3. Si $x_i y_i = a$, $a \neq 0$, rg $U_D = 3$ et

Quatre points de signe alterné sur la courbe d'équation $y = \frac{a}{x}$ forment un ensemble-H dégénéré relatif à V.

4. Si $x_i y_i = ax_i + by_i + c$, $a, b, c \notin \mathbf{R}$ rg $U_D = 3$ et $a + b \neq 0$

Quatre points de signe alterné sur la courbe d'équation $xy = ax + by + c$ forment un ensemble-H dégénéré relatif à V.

Toutes les dégénérescences ont ainsi été étudiées.

2. Construction par le produit cartésien.

Dans les notations du chapitre 2 (2.2.3), on a

- W_1 = eng {1, x} et W_2 = eng {1, y},
- D_1 un ensemble-H relatif à W_1 et D_2 un ensemble-H relatif à W_2.

Appliquant la méthode décrite en 2.2.3, afin d'obtenir des ensembles-H de cinq points, on arrive aux configurations suivantes (fig. 4.13).
Tous les ensembles-H obtenus par la méthode du produit cartésien seront d'un de ces quatre types.

Fixant les quatre points du quadrilatère de chacune de ces configurations aux points de coordonnées (0, 0) (0, 1) (1, 0) et (1, 1) et donnant comme coordonnées au cinquième point : (a, b) avec $a > 0$, $b > 0$, $a \neq 1$, $b \neq 1$, $a \geqslant b$, on obtient les configurations des fig. 4.14.

A partir de ces 3 configurations, on peut trouver tous les ensembles-H engendrés par la méthode du produit cartésien par les transformations des groupes d'invariance (4.5.1).

3. Construction par la méthode utilisant le signe du Jacobien.

Pour avoir un ensemble-H comprenant 5 points distincts, il faut que le système d'équation (théorème 2.7) $P_1(\xi) = P_2(\xi) = 0$ ait exactement $m_1 \cdot m_2 = 5$ racines distinctes, où m_1 et m_2 sont respectivement les degrés de P_1 et P_2. Il s'ensuit que si P_1 est un polynome de degré 5, alors P_2 est un polynome de degré 1. L'ensemble-H engendré (fig. 4.15) n'est donc pas minimal puisque composé de 5 points collinéaires (4.5.2 cas 1 et 2).

4. Construction à l'aide des ensembles-B et des blocs.

V peut être décomposé en eng $\{1, x\} \oplus$ eng $\{y, xy\}$. Il est donc bien de la forme (4.2) désirée pour appliquer la méthode décrite en (4.2.1).
On arrive à former des ensembles de 5 points mais les signes des points de l'ensemble-B ne peuvent être déterminés vu les restrictions apportées aux groupes d'invariance. En effet, on ne peut, par aucune des transformations permises, transporter l'axe des x sur n'importe quelle droite passant par deux points (voir fig. 4.22).

La méthode des blocs ne peut être appliquée car $M(U_D)$ ne peut avoir la structure (2.9).

La méthode de GEARHART (cfr. 2.2.4) ne donne non plus aucun résultat.

5. Conclusions.

Le nombre très restreint des ensembles-H minimaux obtenus par ces différentes méthodes et les restrictions émises quant aux transformations laissant ces ensembles-H invariants ne permettent pas de déterminer tous les ensembles-H minimaux relatifs à V.
Il en est de même pour beaucoup d'autres sous-espaces V.

Chapitre 5 :
Algorithmes et résultats numériques.

5.1. Détermination d'ensemble-H.

Par la formulation duale (1.6), on sait que déterminer si un ensemble donné est un ensemble-H relatif à un sous-espace linéaire donné revient à trouver une solution non identiquement nulle d'un système d'équations homogènes. La méthode utilisée pour résoudre un tel problème est la diagonalisation de GAUSS. L'algorithme employé se base sur une sous-routine, appelée Gauss, résolvant un système d'équations par diagonalisation. On ne le reproduira pas ici. On retrouvera cette sous-routine dans les algorithmes d'échange décrits ultérieurement.

Il faut cependant signaler que le programme a été conçu pour une approximation polynomiale de degré quelconque, dans un espace de dimension quelconque.

Résultats obtenus pour quelques exemples :

1. L'ensemble des points est celui de la figure (5.1)

```
DETERMINATION D'ENSEMBLE-H POUR UNE APPROXIMATION
POLYNOMIALE DE DEGRE 1 DANS UN ESPACE DE DIMENSION 2

L'ENSEMBLE DES 4 POINTS A TESTER EST:
        P( 1 ) : ( -1.00 , 0.00 )
        P( 2 ) : (  1.00 , 0.00 )
        P( 3 ) : (  0.00,-1.00 )
        P( 4 ) : (  0.00 , 1.00 )

LA MATRICE DEFINISSANT LE SYSTEME D'EQUATIONS EST

[  1.00   1.00   1.00   1.00 ]
[ -1.00   1.00   0.00   0.00 ]
[  0.00   0.00  -1.00   1.00 ]
```

L'ENSEMBLE CONSIDERE EST UN ENSEMBLE-H

SA CONFIGURATION DE SIGNE EST:

(-1 , -1 , 1 , 1)

DE PLUS CET ENSEMBLE-H EST MINIMAL

2. L'ensemble des points est celui de la figure (5.2)

DETERMINATION D'ENSEMBLE-H POUR UNE APPROXIMATION

POLYNOMIALE DE DEGRE 1 DANS UN ESPACE DE DIMENSION 3

L'ENSEMBLE DES 5 POINTS A TESTER EST:

P(1) : (2.00 , 0.00 , 0.00)

P(2) : (1.00 , -1.00 , 0.00)

P(3) : (1.00 , 1.00 , 0.00)

P(4) : (0.00 , 0.00 , -1.00)

P(5) : (0.00 , 0.00 , 1.00)

LA MATRICE DEFINISSANT LE SYSTEME D'EQUATIONS EST

$$\begin{bmatrix} 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \\ 2.00 & 1.00 & 1.00 & 0.00 & 0.00 \\ 0.00 & -1.00 & 1.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & -1.00 & 1.00 \end{bmatrix}$$

L'ENSEMBLE CONSIDERE EST UN ENSEMBLE-H

SA CONFIGURATION DE SIGNE EST:

(1 ,-1 ,-1 , 1 , 1)

DE PLUS CET ENSEMBLE-H EST MINIMAL

3. L'ensemble des points est celui de la figure (5.3)

DETERMINATION D'ENSEMBLE-H POUR UNE APPROXIMATION
POLYNOMIALE DE DEGRE 2 DANS UN ESPACE DE DIMENSION 1

L'ENSEMBLE DES 4 POINTS A TESTER EST:

P(1) : (0.00)

P(2) : (1.00)

P(3) : (2.00)

P(4) : (3.00)

LA MATRICE DEFINISSANT LE SYSTEME D'EQUATIONS EST

$$\begin{bmatrix} 1.00 & 1.00 & 1.00 & 1.00 \\ 0.00 & 1.00 & 2.00 & 3.00 \\ 0.00 & 1.00 & 4.00 & 9.00 \end{bmatrix}$$

L'ENSEMBLE CONSIDERE EST UN ENSEMBLE-H
SA CONFIGURATION DE SIGNE EST:

(-1 , 1 ,-1 , 1)

DE PLUS CET ENSEMBLE-H EST MINIMAL

On remarque que la configuration de signes trouvée, associée à ce dernier ensemble est en accord avec le théorème de La Vallée Poussin [2] qui affirme que l'erreur de l'approximation polynomiale dans R atteint sa valeur maximale avec des signes alternés.

4. L'ensemble des points est celui de la figure (5.4)

DETERMINATION D'ENSEMBLE-H POUR UNE APPROXIMATION
POLYNOMIALE DE DEGRE 1 DANS UN ESPACE DE DIMENSION 2

L'ENSEMBLE DES 4 POINTS A TESTER EST:

P(1) : (0.00 , 0.00)
P(2) : (1.00 , 0.00)
P(3) : (0.00 , 1.00)
P(4) : (2.00 ,-1.00)

LA MATRICE DEFINISSANT LE SYSTEME D'EQUATIONS EST

$$\begin{bmatrix} 1.00 & 1.00 & 1.00 & 1.00 \\ 0.00 & 1.00 & 0.00 & 2.00 \\ 0.00 & 0.00 & 1.00 & -1.00 \end{bmatrix}$$

L'ENSEMBLE CONSIDERE EST UN ENSEMBLE-H
SA CONFIGURATION DE SIGNE EST:

(1 , -1 , 1 , 1)

DE PLUS CET ENSEMBLE-H EST NON MINIMAL

5. L'ensemble des points est celui de la figure (5.3)

```
DETERMINATION D'ENSEMBLE-H POUR UNE APPROXIMATION
POLYNOMIALE DE DEGRE 1 DANS UN ESPACE DE DIMENSION 1

L'ENSEMBLE DES 4 POINTS A TESTER EST:
        P( 1 ) :  ( 0.00)
        P( 2 ) :  ( 1.00)
        P( 3 ) :  ( 2.00)
        P( 4 ) :  ( 3.00)

LA MATRICE DEFINISSANT LE SYSTEME D'EQUATIONS EST

        [ 1.00  1.00  1.00  1.00 ]
        [ 0.00  1.00  2.00  3.00 ]

LE SYSTEME A RESOUDRE EST UN SYSTEME DE 2 EQUATION A 4 INCONNUES
ON NE PEUT QUE DONNER UNE BASE DE SOLUTIONS
LA BASE DES SOLUTIONS EST:
        ( 1.00,-2.00, 0.00 ,1.00  )
        ( 2.00,-3.00, 1.00 ,0.00  )
```

5.2. Algorithme de REMES.

L'algorithme que l'on se propose de décrire maintenant a été introduit par E. REMES en 1934. J.P. LAURENT dans [10] en a fait une analyse détaillée.

5.2.1. Description de l'algorithme.

1. Calcul des coefficients

On considère le problème d'approximation posé à la section 1.1. Soit M,

un ensemble-H relatif à W : $M = \{P_1, \ldots, P_{n+1}\}$. On peut choisir parmi les vecteurs $\lambda = (\lambda_1, \ldots, \lambda_{n+1}) \in \ker U_M$ celui qui vérifie

$$\begin{aligned} &\text{i:} \quad \sum_{i=1}^{n+1} |\lambda_i| = 1 \\ &\text{ii:} \quad \sum_{i=1}^{n+1} \lambda_i \, f(P_i) > 0 \end{aligned} \tag{5.1}$$

La configuration de signes associée à M est alors

$$(\varepsilon_1, \ldots, \varepsilon_{n+1}) \text{ telle que } \lambda_i = \varepsilon_i \, \rho_i$$

avec $\varepsilon_i = +1$ ou -1 et $\rho_i > 0$, $i = 1, \ldots, n+1$.

Pour cet ensemble M, on considère le problème d'approximation dans $C[M, \mathbf{R}]$. On a la caractérisation suivante :

> g_o est solution du problème d'approximation ssi il existe n+1 scalaires λ_i non nuls, tels que $(\lambda_1, \ldots, \lambda_{n+1}) \in \ker U_M$ et $\sum_{i=1}^{n+1} |\lambda_i| = 1$ qui vérifient $g_o(P_i) = f(P_i) - \varepsilon_i \cdot \max_{j=1,..,n+1} \{|f(P_j) - g_o(P_j)|\}$ où on a posé ε_i = signe de λ_i.

Cette caractérisation découle de la définition d'une solution d'un problème d'approximation et du théorème d'oscillation (1.2).

On remarque que dans ce cas particulier, on a :

$$\sum_{i=1}^{n+1} \lambda_i \, f(P_i) = \max_{j=1,..,n+1} \{|f(P_j) - g_o(P_j)|\}$$

On note cette somme ρ : c'est en fait l'erreur égalisée sur M. On trouve finalement l'expression de g_o en résolvant le système

$$\sum_{i=1}^{n+1} a_i \, g_i(P_j) = f(P_j) - \varepsilon_j \cdot \rho \qquad j = 1, \ldots, n+1 \tag{5.2}$$

La théorie de l'approximation affirme l'existence et l'unicité de la solution.

2. Description de l'algorithme.

On construit une série d'ensembles-H : $M^k = \{P_1^k, \ldots, P_{n+1}^k\}$ et on détermine les scalaires λ_i^k qui vérifient les conditions i et ii. Puis on calcule la

meilleure approximation sur $C[M^k, \mathbf{R}] = g_o^k$.

Si g_o^k est solution du problème primitif, autrement dit si $||f - g_o^k|| = \max_{i=1,..,n+1} \{|f(P_i^k) - g_o(P_i^k)|\}$, l'algorithme s'arrête.

3. Passage de M^k à M^{k+1}.

On suppose donc que g_o^k n'est pas solution du problème d'approximation général dans $C[B, \mathbf{R}]$.

Critère d'entrée : Il existe donc au moins un point $P_o \in B \setminus M^k$ où l'erreur commise est maximale. P_o vérifie donc la relation

$$|f(P_o) - g^k(P_o)| = ||f - g^k|| \qquad (5.3)$$

Il est important de signaler que ce critère est arbitraire. En effet, il n'existe aucun théorème affirmant que ce critère conduit à une convergence plus rapide de l'algorithme.

Critère de sortie : On note les vecteurs caractéristiques

$$\phi_i^k = \begin{pmatrix} \varepsilon_i^k \, g_1(P_i^k) \\ \vdots \\ \varepsilon_i^k \, g_n(P_i^k) \end{pmatrix} \qquad i = 1, \ldots, n+1$$

et

$$\phi_o^k = \begin{pmatrix} \varepsilon_o^k \, g_o(P_o^k) \\ \vdots \\ \varepsilon_o^k \, g_n(P_o^k) \end{pmatrix}$$

où $\varepsilon_o^k = \dfrac{f(P_o^k) - g^k(P_o^k)}{|f(P_o^k) - g^k(P_o^k)|}$.

Par la formulation duale de la définition d'ensemble-H minimal, il existe n+1 scalaires $\rho_i^k > 0$, $i = 1, \ldots, n+1$ tels que

$$\sum_{i=1}^{n+1} \rho_i^k \, \phi_i^k = 0 \quad \text{et} \quad \sum \rho_i^k = 1$$

De plus, les vecteurs $\phi_1^k, \ldots, \phi_{n+1}^k, \phi_o^k$ de $\mathbf{R}^n$ sont linéairement dépendants. Par conséquent, il existe $w_o, \ldots, w_{n+1}$ non tous nuls tels que

$$\sum_{i=1}^{n+1} w_i \phi_i^k + w_o \phi_o^k = 0 \text{ et } w_o > 0$$

Donc, pour tout $\alpha, \beta \in \mathbf{R}$, on a

$$\alpha \sum_{i=1}^{n+1} \rho_i^k \phi_i^k + \beta \left(\sum_{i=1}^{n+1} w_i \phi_i^k + w_o \phi_o^k \right) = 0$$

ou encore

$$\sum_{i=1}^{n+1} (\alpha \rho_i^k + \beta w_i) \phi_i^k + \beta w_o \phi_o^k = 0$$

On pose $\beta = 1$, $\alpha = - \dfrac{w_{i_o}}{\rho_{i_o}^k}$ tel que $\alpha \rho_{i_o}^k + \beta w_{i_o} = 0$.

On obtient $\sum_{i=1}^{n+1} \rho_i^k \left(\dfrac{w_i}{\rho_i^k} - \dfrac{w_{i_o}}{\rho_{i_o}^k} \right) \phi_i^k + w_o \phi_o^k = 0$.

On choisit l'indice i_o tel que tous les coefficients soient positifs. Autrement, on choisit i_o tel que $\dfrac{w_{i_o}}{\rho_{i_o}^k} = \min_{i=1,..,n+1} \dfrac{w_i}{\rho_i^k}$

Construction de M^{k+1}

$M^{k+1} = \{P_1^{k+1}, \ldots, P_{n+1}^{k+1}\}$ où $P_i^{k+1} = P_i^k$ si $i \neq i_o$

$P_{i_o}^{k+1} = P_o$

M^{k+1} ainsi construit est un ensemble-H relatif à W. En effet, on a déterminé lors du calcul du i_o, des scalaires

$$\rho_i^{k+1} = \rho_i^k \left(\frac{w_i}{\rho_i^k} - \frac{w_{i_o}}{\rho_{i_o}^k} \right) \qquad i \neq i_o$$

$$\rho_{i_o}^{k+1} = w_o$$

qui sont positifs et non tous nuls et qui vérifient la relation

$$\sum_{i=1}^{n+1} \rho_i^{k+1} \varepsilon_i^{k+1} g_j(P_i^{k+1}) = 0$$

où on a posé $\varepsilon_i^{k+1} = \varepsilon_i^k$ $i \neq i_o$ et $\varepsilon_{i_o}^{k+1} = \varepsilon_o$.

En fait, on prend $\rho_i^{k+1} = \dfrac{\rho_i^{k+1}}{\sum_{i=1}^{n+1} \rho_i^{k+1}}$ de sorte que $\sum_{i=1}^{n+1} \rho_i^{k+1} = 1$. La condition i de (5.1) est donc vérifiée. La configuration de signes qui lui est associée est $(\varepsilon_i^{k+1}, i = 1, \ldots, n+1)$.

La condition ii de (5.1) : $\rho^{k+1} = \sum_{i=1}^{n+1} \rho_i^{k+1} \varepsilon_i^{k+1} f(P_i^{k+1}) > 0$ est aussi vérifiée.

Ceci résulte de la proposition 5.1 énoncée ci-dessous. On trouvera le programme réalisé à partir de cet algorithme dans l'annexe 1.

5.2.2. Théorème de convergence.

On se limite à donner les énoncés du théorème de convergence et des propositions permettant de le démontrer. Les preuves de ces affirmations se trouvent dans [10]

Proposition 5.1 :

Soient ρ^k, *l'erreur égalisée à l'itération* k *sur* M^k
g_o^k, *la meilleure approximation sur* $C[M^k, \mathbf{R}]$
g, *la solution du problème d'approximation.*

On a les inégalités suivantes :

$$0 < \rho^k \leqslant \rho^{k+1} \leqslant ||g - f|| \leqslant ||g_o^k - f||$$

et l'égalité : $\rho^{k+1} = \rho^k + \rho_{i_o}^{k+1} (||g_o^k - f|| - \rho^k)$

Proposition 5.2 :

Il existe une constante $s > 0$, *telle que pour chaque itération*
$\rho_i^k \geqslant s \qquad i = 1, \ldots, n+1.$

Théorème 5.1 :

Les suites $(\rho^k)_k$ *et* $(g^k)_k$ *obtenues par l'algorithme de Remes vérifient :*

1. $\lim_{k \to +\infty} \rho^k = ||g - f||$
2. $\lim_{k \to +\infty} g^k = g$

où g *est la solution du problème d'approximation posé en (1.1).*

5.3. Algorithme d'échange point par point.

La principale différence entre cet algorithme et celui de REMES est la manière de déterminer l'indice i_o lors d'une itération. Connaissant exactement, par l'étude précédente, les ensembles-H minimaux relatifs à V_n^1, on a voulu tenir compte de leurs configurations géométriques particulières dans la détermination du point sortant de l'ensemble-H à une itération. L'algorithme d'échange point par point a été écrit et programmé uniquement pour une approximation polynomiale du premier degré, à deux variables.

5.3.1. Critère de sortie.

Le critère d'entrée est identique à celui de l'algorithme de REMES. Le critère de sortie est basé sur la position relative des points de l'ensemble-H : M et du point P_o donné par le critère d'entrée et sur les signes de ces différents points.

Trois cas peuvent se présenter :

c1. M est un ensemble-H du type H_3 (fig. 3.4) et P_o a le même signe que les points sommets du triangle (fig. 5.5, c1);

c2. M est un ensemble-H du type H_3 (fig. 3.4) et P_o a un signe opposé au signe de ces sommets (fig. 5.5, c2);

c3. M est un ensemble-H du type H_2 (fig. 3.4). Le signe de P_o ne donne pas lieu à un nouveau cas (fig. 5.5, c3).

Ainsi, il suffit de traiter ces trois cas de manière différente. Dans l'algorithme, la détermination du cas se base sur le nombre de points de signe positif de l'ensemble-H et sur le signe de P_o.

Etude de chaque cas :

Critère : La détermination de l'indice i_o se fait de façon à garder un ensemble-H minimal après l'échange.

c1 : La représentation donnée à la figure 5.6 illustre le choix suivant.

Si P_o se trouve dans la région I, il est échangé avec N_1.

Si P_o se trouve dans la région II, il est échangé avec N_2.

Si P_o se trouve dans la région III, il est échangé avec N_3.

Ces différentes régions sont délimitées par les demi-droites D_1, D_2 et D_3 passant par les points de l'ensemble-H. Des tests appropriés sur le signe de la valeur des fonctions linéaires définissant ces demi-droites en P_o apportent les renseignements nécessaires quant à la position de P_o.
Le détail de ces tests peut être trouvé dans la sous-routine EXCHAO (annexe 1).

On remarque qu'après l'échange, le nouvel ensemble-H a la même configuration géométrique.

c2 : Les demi-droites D_1, D_2, D_3 et les segments S_1, S_2, S_3 déterminent 4 régions (fig. 5.7). Par différents tests sur la valeur des fonctions linéaires définissant les droites supports de D_1, D_2, D_3, S_1, S_2, S_3, on détermine la région dans laquelle se trouve P_o. Le choix du point sortant s'effectue de la façon suivante :

1. si $P_o \in$ I , P_o est échangé avec N_1,
2. si $P_o \in$ II , P_o est échangé avec N_2,
3. si $P_o \in$ III, P_o est échangé avec N_3,
4. si $P_o \in$ IV , P_o est échangé avec B.

Dans les trois premiers cas, la configuration géométrique de l'ensemble-H change.
Le quatrième échange laisse la même configuration.
Cet échange est traité en détail dans la sous-routine EXCHA1 (annexe 1).

c3 : 4 régions sont délimitées par le segment S_1 et les demi-droites D_1, D_2, D_3, D_4 (fig. 5.8). Des tests appropriés situent P_o dans une de ces régions.

Si $P_o \in$ I , P_o est échangé avec B_1,
si $P_o \in$ II , P_o est échangé avec B_2,
si $P_o \in$ III, P_o est échangé avec N_1,
si $P_o \in$ IV , P_o est échangé avec N_2.

Aux deux premiers échanges correspond un changement de configuration géométrique du nouvel ensemble-H.
Ce cas est traité en détail dans la sous-routine EXCHA2 (annexe 1).

5.3.2. Remarque commune aux trois cas.

Le cas où P_o se trouve sur une des demi-droites D_i ou sur un des segments S_i dans chacun des cas ne permet pas l'échange de P_o avec un des points de l'ensemble-H précédent. En effet dans ce cas, deux points sont candidats à la sortie et

un échange éventuel avec un de ces points donne un ensemble-H non minimal. Dans le programme, cet échange est refusé. Un message est imprimé et l'exécution s'interrompt.

5.3.3. Description de l'algorithme.

Ayant déterminé P_o et l'indice i_o, on peut construire

$$M^{k+1} = \{P_i^{k+1} : i = 1, \ldots, 4\}$$

où $P_i^{k+1} = P_i^k$ si $i \neq i_o$ et $P_{i_o}^{k+1} = P_o$.

La configuration de signes associée à M^{k+1} est $(\varepsilon_i^{k+1} \quad i = 1, \ldots, 4)$ où $\varepsilon_i^{k+1} = \varepsilon_i^k$ $i \neq i_o$ et $\varepsilon_{i_o}^{k+1} = \varepsilon_o$.

Le calcul des ρ_i^{k+1} diffère de celui dans l'algorithme de REMES. La résolution du système (1.6) de la formulation duale permet de calculer les $\rho_i^{k+1} > 0$ $i = 1, \ldots, 4$ tels que

$$- (\varepsilon_i^{k+1} \cdot \rho_i^{k+1}, \; i = 1, \ldots, 4) \in \ker U_{M^{k+1}}$$

$$- \sum_{i=1}^{4} \rho_i^{k+1} = 1$$

$$- \sum_{i=1}^{4} \varepsilon_i^{k+1} \rho_i^{k+1} f(P_i^{k+1}) = \rho^{k+1} > 0$$

On détermine enfin, la meilleure approximation sur $C[M^{k+1}, \mathbf{R}]$: g_o^{k+1}.
Si g_o^{k+1} est solution du problème primitif, autrement dit si $||f - g_o^{k+1}|| = \max_{i=1,\ldots,4} \{|f(P_i^{k+1}) - g_o(P_i^{k+1})|\}$, l'algorithme s'arrête.

5.3.4. Exemple.

Soit $f(x, y) = \cos(\sqrt{x + y + 10})$, la fonction à approximer par un polynome $ax + by + c$ sur le domaine $B = [0, 1] \times [0, 1]$. La référence est construite avec un pas $h = 0.01$.
Autrement dit, les points de la référence sont des points de coordonnées (ih, jh) avec $0 \leqslant i \leqslant 100$ et $0 \leqslant j \leqslant 100$.

Le premier ensemble-H : M^1 est donné. On appelle les ensembles M^k : alternants.
On peut suivre la construction de M^{k+1} sur les figures 5.9, $k = 1, \ldots, \infty$.

ON EST A L'ITERATION NUMERO 1

LES POINTS DE L'ALTERNANT SONT:

(0.01 , 0.02)
(0.02 , 0.86)
(0.18 , 0.37)
(0.60 , 0.03)

LA CONFIGURATION DES SIGNES EST:

(1 1 -1 1)

L'ERREUR EGALISEE SUR L'ALTERNANT EST: 0.00074597

$P_o^1 = (1., 1.)$, $\varepsilon_o^1 = +1$. On est donc dans le cas c1. La sous-routine EXCHAO est appelée. Le point sortant est (0.60, 0.03) d'indice $i_o = 4$.

ON EST A L'ITERATION NUMERO 2

LES POINTS DE L'ALTERNANT SONT:

(0.01 , 0.02)
(0.02 , 0.86)
(0.18 , 0.37)
(1.00 , 1.00)

LA CONFIGURATION DES SIGNES EST:

(1 1 -1 1)

L'ERREUR EGALISEE SUR L'ALTERNANT EST: 0.00298076

$P_o^2 = (1.00, 0.00)$, $\varepsilon_o^1 = -1$. On est donc dans le cas c2. La sous-routine EXCHA1 est appelée. Le point sortant est (0.02, 0.86) d'indice $i_o = 2$.

ON EST A L'ITERATION NUMERO 3

LES POINTS DE L'ALTERNANT SONT:

(0.01 , 0.02)
(1.00 , 0.00)
(0.18 , 0.37)
(1.00 , 1.00)

LA CONFIGURATION DES SIGNES EST:

(1 -1 -1 1)

L'ERREUR EGALISEE SUR L'ALTERNANT EST: 0.00425192

P_o^3 = (0.00, 1.00), ε_o^1 = -1. On est donc dans le cas c3. La sous-routine EXCHA2 est appelée. L'indice i_o du point sortant (0.18, 0.37) égale 3.

ON EST A L'ITERATION NUMERO 4

LES POINTS DE L'ALTERNANT SONT:

(0.01 , 0.02)
(1.00 , 0.00)
(0.00 , 1.00)
(1.00 , 1.00)

LA CONFIGURATION DES SIGNES EST:

(1 -1 -1 1)

L'ERREUR EGALISEE SUR L'ALTERNANT EST: 0.00513542

Les échanges suivants appartiennent au cas c3 :

```
 ON EST A L'ITERATION NUMERO  5

LES POINTS DE L'ALTERNANT SONT:
                                    (   0.00 ,   0.00 )
                                    (   1.00 ,   0.00 )
                                    (   0.00 ,   1.00 )
                                    (   1.00 ,   1.00 )

LA CONFIGURATION DES SIGNES EST:
                                    (    1   -1   -1    1 )

L'ERREUR EGALISEE SUR L'ALTERNANT EST:        0.00530315

 ON EST A L'ITERATION NUMERO  6

LES POINTS DE L'ALTERNANT SONT:
                                    (   0.00 ,   0.00 )
                                    (   0.95 ,   0.02 )
                                    (   0.00 ,   1.00 )
                                    (   1.00 ,   1.00 )

LA CONFIGURATION DES SIGNES EST:
                                    (    1   -1   -1    1 )

L'ERREUR EGALISEE SUR L'ALTERNANT EST:        0.00530520

 ON EST A L'ITERATION NUMERO  7

LES POINTS DE L'ALTERNANT SONT:
                                    (   0.00 ,   0.00 )
                                    (   0.95 ,   0.02 )
                                    (   0.00 ,   0.97 )
                                    (   1.00 ,   1.00 )

LA CONFIGURATION DES SIGNES EST:
                                    (    1   -1   -1    1 )

L'ERREUR EGALISEE SUR L'ALTERNANT EST:        0.00530710
```

Le critère d'entrée n'est plus satisfait, on a :

```
CETTE ALTERNANT DONNE LA MEILLEURE APPROXIMATION SUR LE DOMAINE
LE POLYNOME DE MEILLEURE APPROXIMATION EST:

P(X,Y)=-1.00509317  X** 0 + 0.02567144  X** 1 + 0.02567144  Y** 1
```

5.3.5. Résultats numériques.

1. Remarque :

Si on modifie le critère d'entrée de façon à prendre P_o le premier point de la référence qui satisfait $|f(P_o) - g_o^k(P_o)| > \rho^k$, l'expérience montre que le nombre d'itérations nécessaires pour arriver à la solution est beaucoup plus élevé.

Un essai a été fait pour l'approximation de la fonction $x^2 + \sin y$. A la deuxième itération, avec le critère d'entrée d'origine, l'erreur égalisée vaut : 0.980.
Avec le critère ci-dessus, la valeur 0.965 est obtenue seulement à la 82ème itération.

2. Comparaison des temps de calculs :

Les tableaux ci-dessous donnent les temps nécessaires aux calculs relatifs à l'approximation de fonctions dans les algorithmes de Remes et d'échange point par point.

Tableau 1 : Comparaison des temps (en millisecondes) nécessaires à chaque itération dans l'approximation de la fonction $f(x,y) = e^{5xy}$.

Algorithme de Remes	Algorithme d'échange point par point	Différence
162	161	-1
19.581	19.628	47
19.539	19.066	-473
19.153	19.646	493
19.883	19.430	-453
19.829	19.057	-773
19.946	19.608	-338
19.982	19.215	-767
138075	135.811	-2.264

Tableau 2 : Comparaison des temps (en millisecondes) nécessaires à l'approximation de différentes fonctions.

Algorithme de Remes	Algorithme d'échange point par point	Différence
76.802	74.717	-2.085
130.936	129.516	-1.420
84.031	83.881	-152
122.627	124.041	1.414
117.774	116.573	-1.201
94.184	94.438	254
87.330	86.436	-894

Les différences des temps d'exécution entre les deux algorithmes étant généralement petites, il n'est pas possible de tirer des conclusions valables à propos de leur rapidité respective. En effet, les temps donnés par l'ordinateur sont sujets à des fluctuations pouvant atteindre 1 % .

Le caractère négligeable de ces différences est dû au fait que les opérations prenant beaucoup de temps, comme la résolution d'un système d'équations par la méthode de Gauss, apparaissent de la même façon dans les deux algorithmes.

Il est à remarquer que, dans l'algorithme d'échange point par point, les passages dans la sous-routine qui possède le plus petit nombre de tests, EXCHA0 , ne sont, en général, pas plus rapides que les passages dans les autres sous-routines.

Les gains éventuels sont uniquement dûs à la position géométrique du point d'erreur maximale, en valeur absolue, par rapport à celle de l'alternant. Une autre disposition des tests dans les sous-routines EXCHA0 , EXCHA1, EXCHA2 peut faire gagner du temps pour certains échanges mais en faire perdre pour d'autres.

5.4. Algorithme d'échange global.

Prenant toujours en considération la configuration géométrique des ensembles-H minimaux relatifs à V_2^1, cet algorithme permet de construire à l'itération k+1, l'ensemble-H minimal : M^{k+1} en échangeant un maximum de points.

On n'exposera que l'algorithme permettant l'échange des points de M^k avec un maximum de points de $B \setminus M^k$. Le calcul des coefficients du polynome approximant se fait de la même manière que dans l'algorithme de REMES. De plus, le vecteur $(\lambda_i, i = 1, \ldots, 4)$ qui vérifie $(\lambda_1, \ldots, \lambda_4) \in \ker U_{M^{k+1}}$ et i et ii de 5.2 est une des solutions du système d'équations (1.6), tout comme dans l'algorithme d'échange point par point.

5.4.1. Philosophie générale de l'échange.

Par ce qui précède, on sait que les candidats à l'entrée dans l'ensemble M^{k+1} sont les points de $B \setminus M^k$ où l'erreur en valeur absolue est supérieure à l'erreur égalisée ρ^k sur l'ensemble-H M^k. En fait, dans les algorithmes d'échange point par point, on choisit parmi ces candidats le point où l'erreur est maximale en valeur absolue. On note n, le nombre de candidats à l'entrée. Plusieurs cas peuvent se présenter :

c1 : Si n = 0, aucun point dans $B \setminus M^k$ ne possède une erreur en valeur absolue supérieure à ρ^k. Le polynome approximant trouvé à l'itération k : g_o^k est la meilleure approximation. L'algorithme s'arrête.

c2 : Si n = 1, un seul point est susceptible d'être échangé. Dans ce cas, l'échange point par point (5.3) est appliqué.

c3 : Si n = 2, on construit M^{k+1} avec ces deux points et deux points de M^k, de façon à ce que M^{k+1} soit un ensemble-H minimal relatif à V_2^1.

c4 : Si n = 3, on détermine le point de M^k qui avec ces trois points forme un ensemble-H minimal que l'on nomme M^{k+1}.

c5 : Si $n \geqslant 4$, on essaye de trouver dans les candidats à l'entrée, 4 points qui ont la configuration géométrique et la configuration de signes d'un ensemble-H minimal : M^{k+1}.

Remarque : Dans la suite, on parle de la référence R. Cette référence est un ensemble de discrétisation du domaine B considéré, pour cette approximation par des polynomes de V_2^1, comme un pavé [a, b] x [c, d].

5.4.2. Constatations.

Différents essais ont été effectués pour diminuer le nombre d'itérations et le temps d'exécution d'une itération. Aussi, l'effort a été porté sur une présélection parmi les candidats à l'entrée. Ceci découle des deux remarques suivantes :

1. Classant les candidats dans deux ensembles (vecteurs), l'un noté P, contenant les points d'erreur positive, $> \rho^k$ et l'autre noté M, tous les points d'erreur négative, $< -\rho^k$, on a constaté qu'un grand nombre d'ensembles de 4 points comprenant 3 points de P et 1 point de M, ou 2 points de P et 2 points de M ou 1 point de P et 3 points de M, ne formaient pas un ensemble-H. D'autre part, la convergence de la suite $(\rho^k)_k$ était très lente. Très vite, on a abandonné cette façon de procéder.

2. Construisant les ensembles (vecteurs) P et M comme précédemment, on les a ordonnés : P de façon à ce que l'erreur, en deux points successifs dans P, soit décroissante et M, de façon à ce que cette erreur, pour des points successifs dans M, soit croissante.
   Ainsi donc, on favorisait les points d'erreur, en valeur absolue relativement grande. Mais les essais ont montré que les premiers points de P étaient des points voisins dans la référence. Il en était de même, pour les points de M. Par conséquent, les premiers ensembles de 4 points testés ne formaient pas en général un ensemble-H. Par contre, la valeur de l'erreur égalisée croissait plus rapidement.

Il fallait donc trouver un moyen de sélectionner parmi les candidats, des points assez dispersés dans la référence et où l'erreur était suffisamment grande en valeur absolue par rapport à l'erreur égalisée.

Un mode de présélection a été élaboré et testé. Il est présenté ci-après, de même qu'une description plus détaillée du traitement des cas $n \geqslant 4$, $n = 3$ et $n = 2$.

5.4.3. Construction de M^{k+1}.

1. Echange de 4 points.

On suppose que le nombre de candidats à l'entrée : $n \geqslant 4$.
Mode de présélection : On divise R en quatre sous-ensembles Q_1, Q_2, Q_3, Q_4 tels que

$$Q_1 = R \cap ([a, \frac{b-a}{2}[\times [c, \frac{d-c}{2}])$$

$$Q_2 = R \cap ([\frac{b-a}{2}, b] \times [c, \frac{d-c}{2}])$$

$$Q_3 = R \cap ([a, \frac{b-a}{2}[\times]\frac{d-c}{2}, d])$$

$$Q_4 = R \cap ([\frac{b-a}{2}, b] \times]\frac{d-c}{2}, d])$$

On choisit dans chacun d'eux, le point où l'erreur est maximale et $> \rho^k$ et le point où l'erreur est minimale et $< -\rho^k$.

Ainsi donc, on dispose au plus de huit points, que l'on classe dans deux ensembles (vecteurs) P et M suivant que le signe de l'erreur commise en ces points est positive ou négative.

On appelle NMP le nombre de points présélectionnés, NM, le nombre de points présélectionnés où l'erreur est négative et NP, le nombre de points présélectionnés où l'erreur est positive.

Si NMP = 4 et NM = 0 ou NP = 0, il est inutile de chercher une configuration d'ensemble-H parmi ces 4 points. On ne pourra donc échanger qu'au plus 3 points. Dans le cas contraire, NMP $\geqslant$ 4 et NM $\neq$ 0 et NP $\neq$ 0, on espère trouver un ensemble-H minimal parmi les points sélectionnés. Pour cela, on teste si 4 points dont 2 sont choisis dans M et deux dans P, ne forment pas un ensemble-H minimal, par la sous-routine CARRE (annexe 1). Si ce test ne permet pas de construire M^{k+1}, on regarde alors, à l'aide de la sous-routine TRIGLE (annexe 1), tous les ensembles formés de 3 points de M et 1 point de P ou 3 points de P et 1 point de M.
Si M^{k+1} n'a pu être construit par un échange de 4 points, on essaye d'échanger 3 points, puis 2 points si la conclusion reste négative, et enfin, on essaye d'échanger le point d'erreur maximale en valeur absolue.

2. Echange de 3 points.

On envisage cet échange quand NMP = 3 et dans le cas où M^{k+1} n'a pu être construit

par un échange de 4 points.
La construction de M^{k+1} se base sur la recherche d'un point de M^k qui, combiné avec 3 points de $P \cup M$, donne un ensemble-H minimal. La recherche s'avère différente lorsque les 3 points (choisis) de $P \cup M$ sont de même signe ou de signes différents. Dans le premier cas, les tests sont effectués dans la sous-routine E3EQSI, dans le second cas, ils sont effectués dans la sous-routine E3NESI (annexe 1).

3. Echange de 2 points.

Cet échange est envisagé lorsque NMP = 2 ou lorsqu'un échange de 4 points ou de 3 points ne s'avère pas possible quand NMP $\geqslant$ 4 ou NMP = 3. On recherche donc 2 points de M^k qui avec (les) deux points de $P \cup M$ forment un ensemble-H minimal.
On tient compte de la différence ou de l'égalité des signes des deux points candidats à l'entrée pour orienter la recherche. Les sous-routines E2NESI et E2EQSI effectuent les tests nécessaires respectivement quand les signes sont différents, égaux.

Remarque :
Toutes les sous-routines envisagées pour un échange de plusieurs points appellent les sous-routines CARRE et TRIGLE.
CARRE : Soient A, B les deux points où l'erreur est positive et P, Q les deux points où l'erreur est négative. On regarde si A et B se trouvent de part et d'autre de la droite passant par P et Q, et inversément.
Si A, B, P, Q satisfont à ces conditions, ils forment un ensemble-H minimal.
TRIGLE : Si A, B, C désignent les points où l'erreur est de même signe et Q, le point où l'erreur est de signe opposé, et si Q appartient au domaine triangulaire dont les côtés sont les segments de droite déterminés par les couples de points (A, B) (B, C) et (C, A), alors {A, B, C, Q} est un ensemble-H minimal.

La dégénérescence, autrement l'ensemble-H minimal constitué de trois points en ligne droite où l'erreur est alternée n'est pas retenue pour former un ensemble de la suite $(M^k)_k$.

5.4.4. Exemple.

On traite la fonction $x^2 + \sin y$.

Le premier alternant M^1 est le même qu'en 5.3.4. On peut suivre la construction de M^{k+1} sur la figure 5.10.

```
ON EST A L'ITERATION NUMERO  1

LES POINTS DE L'ALTERNANT SONT:
                                (  0.01 ,  0.02 )
                                (  0.02 ,  0.86 )
                                (  0.18 ,  0.37 )
                                (  0.60 ,  0.03 )

LA CONFIGURATION DES SIGNES EST:
                                (   1   1  -1   1 )

L'ERREUR EGALISEE SUR L'ALTERNANT EST:      0.01758097

                ON A ECHANGE 4 POINTS
```

Les quatre points sont échangés. La sous-routine "carrée" est appelée.

```
ON EST A L'ITERATION NUMERO  2

LES POINTS DE L'ALTERNANT SONT:
                                (  0.31 ,  0.00 )
                                (  0.31 ,  1.00 )
                                (  0.00 ,  0.50 )
                                (  1.00 ,  0.50 )

LA CONFIGURATION DES SIGNES EST:
                                (  -1  -1   1   1 )

L'ERREUR EGALISEE SUR L'ALTERNANT EST:      0.13629502

                ON A ECHANGE 4 POINTS
```

Les quatre points sont échangés. La sous-routine carrée est appelée.

```
ON EST A L'ITERATION NUMERO  3

LES POINTS DE L'ALTERNANT SONT:
                                (   0.50 ,   0.00 )
                                (   0.50 ,   1.00 )
                                (   0.00 ,   0.57 )
                                (   1.00 ,   0.57 )

LA CONFIGURATION DES SIGNES EST:
                                (  -1   -1    1    1 )

L'ERREUR EGALISEE SUR L'ALTERNANT EST:        0.15499679
```

Le critère d'entrée n'est pas satisfait : on a

```
CETTE ALTERNANT DONNE LA MEILLEURE APPROXIMATION
LE POLYNOME DE MEILLEURE APPROXIMATION EST:

P(X,Y)= -0.09500321  X** 0 + 1.00000000  X** 1 + 0.84147098  Y** 1
```

5.4.5. Résultats numériques.

L'algorithme a été testé sur quelques exemples. Le tableau ci-dessous donne la comparaison des temps d'exécution d'un problème entier et la comparaison des nombres d'itérations.

Algorithme de Remes		Algorithme d'échange global		Différence	
Temps	Itération	Temps	Itération	Temps	Itération
136.686	8	39.639	3	-97.047	-5
130.936	7	48.187	5	-82.749	-2
122.627	7	118.989	7	-3.638	-0
117.774	8	83.395	6	-34.379	-2
87.331	6	52.035	4	-35.296	-2
84.031	6	97.483	7	13.452	+1
94.184	7	83.395	6	-13.945	-1

On remarque que les différences de temps sont relativement variables.

Le mode de présélection adopté ne donne pas toujours des points bien disposés dans la référence. De ce fait, on ne peut souvent qu'échanger un, deux ou trois points, et donc l'erreur égalisée croît moins rapidement que lors d'un échange de quatre points.

En outre, on a remarqué que le fait de tester beaucoup d'ensembles de quatre points, toujours relativement dispersés dans la référence, n'augmente guère le temps nécessaire à une itération.

Les résultats obtenus montrent l'efficacité de l'algorithme dans la plupart des cas. Une étude poussée des cas d'exception et de variantes dans le mode de présélection (diviser en neuf sous-ensembles) pourrait permettre de rendre cet algorithme plus efficace.

Annexe 1 : Programmes.

```
      PROGRAM APPROX

C     * * * * * * * * * * * * * * * * * * * * * * * * * * * *
C     * LE PROGRAMME SUIVANT CALCULE LA MEILLEURE           *
C     * APPROXIMATION POLYNOMIALE D'UNE FONCTION A          *
C     * DEUX VARIABLES PAR L'ALGORITHME DE REMEZ            *
C     * * * * * * * * * * * * * * * * * * * * * * * * * * * *

C     TABLES DE VARIABLES
C     ←.←.←.←.←.←.←.←.←.←.←

C     REELS
C     ← ← ←
C       PASX   : PAS DE CONSTRUCTION DE LA REFERENCE
C                SUR L'AXE DES ABSCISSES
C       PASY   : PAS DE CONSTRUCTION DE LA REFERENCE
C                SUR L'AXE DES ORDONNEES
C       TOLER  : INDICE DE TOLERANCE SUR LA PRECISION
C       XR     : VECTEUR CONTENANT LES ABSCISSES DES
C                POINTS DE LA REFERENCE
C       XMIN   : ABSCISSE MINIMALE DES POINTS DE LA
C                REFERENCE
C       XMAX   : ABSCISSE MAXIMALE DES POINTS DE LA
C                REFERENCE
C       YR     : VECTEUR CONTENANT LES ORDONNEES DES
C                POINTS DE LA REFERENCE
C       YMIN   : ORDONNEE MINIMALE DES POINTS DE LA
C                REFERENCE
C       YMAX   : ORDONNEE MAXIMALE DES POINTS DE LA
C                REFERENCE
C     ENTIERS
C     ← ← ← ←
C       DEGRE  : DEGRE DU POLYNOME APPROXIMANT
C       IP     : INDICE D'IMPRESSION
C       IA     : VECTEUR CONTENANT LES INDICES DES
C                POINTS DE L'ALTERNANT
C       MAXIT  : NOMBRE MAXIMAL  D'ITERATIONS PERMISES
C       NF     : NOMBRE DE PROBLEMES TRAITES
C       NX     : NOMBRE D'ABSCISSES DIFFERENTES
C       NY     : NOMBRE D'ORDONNEES DIFFERENTES
C       NFE    : NOMBRE DE FONCTIONS DE BASE
C       NAL    : NOMBRE DE POINTS DE L'ALTERNANT
C       NR     : NOMBRE DE POINTS DE LA REFERENCE

C     LES VARIABLES QUI NE SONT PAS REPRISES DANS
C     CETTE TABLE SONT DES VARIABLES INTERMEDIAIRES
```

```
C       DECLARATIONS

        IMPLICIT REAL*8 (A-H,O-Z)
        COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
        INTEGER DEGRE
        CALL TIME(NOTIME)

C       LECTURES

  100   READ 1001,DEGRE,IP,MAXIT,NF
        IF(DEGRE.LT.0) STOP
        READ 1000,XMIN,XMAX,NX
        READ 1000,YMIN,YMAX,NY
        READ 1002,TOLER
 1000   FORMAT(2F6.2,I4)
 1001   FORMAT(4(10X,I4))
 1002   FORMAT(D10.3)

C       CONSTRUCTION DE LA REFERENCE

        PASX=(XMAX-XMIN)/NX
        PASY=(YMAX-YMIN)/NY
        NX=NX+1
        NY=NY+1
        DO 6 J=1,NX
        J1=J-1
        X=XMIN+J1*PASX
        DO 7 I=1,NY
        M=J1*NY+I
        XR(M)=X
        YR(M)=YMIN+(I-1)*PASY
    7   CONTINUE
    6   CONTINUE

C       NOMBRE DE POINTS DANS LA REFERENCE

        NR=NX*NY

C       NOMERE DE FONCTIONS DE BASE.

        NFB=(DEGRE+1)*(DEGRE+2)/2

C       NOMERE DE POINTS DANS L'ALTERNANT

        NAL=NFB+1
```

```
C       APPEL A LA SOUS-ROUTINE REMEZ POUR CHAQUE FONCTION

      DO 1 NUMF=1,NF
      PRINT 1108,NUMF
 1108 FORMAT(1H1,1X,' APPROXIMATION DE LA FONCTION ',I4)
      CALL TIME(NOTIME)

C       LECTURE DE L'ALTERNANT.

      READ 1003,(IA(I),I=1,NAL)
 1003 FORMAT(10I4)
      CALL REMEZ(NUMF,TOLER,IP,MAXIT)
    1 CONTINUE
      GOTO 100
      END

      DOUBLE PRECISION FUNCTION FB(I,X,Y)

      IMPLICIT REAL*8 (A-H,O-Z)

C       FONCTIONS DE BASE DE L'APPROXIMATION POLYNOMIALE.

      GOTO(1,2,3,4,5,6,7,8,9,10),I
    1 FB=1.D00
      RETURN
    2 FB=X
      RETURN
    3 FB=Y
      RETURN
    4 FB=X**2
      RETURN
    5 FB=Y**2
      RETURN
    6 FB=X*Y
      RETURN
    7 FB=X**3
      RETURN
    8 FB=X**2*Y
      RETURN
    9 FB=X*Y**2
      RETURN
   10 FB=Y**3
      RETURN
      END
```

```
      SUBROUTINE GAUSS(A,B,N,EPS,IERR)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      DIMENSION A(100,100),B(100),IL(50)

C       RESOLUTION PAR DIAGONALISATION DE GAUSS DU SYSTEME
C       DE N EQUATIONS A N INCONNUES.
C       LA MATRICE A SE TRANSFORME LORS DES CALCULS.
C       LE RESULTAT SE TROUVE DANS B.

      DO 5 I=1,N
    5 IL(I)=0
      IERR=0
      DO 100 JJ=1,N
      AMAX=0.D00
      DO 10 I=1,N
      IF(IL(I).NE.0) GOTO 10
    7 IF(DABS(A(I,JJ)).LE.DABS(AMAX))GOTO10
      AMAX=A(I,JJ)
      IMEM=I
   10 CONTINUE
      IF(DABS(AMAX).LT.EPS) GOTO 200
      IL(IMEM)=JJ
      DO 20 J=1,N
   20 A(IMEM,J)=A(IMEM,J)/AMAX
      B(IMEM)=B(IMEM)/AMAX
      DO 40 I=1,N
      IF(I.EQ.IMEM) GOTO 40
      DO 30 J=1,N
      IF(J.EQ.JJ) GOTO 30
      A(I,J)=A(I,J)-A(I,JJ)*A(IMEM,J)
   30 CONTINUE
      B(I)=B(I)-A(I,JJ)*B(IMEM)
   40 CONTINUE
      DO 50 I=1,N
   50 A(I,JJ)=0.D00
      A(IMEM,JJ)=1.D00
      GOTO 100
  200 IERR=JJ
      PRINT 1001,JJ
 1001 FORMAT(1X,'COLONNE',I2,'NULLE')
      RETURN
  100 CONTINUE
      DO 110 I=1,N
  110 A(I,1)=B(I)
      DO 120 I=1,N
      IL1=IL(I)
  120 B(IL1)=A(I,1)
      RETURN
      END
```

```
      DOUBLE PRECISION FUNCTION E(X,Y,ALPHA,NUMF)


      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
      DIMENSION ALPHA(100)


C       FONCTION QUI DONNE L'ERREUR ENTRE LE POLYNOME
C       APPROXIMANT ET LA FONCTION OBJECTIF.


      E=ALPHA(1)*FB(1,X,Y)
      DO 1 I=2,NFB
    1 E=E+ALPHA(I)*FB(I,X,Y)
      E=F(X,Y,NUMF)-E
      RETURN
      END



      DOUBLE PRECISION FUNCTION F(X,Y,I)


      IMPLICIT REAL*8 (A-H,O-Z)


C       LES FONCTIONS OBJECTIFS.


      GOTO (1,2,3,4,5,6,7,8,9,10),I
    1 F=DEXP(X*Y)
      RETURN
    2 F=DCOS(DSQRT(X+Y+10.))
      RETURN
    3 F=X**2+Y**3
      RETURN
    4 F=DSIN(X+Y**2)
      RETURN
    5 F=(10*X+DSQRT(Y))/(1+X+Y)
      RETURN
    6 F=DSQRT(X+Y+3)
      RETURN
    7 F=DEXP(5*X*Y)
      RETURN
    8 F=(X+Y+10)/(X**2+Y**2+1)
      RETURN
    9 F=DEXP(X*Y)
      RETURN
   10 F=DCOS(DSQRT(X+Y+10))
      RETURN
      END
```

```
      SUBROUTINE REMEZ(NUMF,TOLER,IP,MAXIT)

C       TABLES DE VARIABLES
C       +.+.+.+.+.+.+.+.+.+.+

C       REELS
C       + + +
C         AUX    : MATRICE ASSOCIEE A L'HOMOMORPHISME U
C         A      : MATRICE AUXILIAIRE
C         ALPHA  : VECTEUR DES COEFFICIENTS DU POLYNOME
C         B      : VECTEUR SECOND MEMBRE A L'ENTREE DE
C                  GAUSS,
C                  VECTEUR SOLUTION A LA SORTIE
C                  GAUSS
C         IMP    : VECTEUR SERVANT A L'IMPRESSION
C         ROI    : VECTEURS DES VALEURS ABSOLUES DES
C                  COMPOSANTES DU VECTEUR SOLUTION B
C         RC     : ERREUR EGALISEE SUR L'ALTERNANT
C         U      : MATRICE SERVANT AU CALCUL DE IOLD

C       ENTIERS
C       + + + +
C         ISGNO  : SIGNE DE L'ERREUR AU POINT QUI ENTRE
C                  DANS L'ALTERNANT
C         ISGN   : VECTEUR DES SIGNES DE L'ERREUR AUX
C                  POINTS DE L'ALTERNANT
C         INEW   : INDICE DANS LA REFERENCE DU POINT
C                  QUI ENTRE DANS L'ALTERNANT
C         IOLD   : INDICE DANS L'ALTERNANT DU POINT
C                  QUI SORT DE L'ALTERNANT
C         NUMIT  : NUMERO DE L'ITERATION EN COURS
C         NUMF   : NUMERO DU PROBLEME TRAITE

C       FONCTIONS
C       + + + + +
C         E      : FONCTION ERREUR
C         FB     : FONCTIONS DE BASE
C         F      : FONCTION A APPROXIMER


C       DECLARATIONS


      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
      DIMENSION AUX(100,100),A(100,100),ALPHA(100),
     CISGN(100),U(100,100),ROI(100),B(100),IMP(100)
      NUMIT=1


C       PREMIERE ITERATION.

C       LES CALCULS EXECUTES JUSQU'A L'APPEL A LA SOUS-
C       ROUTINE COEFF SONT SPECIFIQUES A CETTE ITERATION.


      CALL TIME(NOTIME)
```

```
C       CONSTRUCTION DE LA MATRICE AUX.

      DO 1 I=1,NFB
      DO 2 J=1,NAL
      K=IA(J)
    2 AUX(I,J)=FB(I,XR(K),YR(K))
    1 CONTINUE

C       RESOLUTION PAR DIAGONALISATION DE GAUSS DU
C       SYSTEME D'EQUATIONS DE LA FORMULATION DUALE (1,6)

      DO 3 I=1,NFB
      B(I)=-AUX(I,NAL)
      DO 3 J=1,NFB
    3 A(I,J)=AUX(I,J)
      CALL GAUSS(A,B,NFB,1.D-13,JERR)
      IF(JERR.EQ.0) GOTO 10
      PRINT 2010,JERR
 2010 FORMAT(///,1X,'L''ALTERNANT DE LA PREMIERE',
     C'ITERATION DONNE UN SYSTEME INSOLUBLE',/,
     C'CAR ',I3,' COLONNE EST NULLE')
      RETURN

C       CALCUL DE RO,ISGN,ROI

   10 B(NAL)=1.D00
      S=0.D00
      DO 11 I=1,NAL
   11 S=S+DABS(B(I))
      S1=S
      DO 14 I=1,NAL
      ROI(I)=DABS(B(I))/S
      IF(B(I)) 13,14,12
   12 ISGN(I)=1
      GOTO 14
   13 ISGN(I)=-1
   14 CONTINUE
      S=0.D00
      DO 15 I=1,NAL
      K=IA(I)
   15 S=S+B(I)*F(XR(K),YR(K),NUMF)
      IF(S.LT.0.D00) GOTO 16
      RO=S/S1
      GOTO 19
   16 RO=-S/S1
      DO 17 I=1,NAL
   17 ISGN(I)=-ISGN(I)

C       CONSTRUCTION DE LA MATRICE U

   19 DO 20 I=1,NFB
      DO 20 J=1,NAL
   20 U(I,J)=ISGN(J)*AUX(I,J)
```

```
C       LA SUITE DES CALCULS EST COMMUNE A TOUTES LES
C       ITERATIONS.

C       APPEL A LA SOUS-ROUTINE QUI CALCULE LES
C       COEFFICIENTS DU POLYNOME APPROXIMANT.


  100 CALL COEFF(TOLER,AUX,ALPHA,ISGN,RO,IERR,NUMF)
      IF(IERR.EQ.0) GOTO 28
      GOTO (25,26),IERR
   25 PRINT 2025
 2025 FORMAT(///,1X,'SYSTEME DE COEFF EST INSOLUBLE')
      RETURN
   26 PRINT 2026
 2026 FORMAT(///,1X,'LA DERNIERE EQUATION DU SYSTEME',
     C'DE COEFF N''EST PAS VERIFIEE')


C       IMPRESSION.


   28 IF(IP.EQ.0) GOTO 30
      IF(NUMIT.EQ.1) GOTO 29
      IRAP=MOD(NUMIT,IP)
      IF(IRAP.NE.0) GOTO 30
   29 PRINT 2004,NUMIT
      PRINT 2002
      DO 24 I=1,NAL
   24 PRINT 2003,XR(K),YR(K)
      PRINT 2005
      PRINT 2006,(ISGN(I),I=1,NAL)
      PRINT 2007,RO
 2002 FORMAT(///,1X,'LES POINTS DE L''ALTERNANT SONT:')
 2003 FORMAT(37X,'( ',F6.2,' , ',F6.2,' )')
 2004 FORMAT(1H1,' ON EST A L''ITERATION NUMERO ',I2)
 2005 FORMAT(//,1X,'LA CONFIGURATION DES SIGNES EST:')
 2006 FORMAT(37X,'( ',4I4,' )')
 2007 FORMAT(//,1X,'L''ERREUR EGALISEE SUR ',
     C'L''ALTERNANT EST:',2X,F15.8)
      CALL TIME(NOTIME)


C       CALCUL DU POINT D'ERREUR MAXIMALE SUR LA
C       REFERENCE,ON NOTE INEW L'INDICE DE CE POINT


   30 INEW=1
      ERR=E(XR(1),YR(1),ALPHA,NUMF)
      DO 31 J=2,NR
      EAUX=E(XR(J),YR(J),ALPHA,NUMF)
      IF(DABS(EAUX).LT.DABS(ERR)) GOTO 31
      ERR=EAUX
      INEW=J
   31 CONTINUE
```

```
      IF(DABS(RO-DABS(ERR)).GT.1.D-08) GOTO 33
   35 PRINT 2011
      PRINT 2012
      DO 310 I=1,NFB
  310 IMP(I)=I-1
      PRINT 2013,(ALPHA(I),IMP(I),I=1,NFB)
 2011 FORMAT(//,1X,'CET  ALTERNANT DONNE LA MEILLEURE',
     C' APPROXIMATION SUR LE DOMAINE (A,B)*(C,D) ')
 2012 FORMAT(/,1X,'LE POLYNOME DE MEILLEURE APPROXIMATION EST:')
 2013 FORMAT(//,1X,'P(X,Y)= ',6(F16.8,2X,'X**',I2 ' + '))
      RETURN
   33 IF(ERR.LT.0.D00) GOTO 32
      ISGNO=+1
      GOTO 34
   32 ISGNO=-1


C        PASSAGE A L'ITERATION SUIVANTE.


   34 NUMIT=NUMIT+1
      IF(NUMIT.LT.MAXIT) GOTO 101
      PRINT 2030,MAXIT
 2030 FORMAT(///,'LE CALCUL DE LA MEILLEURE APPROXIMATION DEMANDE ',
     C'PLUS DE',I4,' ITERATIONS.')
      RETURN


C        APPEL A LA SOUS-ROUTINE QUI ECHANGE DEUX POINTS.


  101 CALL EXCH(NUMF,AUX,U,ISGN,RO,IOLD,INEW,IERR,ROI,SUM,ISGNO)
      IF (IERR.EQ.1) GOTO 60


C        CALCUL DE RO.


      RO=RO+(DABS(ERR)-RO)/SUM


C        L'ETAPE SUIVANTE EST LE CALCUL DES COEFFICIENTS.


      GOTO 100
   60 PRINT 2024
 2024 FORMAT(//,1X,'LE SYSTEME DANS LA SOUS-ROUTINE EXCH EST INSOLUBLE')
      RETURN
      END
```

```
      SUBROUTINE COEFF(TOLER,AUX,ALPHA,ISGN,RO,IERR,NUMF)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
      DIMENSION AUX(100,100),ALPHA(100),ISGN(100),
     CA(100,100)
      IERR=0

C       CALCUL DES COEFFICIENTS DU POLYNOME APPROXIMANT
C       PAR LA RESOLUTION D'UN SYSTEME DE NAL EQUATIONS
C       A NFB INCONNUES.

      DO 1 I=1,NFB
      DO 2 J=1,NFB
    2 A(I,J)=AUX(J,I)
      K=IA(I)
    1 ALPHA(I)=F(XR(K),YR(K),NUMF)-RO*ISGN(I)
      CALL GAUSS(A,ALPHA,NFB,1.D-13,JERR)
      IF(JERR.NE.0) GOTO 5

C       VERIFICATION DE LA COMPATIBILITE DE LA DERNIERE
C       EQUATION AVEC LA SOLUTION DU SYSTEME.
C       LA THEORIE AFFIRME CETTE COMPATIBILITE.

      S=0.D00
      DO 3 I=1,NFB
    3 S=S+ALPHA(I)*AUX(I,NAL)
      K=IA(NAL)
      VALUE=F(XR(K),YR(K),NUMF)
   30 IF(DABS(S+ISGN(NAL)*RO-VALUE).LT.TOLER) GOTO 6
      IERR=2
      RETURN
    5 IERR=1
    6 RETURN
      END
```

```
      SUBROUTINE EXCH(NUMF,AUX,U,ISGN,RO,IOLD,INEW,IERR,ROI,S,ISGNO)
C
C       DECLARATIONS.
C
      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
      DIMENSION AUX(100,100),ISGN(100),U(100,100)
      DIMENSION A(100,100),W(100),UO(100),ROI(100)
C
C       CALCUL DE L'INDICE DU POINT QUI SORT DE
C       L'ALTERNANT.
C
      DO 10 I=1,NFB
      UO(I)=ISGNO*FB(I,XR(INEW),YR(INEW))
      DO 9 J=1,NFB
    9 A(I,J)=U(I,J)
   10 W(I)=-UO(I)
      EPS1=1.D-14
      CALL GAUSS(A,W,NFB,EPS1,JERR)
      IF(JERR.EQ.0) GOTO 11
      IERR=1
      RETURN
   11 AMIN=W(1)/ROI(1)
      IOLD=1
      W(NAL)=0.D00
      DO 12 I=2,NAL
      IF(AMIN.LE.W(I)/ROI(I)) GOTO 12
      AMIN=W(I)/ROI(I)
      IOLD=I
   12 CONTINUE
C
C       CALCUL DE ROI
C
      S=0.D00
      DO 15 I=1,NAL
      IF (I.EQ.IOLD) GOTO 15
      ROI(I)=W(I)-ROI(I)*AMIN
      S=S+ROI(I)
   15 CONTINUE
      ROI(IOLD)=1.D00
      S=S+ROI(IOLD)
      DO 16 I=1,NAL
   16 ROI(I)=ROI(I)/S
C
C       ECHANGE PROPREMENT DIT ENTRE LES DEUX POINTS.
C       MODIFICATION DES MATRICES AUX ET U,DU VECTEUR
C       DES SIGNES ISGN.
C
      IA(IOLD)=INEW
      DO 13 I=1,NFB
   13 AUX(I,IOLD)=FB(I,XR(INEW),YR(INEW))
      ISGN(IOLD)=ISGNO
      DO 14 I=1,NFB
   14 U(I,IOLD)=UO(I)
      RETURN
```

```
C       * * * * * * * * * * * * * * * * * * * * * * * * *
C       * LE PROGRAMME SUIVANT CALCULE LA MEILLEURE      *
C       * APPROXIMATION POLYNOMIALE D'UNE FONCTION A     *
C       * DEUX VARIABLES PAR UNE METHODE UTILISANT       *
C       * LES CONFIGURATIONS PARTICULIERES DES           *
C       * ENSEMBLES-H.                                   *
C       * * * * * * * * * * * * * * * * * * * * * * * * *


      SUBROUTINE REMEZ(NUMF,TOLER,IP,MAXIT)


C       TABLES DE VARIABLES
C       ←.←.←.←.←.←.←.←.←.←

C       REELS
C       ← ← ←
C          A      : MATRICE AUXILIAIRE
C          AUX    : MATRICE ASSOCIEE A L'HOMOMORPHISME U
C          ALPHA  : VECTEUR DES COEFFICIENTS DU POLYNOME
C                   VECTEUR SOLUTION A LA SORTIE
C          RO     : ERREUR EGALISEE SUR L'ALTERNANT

C       ENTIERS
C       ← ← ← ←
C          IMP    : VECTEUR SERVANT A L'IMPRESSION
C          INEW   : INDICE DANS LA REFERENCE DU POINT
C                   QUI ENTRE DANS L'ALTERNANT
C          IOLD   : INDICE DANS L'ALTERNANT DU POINT
C                   QUI SORT DE L'ALTERNANT
C          ISGNO  : SIGNE DE L'ERREUR AU POINT QUI ENTRE
C                   DANS L'ALTERNANT
C          ISGN   : VECTEUR DES SIGNES DE L'ERREUR AUX
C                   POINTS DE L'ALTERNANT
C          IP     : VARIABLE D'IMPRESSION
C          KP     : VARIABLE QUI COMPTE LE NOMBRE DE POINTS
C                   D'ERREUR POSITIVE
C          NUMIT  : NUMERO DE L'ITERATION EN COURS
C          NUMF   : NUMERO DU PROBLEME TRAITE

C       BOOLEENNES
C       ← ← ← ← ←
C          DG     : SI DG=TRUE ALORS ON NE PEUT ECHANGER
C                   LE POINT D'ERREUR MAXIMALE

C       FONCTIONS
C       ← ← ← ← ←
C          F      : FONCTION A APPROXIMER
C          FB     : FONCTIONS DE BASE
C          E      : FONCTION ERREUR
C          D      : EQUATION DE LA DROITE PASSANT PAR DEUX POINTS
```

```
C       DECLARATIONS.

        IMPLICIT REAL*8 (A-H,O-Z)
        COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
        DIMENSION AUX(10,10),ALPHA(10),ISGN(10),IMP(10)

        NUMIT=1
        CALL TIME(NOTIME)

C       CONSTRUCTION DE LA MATRICE AUX.

        DO 1 I=1,NFB
        DO 2 J=1,NAL
        K=IA(J)
    2   AUX(I,J)=FB(I,XR(K),YR(K))
    1   CONTINUE

C       APPEL A LA SOUS-ROUTINE QUI CALCULE LES
C       COEFFICIENTS DU POLYNOME APPROXIMANT.

  100   CALL COEFF(NUMF,TOLER,AUX,NUMIT,ALPHA,ISGN,RO,IERR,KP)
        IF(IP.EQ.0) GOTO 10
        IF(NUMIT.EQ.1) GOTO 29
        IRAP=MOD(NUMIT,IP)
        IF(IRAP.NE.0) GOTO 10

C       IMPRESSION.

   29   PRINT 2004,NUMIT
        PRINT 2002
        DO 24 I=1,NAL
        K=IA(I)
   24   PRINT 2003,XR(K),YR(K)
        PRINT 2005
        PRINT 2006,(ISGN(I),I=1,NAL)
        PRINT 2007,RO
        GOTO(10,5,6,7,8),IERR
 2002   FORMAT(///,1X,'LES POINTS DE L''ALTERNANT SONT:')
 2003   FORMAT(37X,'( ',F6.2,' , ',F6.2,' )')
 2004   FORMAT(1H1,'ON EST A L''ITERATION NUMERO ',I2)
 2005   FORMAT(//,1X,'LA CONFIGURATION DES SIGNES EST:')
 2006   FORMAT(37X,'( ',4I4,' )')
 2007   FORMAT(//,1X,'L''ERREUR EGALISEE SUR ',
     C'L''ALTERNANT EST:',2X,F15.8)
    5   PRINT 2010
 2010   FORMAT(///,1X,'L''ALTERNANT DE LA PREMIERE ',
     C'ITERATION DONNE UN SYSTEME INSOLUBLE')
        RETURN
    6   PRINT 2011,NUMIT
 2011   FORMAT(///,1X,'LE SYSTEME CALCULANT ISGN A ',
     C'L''ITERATION',I3,' DANS COEFF EST INSOLUBLE')
        RETURN
```

```
    7 PRINT 2026
 2026 FORMAT(///,1X,'LA DERNIERE EQUATION DU SYSTEME DE ',
     C'COEFF EST INCOMPATIBLE AVEC LA SOLUTION TROUVEE')
      GOTO 10
    8 PRINT 2025
 2025 FORMAT(///,1X,'SYSTEME DE COEFF EST INSOLUBLE')
      RETURN
   10 CALL TIME(NOTIME)


C        APPEL A LA SOUS-ROUTINE QUI ECHANGE DEUX POINTS.


      CALL EXCH(NUMF,AUX,ALPHA,ISGN,RO,IOLD,INEW,KERR,KP,TOLER)


C        IMPRESSION.


      IF(KERR.EQ.0) GOTO 30
      GOTO(25,26),KERR
   26 PRINT 2027
 2027 FORMAT(///,1X,'CE CAS EST IMPOSSIBLE')
      RETURN
   25 PRINT 2014
      PRINT 2012
      DO 310 I=1,NFB
  310 IMP(I)=I-1
      PRINT 2013,(ALPHA(I),IMP(I),I=1,NFB)
 2014 FORMAT(//,1X,'CET   ALTERNANT DONNE LA MEILLEURE',
     C' APPROXIMATION SUR LE DOMAINE (A,B)*(C,D) ')
 2012 FORMAT(/,1X,'LE POLYNOME DE MEILLEURE APPROXIMATION EST:')
 2013 FORMAT(1X,'P(X,Y)= ',6(F16.8,2X,'X**',I2,' + '))
      RETURN


C        PASSAGE A L'ITERATION SUIVANTE.


   30 NUMIT=NUMIT+1


C        L'ETAPE SUIVANTE EST LE CALCUL DES COEFFICIENTS.


      IF(NUMIT.LT.MAXIT) GOTO 100
      PRINT 2030,MAXIT
 2030 FORMAT(///,1X,'LE CALCUL DE LA MEILLEURE APPROXIMATION DEMANDE '
     C'PUS DE',I4,' ITERATIONS.')
      RETURN
      END
```

```
      SUBROUTINE COEFF(NUMF,TOLER,AUX,NUMIT,ALPHA,ISGN,RO,IERR,KP)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
      DIMENSION AUX(10,10),ALPHA(10),ISGN(10)
      DIMENSION A(10,10),B(10)
      IERR=1

C       RESOLUTION PAR DIAGONALISATION DE GAUSS DU
C       SYSTEME D'EQUATIONS DE LA FORMULATION DUALE (1.6).

      DO 1 I=1,NFB
      B(I)=-AUX(I,NAL)
      DO 2 J=1,NFB
    2 A(I,J)=AUX(I,J)
    1 CONTINUE
      CALL GAUSS(A,B,NFB,1.D-13,JERR)
      IF(JERR.EQ.0) GOTO 5
      IF(NUMIT.NE.1) GOTO 4
      IERR=2
      RETURN
    4 IERR=3
      RETURN
    5 B(NAL)=1.D00
      S=1.D00
      DO 6 I=1,NFB
    6 S=S+DABS(B(I))

C       CALCUL DE RO ET ISGN.

      K=IA(NAL)
      RO=F(XR(K),YR(K),NUMF)
      DO 11 I=1,NFB
      k=IA(I)
   11 RO=RO+B(I)*F(XR(K),YR(K),NUMF)
      IF(RO.LT.0.D00) GOTO 15
      RO=RO/S
      IF(NUMIT.GT.1) GOTO 19
      KP=0
      DO 14 I=1,NAL
      IF(B(I)) 13,14,12
   12 ISGN(I)=1
      KP=KP+1
      GOTO 14
   13 ISGN(I)=-1
   14 CONTINUE
      GOTO 19
```

```
15 RO=-RO/S
   IF(NUMIT.GT.1) GOTO 19
   KP=0
   DO 18 I=1,NAL
   IF(B(I)) 17,18,16
16 ISGN(I)=-1
   GOTO 18
17 ISGN(I)=+1
   KP=KP+1
18 CONTINUE


C        CALCUL DES COEFFICIENTS DU POLYNOME APPROXIMANT
C        PAR LA RESOLUTION D'UN SYSTEME DE NAL EQUATIONS
C        A NFB INCONNUES.


19 DO 20 I=1,NFB
   DO 21 J=1,NFB
21 A(I,J)=AUX(J,I)
   K=IA(I)
20 ALPHA(I)=F(XR(K),YR(K),NUMF)-RO*ISGN(I)
   CALL GAUSS(A,ALPHA,NFB,1,D-13,JERR)
   IF(JERR.NE.0) GOTO 25


C        VERIFICATION DE LA COMPATIBILITE DE LA DERNIERE
C        EQUATION AVEC LA SOLUTION DU SYSTEME.
C        LA THEORIE AFFIRME CETTE COMPATIBILITE.


   S=0.D00
   DO 3 I=1,NFB
 3 S=S+ALPHA(I)*AUX(I,NAL)
   K=IA(NAL)
   VALUE=F(XR(K),YR(K),NUMF)
   IF(DABS(S+ISGN(NAL)*RO-VALUE).LT.TOLER) RETURN
   IERR=4
   RETURN
25 IERR=5
   RETURN
   END
```

```
      SUBROUTINE EXCH(NUMF,AUX,ALPHA,ISGN,RO,IOLD,INEW,KERR,KP,TOLER)


C       DECLARATIONS.


      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
      DIMENSION AUX(10,10),ALPHA(10),ISGN(10)
      LOGICAL DG


      KERR=0
      K=KP-2


C       CALCUL DU POINT D'ERREUR MAXIMALE SUR LA
C       REFERENCE,ON NOTE INEW L'INDICE DE CE POINT.


      INEW=1
      ERR=E(XR(1),YR(1),ALPHA,NUMF)
      DO 1 J=2,NR
      EAUX=E(XR(J),YR(J),ALPHA,NUMF)
      IF(DABS(EAUX).LT.DABS(ERR)) GOTO 1
      ERR=EAUX
      INEW=J
    1 CONTINUE
      IF(DABS(RO-DABS(ERR)).LE.1.D-08) GOTO 8
      IF(ERR.LT.0.D00) GOTO 2
      ISGNO=1
      IF(K) 4,10,3
    2 ISGNO=-1
      IF(K) 3,10,4


C       TROIS POINTS DE L'ALTERNANT ONT UN ERREUR DE
C       MEME SIGNE, LE POINT QUI ENTRE A UNE ERREUR DE
C       CE SIGNE.


    3 CALL EXCHA0(INEW,ISGN,DG,IOLD,ISGNO)
      IF(DG) GOTO 11
      CALL ECHA(INEW,IOLD,AUX,ISGN,ISGNO,KP)
      RETURN


C       TROIS POINTS DE L'ALTERNANT ONT UNE ERREUR DE
C       MEME SIGNE, LE POINT QUI ENTRE A UNE ERREUR DE
C       SIGNE OPPOSE


    4 CALL EXCHA1(INEW,ISGN,DG,IOLD,ISGNO)
      IF(DG) GOTO 11
      CALL ECHA(INEW,IOLD,AUX,ISGN,ISGNO,KP)
      RETURN
```

```
C       DEUX POINTS DE L'ALTERNANT ONT UNE ERREUR DE
C       SIGNE PLUS, LES DEUX AUTRES POINTS ONT UN ERREUR
C       DE SIGNE MOINS.


   10 CALL EXCHA2(INEW,ISGN,DG,IOLD,ISGNO)
      IF(DG) GOTO 11
      CALL ECHA(INEW,IOLD,AUX,ISGN,ISGNO,KP)
      RETURN
    8 KERR=1
      RETURN
   11 PRINT 900
  900 FORMAT(///,1X,'ON NE PEUT ECHANGER LE POINT D''ERREUR MAX',///)
      RETURN
      END




      SUBROUTINE ECHA(INEW,IOLD,AUX,ISGN,ISGNO,KP)


C       DECLARATIONS.


      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
      DIMENSION AUX(10,10),ISGN(10)


C       ECHANGE PROPREMENT DIT ENTRE LES DEUX POINTS.
C       MODIFICATION DE LA MATRICE AUX ET DU VECTEUR DES SIGNES ISGN.

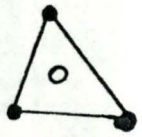

      IA(IOLD)=INEW
      DO 1 I=1,NFB
    1 AUX(I,IOLD)=FB(I,XR(INEW),YR(INEW))
      IF(ISGN(IOLD).NE.ISGNO) KP=KP+ISGNO
      ISGN(IOLD)=ISGNO
      RETURN
      END
```

```
      SUBROUTINE EXCHAO(INEW,ISGN,DG,IOLD,ISGNO)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
      DIMENSION ISGN(10),N(10),NI(10)
      LOGICAL DG
      INTEGER B

C       LA CONFIGURATION EST DU TYPE
```



```
C       ET LE POINT QUI ENTRE EST UN POINT ●
C       ON CALCULE LE POINT QUI SORT DE L'ENSEMBLE-H

      DG=.FALSE.
      II=1
      DO 1 I=1,NAL
      IF(ISGN(I).EQ.ISGNO) GOTO 2
      B=IA(I)
      GOTO 1
    2 N(II)=IA(I)
      NI(II)=I
      II=II+1
    1 CONTINUE
      DIV=D(INEW,B,N(1))/D(N(2),B,N(1))
      IF(DIV) 5,4,3
    3 DIV=D(INEW,B,N(3))/D(N(2),B,N(3))
      IF(DIV) 6,7,8
    4 DIV=D(INEW,B,N(3))/D(N(2),B,N(3))
      IF(DIV) 6,7,7
    5 DIV=D(INEW,B,N(2))/D(N(3),B,N(2))
      IF(DIV) 6,7,9
    6 IOLD=NI(1)
      RETURN
    7 DG=.TRUE.
      RETURN
    8 IOLD=NI(2)
      RETURN
    9 IOLD=NI(3)
      RETURN
      END
```

```
      SUBROUTINE EXCHA1(INEW,ISGN,DG,IOLD,ISGNO)


C       DECLARATIONS.


      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
      DIMENSION ISGN(10),N(10),NI(10)
      LOGICAL DG
      INTEGER B,BI


C       LA CONFIGURATION EST DU TYPE
```



```
C       ET LE POINT QUI ENTRE EST UN POINT o
C       ON CALCULE LE POINT QUI SORT DE L'ENSEMBLE-H.

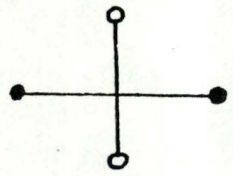

      DG=.FALSE.
      II=1
      DO 1 I=1,NAL
      IF(ISGN(I).NE.ISGNO) GOTO 2
      B=IA(I)
      BI=I
      GOTO 1
    2 N(II)=IA(I)
      NI(II)=I
      II=II+1
    1 CONTINUE
      DIV=D(INEW,N(1),N(3))/D(B,N(1),N(3))
      IF(DIV) 15,10,5
    5 DIV=D(INEW,N(1),N(2))/D(B,N(1),N(2))
      IF(DIV) 6,7,8
    6 DIV=D(INEW,B,N(2))/D(N(1),B,N(2))
      IF(DIV) 21,20,23
    7 DIV=D(INEW,N(2),N(3))/D(B,N(2),N(3))
      IF(DIV) 21,20,20
    8 DIV=D(INEW,N(2),N(3))/D(B,N(2),N(3))
      IF(DIV) 21,20,24
   10 DIV=D(INEW,B,N(3))/D(N(2),B,N(3))
      IF(DIV) 11,20,21
   11 DIV=D(INEW,B,N(1))/D(N(3),B,N(1))
      IF(DIV) 23,20,20
   15 DIV=D(INEW,B,N(3))/D(N(2),B,N(3))
      IF(DIV) 16,20,21
   16 DIV=D(INEW,B,N(1))/D(N(3),B,N(1))
      IF(DIV) 23,20,22
   20 DG=.TRUE.
      RETURN
   21 IOLD=NI(1)
      RETURN
   22 IOLD=NI(2)
      RETURN
   23 IOLD=NI(3)
      RETURN
   24 IOLD=BI
      RETURN
      END
```

```
      SUBROUTINE EXCHA2(INEW,ISGN,DG,IOLD,ISGNO)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
      DIMENSION ISGN(10),NI(10),N(10),B(10),BI(10)
      LOGICAL DG
      INTEGER B,BI

C       LA CONFIGURATION EST DU TYPE
```



```
C       ET LE POINT QUI ENTRE EST UN POINT ●
C       ON CALCULE LE POINT QUI SORT DE L'ENSEMBLE-H.

      DG=.FALSE.
      JJ=1
      II=1
      DO 1 I=1,NAL
      IF(ISGN(I).NE.ISGNO) GOTO 2
      N(II)=IA(I)
      NI(II)=I
      II=II+1
      GOTO 1
    2 B(JJ)=IA(I)
      BI(JJ)=I
      JJ=JJ+1
    1 CONTINUE
      DIV=D(INEW,B(1),B(2))/D(N(1),B(1),B(2))
      IF(DIV) 5,10,15
    5 DIV=D(INEW,N(1),B(2))/D(N(2),N(1),B(2))
      IF(DIV) 21,20,6
    6 DIV=D(INEW,N(1),B(1))/D(N(2),N(1),B(1))
      IF(DIV) 22,20,32
   10 DIV=D(INEW,N(2),B(2))/D(N(1),N(2),B(2))
      IF(DIV) 21,20,11
   11 DIV=D(INEW,N(2),B(1))/D(N(1),N(2),B(1))
      IF(DIV) 22,20,20
   15 DIV=D(INEW,N(2),B(2))/D(N(1),N(2),B(2))
      IF(DIV) 21,20,16
   16 DIV=D(INEW,N(2),B(1))/D(N(1),N(2),B(1))
      IF(DIV) 22,20,31
   20 DG=.TRUE.
      RETURN
   21 IOLD=BI(1)
      RETURN
   22 IOLD=BI(2)
      RETURN
   31 IOLD=NI(1)
      RETURN
   32 IOLD=NI(2)
      RETURN
      END
```

```
      DOUBLE PRECISION FUNCTION D(P,L,M)

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON XR(11000),YR(11000),IA(10),NR,NAL,NFB,NF
      INTEGER P

C       FONCTION QUI CALCULE LA VALEUR DE LA DROITE
C       PASSANT PAR (XR(L),YR(L)) ET (XR(M),YR(M)) EN
C       (XR(P),YR(P)).

      D=(XR(P)-XR(L))*(YR(L)-YR(M))-(YR(P)-YR(L))*(XR(L)-XR(M))
      RETURN
      END

      DOUBLE PRECISION FUNCTION FB(I,X,Y)

      IMPLICIT REAL*8 (A-H,O-Z)

C       FONCTIONS DE BASE DE L'APPROXIMATION POLYNOMIALE
C       DE DEGRE 1 DANS RXR

      GOTO(1,2,3),I
    1 FB=1.D00
      RETURN
    2 FB=X
      RETURN
    3 FB=Y
      RETURN
      END
```

```
      PROGRAM HSETS
      SUBROUTINE REMEZ

C
C     TABLES DE VARIABLES
C     - - - - - - - - - -
C
C     REELS
C     - - -
C
C        AUX    : MATRICE ASSOCIEE A L'HOMOMORPHISME U
C        A      : MATRICE AUXILIAIRE
C        ALPHA  : VECTEUR DES COEFFICIENTS DU POLYNOME
C        B      : VECTEUR SECOND MEMBRE A L'ENTREE DE GAUSS
C                 VECTEUR SOLUTION A LA SORTIE
C        ERM    : VECTEUR DE L'ERREUR NEGATIVE MAXIMALE EN
C                 VALEUR ABSOLUE ET > RO DANS CHACUN DES 4
C                 SOUS-ENSEMBLES DU DOMAINE D'APPROXIMATION
C        ERP    : VECTEUR DE L'ERREUR POSITIVE MAXIMALE ET > RO
C                 DANS CHACUN DE CES 4 SOUS-ENSEMBLES
C        RO     : ERREUR EGALISEE SUR L'ALTERNANT
C
C     ENTIERS
C     - - - - -
C
C        IMP    : VECTEUR SERVANT A L'IMPRESSION
C        ISGNO  : SIGNE DE L'ERREUR AU POINT QUI ENTRE
C                 DANS L'ALTERNANT
C        ISGN   : VECTEUR DES SIGNES DE L'ERREUR AUX
C                 POINTS DE L'ALTERNANT
C        INEW   : INDICE DANS LA REFERENCE DU POINT
C                 QUI ENTRE DANS L'ALTERNANT
C        IOLD   : INDICE DANS L'ALTERNANT DU POINT
C                 QUI SORT DE L'ALTERNANT
C        IM     : VECTEUR DES INDICES DES POINTS DE L'ALTERNANT
C                 OU L'ERREUR EST NEGATIVE
C        IP     : VECTEUR DES INDICES DES POINTS DE L'ALTERNANT
C                 OU L'ERREUR EST POSITIVE
C        NUMIT  : NUMERO DE L'ITERATION EN COURS
C        NUMF   : NUMERO DU PROBLEME TRAITE
C        NMP    : NOMBRE DE POINTS PRESELECTIONNES
C        NM     : NOMBRE DE POINTS PRESELECTIONNES OU L'ERREUR
C                 EST NEGATIVE
C        NP     : NOMBRE DE POINTS PRESELECTIONNES OU L'ERREUR
C                 EST POSITIVE
C        NQ     : NUMERO DU QUADRANT
C        MAL    : NOMBRE DE POINTS OU L'ERREUR EST NEGATIVE
C                 DANS L'ALTERNANT
C        PAL    : NOMBRE DE POINTS OU L'ERREUR EST POSITIVE
C                 DANS L'ALTERNANT
C        M      : VECTEUR DES INDICES DES POINTS PRESELECTIONNES
C                 OU L'ERREUR EST NEGATIVE
C        P      : VECTEUR DES INDICES DES POINTS PRESELECTIONNES
C                 OU L'ERREUR EST POSITIVE
```

```
C         FONCTIONS
C         - - - - -
C
C           E       : FONCTION ERREUR
C           D       : EQUATION DE LA DROITE PASSANT PAR DEUX POINTS
C           FB      : FONCTIONS DE BASE
C           F       : FONCTION A APPROXIMER

C         LES VARIABLES QUI NE SONT PAS REPRISES DANS
C         CETTE TABLE SONT DES VARIABLES INTERMEDIAIRES

C         DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /REF/XR(1100),YR(1100),NR
      COMMON /ALT/IA(10),NAL,NFB
      COMMON /BORN/XMAX,XMIN,YMAX,YMIN
      COMMON /PLUS/IP(10),PAL
      COMMON /MOINS/IM(10),MAL
      COMMON /RESULT/ISGN(10),ALPHA(10),PO
      COMMON /DONNEE/MAXIT,IMP,TOLER,NUMF
      COMMON /INTERM/AUX(10,10),KP
      DIMENSION A(10,10)
      NUMIT=1
      CALL TIME(NOTIME)

C         CONSTRUCTION DE LA MATRICE AUX.

      DO 1 I=1,NFB
      DO 2 J=1,NAL
      K=IA(J)
    2 AUX(I,J)=FB(I,XR(K),YR(K))
    1 CONTINUE

C         APPEL A LA SOUS-ROUTINE QUI CALCULE LES
C         COEFFICIENTS DU POLYNOME APPROXIMANT.

  100 CALL COEFF(NUMIT,IERR)

C         IMPRESSION.

      IF(IMP.EQ.0) GOTO 10
      IF(NUMIT.EQ.1) GOTO 29
      IRAP=MOD(NUMIT,IMP)
      IF(IRAP.NE.0) GOTO 10
```

```
   29 PRINT 2004,NUMIT
      PRINT 2002
      DO 24 I=1,NAL
      K=IA(I)
   24 PRINT 2003,XR(K),YR(K)
      PRINT 2005
      PRINT 2006,(ISGN(I),I=1,NAL)
      PRINT 2007,RQ
      GOTO(4,5,6,7,8),IERR
    4 PRINT 2008
      PRINT 2009,(ALPHA(I),I=1,NFB)
 2002 FORMAT(///,1X,'LES POINTS DE L''ALTERNANT SONT:')
 2003 FORMAT(37X,'( ',F6.2,' , ',F6.2,' )')
 2004 FORMAT(1H1,'ON EST A L''ITERATION NUMERO ',I2)
 2005 FORMAT(//,1X,'LA CONFIGURATION DES SIGNES EST:')
 2006 FORMAT(37X,'( ',4I4,' )')
 2007 FORMAT(//,1X,'L''ERREUR EGALISEE SUR ',
     C'L''ALTERNANT EST:',2X,F15.8)
 2008 FORMAT(//,1X,'LES COEFFICIENTS DU POLYNOME APPROXIMANT SONT:')
 2009 FORMAT(37X,5(F11.8,5X))
      GOTO 10
    5 PRINT 2010
 2010 FORMAT(///,1X,'L''ALTERNANT DE LA PREMIERE ',
     C'ITERATION DONNE UN SYSTEME INSOLUBLE')
      RETURN
    6 PRINT 2011,NUMIT
 2011 FORMAT(///,1X,'LE SYSTEME CALCULANT ISGN A ',
     C'L''ITERATION',I3,' DANS COEFF EST INSOLUBLE')
      RETURN
    7 PRINT 2026
 2026 FORMAT(///,1X,'LA DERNIERE EQUATION DU SYSTEME DE ',
     C'COEFF EST INCOMPATIBLE AVEC LA SOLUTION TROUVEE')
      GOTO 10
    8 PRINT 2025
 2025 FORMAT(///,1X,'SYSTEME DE COEFF EST INSOLUBLE')
      RETURN


C        APPEL A LA SOUS-ROUTINE QUI ECHANGE LES POINTS.


   10 CALL TIME(NOTIME)
      CALL GLOB(IERR)


C        IMPRESSION.

      IF(IERR.EQ.0) GOTO 25
      GOTO(31,32,33,34),IERR
   34 PRINT 3000
 3000 FORMAT(///,1X,'ON A ECHANGE 4 PONTS')
      GOTO 30
   33 PRINT 3001
 3001 FORMAT(///,1X,'ON A ECHANGE 3 POINTS')
      GOTO 30
```

```
   32 PRINT 3002
 3002 FORMAT(///,1X,'ON A ECHANGE 2 POINTS')
      GOTO 30
   31 PRINT 3003
 3003 FORMAT(///,1X,'ON A ECHANGE 1 POINT')
      GOTO 30
   25 PRINT 2004,NUMIT
      PRINT 2027
 2027 FORMAT(//,1X,'CET ALTERNANT DONNE LA MEILLEURE',
     C' APPROXIMATION SUR LE DOMAINE (A,B)*(C,D) ')
      RETURN

C        PASSAGE A L'ITERATION SUIVANTE.

   30 NUMIT=NUMIT+1

C        L'ETAPE SUIVANTE EST LE CALCUL DES COEFFICIENTS.

      IF(NUMIT.LT.MAXIT) GOTO 100
      PRINT 2030,MAXIT
 2030 FORMAT(///,1X,'LE CALCUL DE LA MEILLEURE APPROXIMATION DEMANDE ',
     C'PLUS DE',I4,' ITERATIONS.')
      RETURN
      END
```

```
      SUBROUTINE COEFF(NUMIT,IERR)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /REF/XR(11000),YR(11000),NR
      COMMON /ALT/IA(10),NAL,NFB
      COMMON /PLUS/IP(10),PAL
      COMMON /MOINS/IM(10),MAL
      COMMON /RESULT/ISGN(10),ALPHA(10),RO
      COMMON /DONNEE/MAXIT,IMP,TOLER,NUMF
      COMMON /INTERM/AUX(10,10),KP
      DIMENSION A(10,10),B(10)

      IERR=1
      KM=0
      KP=0

C       RESOLUTION PAR DIAGONALISATION DE GAUSS DU
C       SYSTEME D'EQUATIONS DE LA FORMULATION DUALE (1.6).

      DO 1 I=1,NFB
      B(I)=-AUX(I,NAL)
      DO 2 J=1,NFB
    2 A(I,J)=AUX(I,J)
    1 CONTINUE
      CALL GAUSS(A,B,NFB,1.D-13,JERR)
      IF(JERR.EQ.0) GOTO 5

C     CALCUL DE RO ET ISGN.

C       CALCUL DE KP,KM,IP,IM.

      IF(NUMIT.NE.1) GOTO 4
      IERR=2
      RETURN
    4 IERR=3
      RETURN
    5 B(NAL)=1.D00
      S=1.D00
      DO 6 I=1,NFB
    6 S=S+DABS(B(I))
C     CALCULATION OF RO AND ISGN
      K=IA(NAL)
      RO=F(XR(K),YR(K),NUMF)
      DO 11 I=1,NFB
      K=IA(I)
   11 RO=RO+B(I)*F(XR(K),YR(K),NUMF)
```

```
      IF(RO.LT.0.D00) GOTO 15
      RO=RO/S
      DO 14 I=1,NAL
      IF(B(I)) 13,14,12
   12 ISGN(I)=1
      KP=KP+1
      IP(KP)=IA(I)
      GOTO 14
   13 ISGN(I)=-1
      KM=KM+1
      IM(KM)=IA(I)
   14 CONTINUE
      GOTO 19
   15 RO=-RO/S
      DO 18 I=1,NAL
      IF(B(I)) 17,18,16
   16 ISGN(I)=-1
      KM=KM+1
      IM(KM)=IA(I)
      GOTO 18
   17 ISGN(I)=+1
      KP=KP+1
      IP(KP)=IA(I)
   18 CONTINUE


C        CALCUL DES COEFFICIENTS DU POLYNOME APPROXIMANT
C        PAR LA RESOLUTION D'UN SYSTEME DE NAL EQUATIONS
C        A NFB INCONNES.


   19 DO 20 I=1,NFB
      DO 21 J=1,NFB
   21 A(I,J)=AUX(J,I)
      K=IA(I)
   20 ALPHA(I)=F(XR(K),YR(K),NUMF)-RO*ISGN(I)
      CALL GAUSS(A,ALPHA,NFB,1.D-13,JERR)
      IF(JERR.NE.0) GOTO 25


C        VERIFICATION DE LA COMPATIBILITE DE LA DERNIERE
C        EQUATION AVEC LA SOLUTION DU SYSTEME.
C        LA THEORIE AFFIRME CETTE COMPATIBILITE.


      S=0.D00
      DO 3 I=1,NFB
    3 S=S+ALPHA(I)*AUX(I,NAL)
      K=IA(NAL)
      VALUE=F(XR(K),YR(K),NUMF)
      IF(DABS(S+ISGN(NAL)*RO-VALUE).LT.TOLER) RETURN
      IERR=4
      RETURN
   25 IERR=5
      RETURN
      END
```

```
      SUBROUTINE GLOB(IERR)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /REF/XR(11000),YR(11000),NR
      COMMON /ALT/IA(10),NAL,NPB
      COMMON /BORN/XMAX,XMIN,YMAX,YMIN
      COMMON /PLUS/IP(10),PAL
      COMMON /MOINS/IM(10),MAL
      COMMON /RESULT/ISGN(10),ALPHA(10),RO
      COMMON /INTERM/AUX(10,10),KP
      DIMENSION ERM(10),ERP(10)
      DIMENSION M(10)
      INTEGER P(10)
      INTEGER PAL

C       INITIALISATION.

      NM=0
      NP=0
      DO 2 I=1,4
      M(I)=0
      P(I)=0
      ERM(I)=0.D00
      ERP(I)=0.D0C
    2 CONTINUE
      PAL=KP
      MAL=NAL-KP

C       CALCUL DES POINTS D'ERREUR MAXIMALE DANS LES QUADRANTS.

      XMID=(XMAX-XMIN)/2
      YMID=(YMAX-YMIN)/2
      EAUX=RO
      DO 1 I=1,NR
      ER=E(XR(I),YR(I))
      IF((DABS(ER)-RO).LE.1.D-05) GOTO 1

C       DETERMINATION DU QUADRANT
```

```
      IF(XR(I).LT.XMID) GOTO 70
      IF(YR(I).GT.YMID) GOTO 74
      NQ=2
      GOTO 75
   70 IF(YR(I).GT.YMID) GOTO 73
      NQ=1
      GOTO 75
   73 NQ=3
      GOTO 75
   74 NQ=4

C        DETERMINATION DU MAX DE L'ERREUR - ET + DANS NQ

C        DETERMINATION DU POINT D'ERREUR MAXIMALE SUR LA REFERENCE.

   75 IF(ER.GT.0.000) GOTO 76
      IF(ER.GT.ERM(NQ)) GOTO 1
      ERM(NQ)=ER
      M(NQ)=I
      IF(DABS(ER).LE.EAUX) GOTO 1
      ISGNQ=-1
      GOTO 79
   76 IF(ER.LT.ERP(NQ)) GOTO 1
      ERP(NQ)=ER
      P(NQ)=I
      IF(ER.LE.EAUX) GOTO 1
      ISGNQ=+1
   79 EAUX=DABS(ER)
      INEW=I
    1 CONTINUE

C        ARRIVE A CE STADE, ON A LE POINT DU QUADRANT I OU
C        L'ERREUR POSITIVE EST MAXIMALE DANS P(I) ET LE POINT
C        OU L'ERREUR NEGATIVE EST,EN VALEUR ABSOLUE, MAXIMALE
C        DANS M(I)

C        LES QUADRANT N'AYANT AUCUNE IMPORTANCE PAR APRES, ON
C        ORDONNE M ET P POUR NE PLUS AVOIR DE COMPOSANTE NULLE
C        AU CAS OU IL Y EN AURAIT EU


      II=0
      JJ=0
      DO 77 I=1,4
      IF(M(I).EQ.0) GOTO 78
      II=II+1
      M(II)=M(I)
```

```
      NM=NM+1
   78 IF(P(I).EQ.0) GOTO 77
      JJ=JJ+1
      P(JJ)=P(I)
      NP=NP+1
   77 CONTINUE
      NMP=NM+NP

C        ON A DONC NMP POINTS DONT NM POINTS NEGATIFS ET NP POINTS
C        POSITIFS.

C        CHOIX DES ECHANGES A ENVISAGER SUIVANT LES VALEURS
C        DE NMP,NM,NP.

      IF(NMP.EQ.0) GOTO 900
      IF(NMP.EQ.1) GOTO 100
      IF(NM.EQ.0) GOTO 600
      IF(NP.EQ.0) GOTO 500
      IF(NMP.GE.4) GOTO 400
      GOTO(100,200,300),NMP

C        NMP>3, NM>0,NP>0 , ON PEUT ESPERER UN ECHANGE DE 4 POINTS

C        1. EN CARRE

  400 DO 401 I1=1,NM
      I11=I1+1
      IF(I11.GT.NM) GOTO 401
      DO 402 I2=I11,NM
      DO 403 I3=1,NP
      I31=I3+1
      IF(I31.GT.NP) GOTO 403
      DO 404 I4=I31,NP
      CALL CARRE(JIERR,JERR,M(I1),M(I2),P(I3),P(I4))
      IF(JERR.EQ.0)  GOTO 407
      IERR=4
      RETURN

C        2. EN TRIANGLE

  407 I41=I4+1
      IF(I41.GT.NP) GOTO 404
      DO 405 I5=I41,NP
      CALL TRIGLE(IJERR,IERR,P(I3),P(I4),P(I5),M(I1))
```

```
      IF(IERR.EQ.0) GOTO 405
      IERR=4
      RETURN
  405 CONTINUE
  404 CONTINUE
      I21=I2+1
      IF(I21.GT.NM) GOTO 402
      DO 406 I6=I21,NM
      CALL TRIGLE(IJERR,IERR,M(I1),M(I2),M(I6),P(I3))
      IF(IERR.EQ.0) GOTO 406
      IERR=4
      RETURN
  406 CONTINUE
  403 CONTINUE
  402 CONTINUE
  401 CONTINUE


C         UN ECHANGE DE QUATRE DE CES POINTS EST IMPOSSIBLE.
C         ON ESSAYE D'ECHANGER TROIS POINTS.


      DO 411 I1=1,NM
      I11=I1+1
      IF(I11.GT.NM) GOTO 411
      DO 412 I2=I11,NM
      I21=I2+1
      IF(I21.GT.NM) GOTO 412
      DO 413 I3=I21,NM
      CALL E3EQSI(PAL,M(I1),M(I2),M(I3),IP,KERR)
      IF(KERR.EQ.1) GOTO 413
      IERR=3
      RETURN
  413 CONTINUE
      DO 414 I4=1,NP
      CALL E3NESI(MAL,PAL,M(I1),M(I2),P(I4),IP,IM,KERR)
      IF(KERR.EQ.1) GOTO 414
      IERR=3
      RETURN
  414 CONTINUE
  412 CONTINUE
      DO 415 I5=1,NP
      I51=I5+1
      IF(I51.GT.NP) GOTO 415
      DO 416 I6=I51,NP
      CALL E3NESI(PAL,MAL,P(I5),P(I6),M(I1),IM,IP,KERR)
      IF(KERR.EQ.1) GOTO 418
      IERR=3
      RETURN
  418 I61=I6+1
```

```
      IF(I61.GT.NP) GOTO 416
      DO 417 I7= I61,NP
      CALL E3EQSI(MAL,P(I5),P(I6),P(I7),IM,KERR)
      IF(KERR.EQ.1) GOTO 417
      IERR=3
      RETURN
  417 CONTINUE
  416 CONTINUE
  415 CONTINUE
  411 CONTINUE


C       UN ECHANGE DE TROIS DE CES POINTS EST IMPOSSIBLE.
C       ON ESSAYE UN ECHANGE DE DEUX POINTS


      DO 421 I1=1,NM
      I11=I1+1
      IF(I11.GT.NM) GOTO 421
      DO 422 I2=I11,NM
      CALL E2EQSI(M(I1),M(I2),PAL,MAL,IM,IP,KERR)
      IF(KERR.EQ.1) GOTO 422
      IERR=2
      RETURN
  422 CONTINUE
      DO 423 I3=1,NP
      CALL E2NESI(M(I1),P(I3),IP,IM,PAL,MAL,KERR)
      IF(KERR.EQ.1) GOTO 423
      IERR=2
      RETURN
  423 CONTINUE
  421 CONTINUE
      DO 424 I4=1,NP
      I41=I4+1
      IF(I41.GT.NP) GOTO 424
      DO 425 I5=I41,NP
      CALL E2EQSI(P(I4),P(I5),MAL,PAL,IP,IM,KERR)
      IF(KERR.EQ.1) GOTO 425
      IERR=2
      RETURN
  425 CONTINUE
  424 CONTINUE


C       UN ECHANGE DE DEUX DE CES POINTS EST IMPOSSIBLE
C       ON FAIT L'ECHANGE D'UN SEUL POINT.


      GOTO 100
```

```
C       NMP<5,NP=0, ON ESSAYE D'ECHANGER 3,2 OU 1 POINT.

  500 DO 501 I1=1,NM
      I11=I1+1
      IF(I11.GT.NM) GOTO 501
      DO 502 I2=I11,NM
      I21=I2+1
      IF(I21.GT.NM) GOTO 504
      DO 503 I3=I21,NM
      CALL E3EQSI(PAL,M(I1),M(I2),M(I3),IP,KERR)
      IF(KERR.EQ.1) GOTO 503
      IERR=3
      RETURN
  503 CONTINUE
  504 CALL E2EQSI(M(I1),M(I2),PAL,MAL,IM,IP,KERR)
      IF(KERR.EQ.1) GOTO 502
      IERR=2
      RETURN
  502 CONTINUE
  501 CONTINUE
      GOTO 100

C       NMP<5,NM=0, ON ESSAYE D'ECHANGER 3,2 OU 1 POINTS

  600 DO 601 I1=1,NP
      I11=I1+1
      IF(I11.GT.NP) GOTO 601
      DO 602 I2=I11,NP
      I21=I2+1
      IF(I21.GT.NP) GOTO 604
      DO 603 I3=I21,NP
      CALL E3EQSI(MAL,P(I1),P(I2),P(I3),IM,KERR)
      IF(KERR.EQ.1) GOTO 603
      IERR=3
      RETURN
  603 CONTINUE
  604 CALL E2EQSI(P(I1),P(I2),MAL,PAL,IP,IM,KERR)
      IF(KERR.EQ.1) GOTO 602
      IERR=2
      RETURN
  602 CONTINUE
  601 CONTINUE
      GOTO 100

C       NMP=3,NM=2,NP=1 OU NM=1,NP=2
```

```
C       ESSAI DE 3 POINTS

  300 IF(NM.GT.NP) GOTO 301
      CALL E3NESI(PAL,MAL,P(1),P(2),M(1),IM,IP,KERR)
      IF(KERR.EQ.1) GOTO 302
      IERR=3
      RETURN
  301 CALL E3NESI(MAL,PAL,M(1),M(2),P(1),IP,IM,KERR)
      IF(KERR.EQ.1) GOTO 304
      IERR=3
      RETURN

C       ESSAI DE DEUX POINTS

  302 CALL E2EQSI(P(1),P(2),MAL,PAL,IP,IM,KERR)
      IF(KERR.EQ.1) GOTO 303
      IERR=2
      RETURN
  303 CALL E2NESI(M(1),P(2),IP,IM,PAL,MAL,KERR)
      IF(KERR.EQ.1) GOTO 306
      IERR=2
      RETURN
  304 CALL E2NESI(M(2),P(1),IP,IM,PAL,MAL,KERR)
      IF(KERR.EQ.1) GOTO 305
      IERR=2
      RETURN
  305 CALL E2EQSI(M(1),M(2),PAL,MAL,IM,IP,KERR)
      IF(KERR.EQ.1) GOTO 306
      IERR=2
      RETURN
  306 CALL E2NESI(M(1),P(1),IP,IM,PAL,MAL,KERR)
      IF(KERR.EQ.1) GOTO 100
      IERR=2
      RETURN

C       NMP=2,NM=1,NP=1

  200 CALL E2NESI(M(1),P(1),IP,IM,PAL,MAL,KERR)
      IF(KERR.EQ.1) GOTO 100
      IERR=2
      RETURN

C       NMP=1

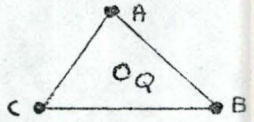
  100 CALL E1(IOLD,INEW,ISGNO)
      IERR=1
      RETURN
  900 IERR=0
      RETURN
      END
```

```
      SUBROUTINE TRIGLE(IJERR,IERR,A,B,C,Q)

C        DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /REF/XR(11000),YR(11000),NR
      COMMON /ALT/IA(10),NAL,NPB
      COMMON /RESULT/ISGN(10),ALPHA(10),RO
      COMMON /INTERM/AUX(10,10),KP
      INTEGER A,B,C,Q

C        SOUS-ROUTINE QUI VERIFIE SI LES QUATRES POINTS
C        A,B,C,Q ONT LA CONFIGURATION SUIVANTE:
```



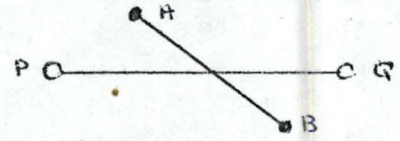

```
      IJERR=1
      IERR=1
      DIV=D(Q,A,B)
      IF(DABS(DIV).LT.1.D-08) GOTO 10
      DIV=D(C,A,B)/DIV
      IF(DIV) 11,11,1
    1 DIV=D(Q,B,C)
      IF(DABS(DIV).LT.1.D-08) GOTO 11
      DIV=D(A,B,C)/DIV
      IF(DIV) 11,11,2
    2 DIV=D(Q,A,C)
      IF(DABS(DIV).LT.1.D-08) GOTO 11
      DIV=D(B,A,C)/DIV
      IF(DIV) 11,11,3
    3 IA(1)=A
      IA(2)=B
      IA(3)=C
      IA(4)=Q
      CALL CONS(AUX)
      RETURN
   10 IJERR=0
   11 IERR=0
      RETURN
      END
```

```
      SUBROUTINE CARRE(JIERR,JERR,A,B,P,Q)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /REF/XR(10000),YR(10000),NR
      COMMON /ALT/IA(10),NAL,NFB
      COMMON /RESULT/ISGN(10),ALPHA(10),RO
      COMMON /INTERM/AUX(10,10),KP
      INTEGER A,B,P,Q

C       SOUS-ROUTINE QUI VERIFIE SI LES QUATRES POINTS
C       A,B,P,Q ONT LA CONFIGURATION SUIVANTE:
```




```
      JIERR=1
      JERR=1
      DIV=D(P,A,B)
      IF(DABS(DIV).GE.1.D-08) GOTO 1
      JIERR=0
      JERR=0
      RETURN
    1 DIV=D(Q,A,B)/DIV
      IF(DIV) 2,10,10
    2 DIV=D(A,P,Q)
      IF(DABS(DIV).LT.1.D-08) GOTO 10
      DIV=D(B,P,Q)/DIV
      IF(DIV) 4,10,10
    4 IA(1)=A
      IA(2)=B
      IA(3)=P
      IA(4)=Q
      CALL CONS(AUX)
      RETURN
   10 JERR=0
      RETURN
      END
```

```
      SUBROUTINE CONS(AUX)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /REF/XR(11000),YR(11000),NR
      COMMON /ALT/IA(10),NAL,NFB
      DIMENSION AUX(10,10)

C       CONSTRUCTION DE LA MATRICE AUX.

      DO 1 I=1,NFB
      DO 2 J=1,NAL
      K=IA(J)
    2 AUX(I,J)=FB(I,XR(K),YR(K))
    1 CONTINUE
      RETURN
      END


      SUBROUTINE E3EQSI(PAUX,LR1,LR2,LR3,IPAUX,KERR)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /REF/XR(11000),YR(11000),NR
      COMMON /ALT/IA(10),NAL,NFB
      COMMON /PLUS/IP(10),PAL
      COMMON /MOINS/IM(10),MAL
      COMMON /RESULT/ISGN(10),ALPHA(10),RO
      COMMON /INTERM/AUX(10,10),KP
      DIMENSION IPAUX(10)
      INTEGER PAUX

C       LES TROIS POINTS ONT DES ERREURS DE MEME SIGNE.
C       UNE SEULE CONFIGURATION POSSIBLE: LE TRIANGLE.

      KERR=0
      DO 1 I=1,PAUX
      CALL TRIGLE(IJERR,IERR,LR1,LR2,LR3,IPAUX(I))
      IF(IERR.EQ.0) GOTO 1
      RETURN
    1 CONTINUE
      KERR=1
      RETURN
      END
```

```
      SUBROUTINE ESNESI(MAUX,PAUX,LR1,LR2,KR1,IPAUX,IMAUX,KERR)

C       IMPRESSION.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /REF/XR(11000),YR(11000),NR
      COMMON /ALT/IA(10),NAL,NFB
      COMMON /PLUS/IP(10),PAL
      COMMON /MOINS/IM(10),MAL
      COMMON /RESULT/ISGN(10),ALPHA(10),RO
      COMMON /INTERM/ALX(10,10),KP
      DIMENSION IPAUX(10),IMAUX(10)
      INTEGER PAUX

C       LES TROIS POINTS ONT DES ERREURS DE SIGNES DIFFERENTS.

      KERR=0
      DO 1 J=1,MAUX
      CALL TRIGLE(IJERR,IERR,LR1,LR2,IMAUX(J),KR1)
      IF(IJERR.EQ.0) GOTO 3
      IF(IERR.EQ.0) GOTO 1
      RETURN
    1 CONTINUE
      DO 2 I=1,PAUX
      CALL CARRE(JIERR,JERR,LR1,LR2,KR1,IPAUX(I))
      IF(IERR.EQ.0) GOTO 2
      RETURN
    2 CONTINUE
    3 KERR=1
      RETURN
      END
```

```
      SUBROUTINE E2TYPE(LR1,KR1,KA1,IMAUX,PAUX,MAUX,KERR)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /REF/XR(11000),YR(11000),NR
      COMMON /ALT/IA(10),NAL,NFB
      COMMON /PLUS/IP(10),PAL
      COMMON /MOINS/IM(10),MAL
      COMMON /RESULT/ISGN(10),ALPHA(10),RO
      COMMON /INTERM/AUX(10,10),KP
      DIMENSION IMAUX(10)
      INTEGER PAUX

C       LES DEUX POINTS ENTRANT DANS L'ALTERNANT ONT DES
C       ERREURS DE SIGNES OPPOSES
C       LA CONFIGURATION DE L'ALTERNANT EST UN TRIANGLE.

      KERR=0
      DO 1 I=1,MAUX
      CALL CARRE(JIERR,JERR,KR1,KA1,LR1,IMAUX(I))
      IF(JIERR.EQ.0) GOTO 2
      IF(JERR.EQ.0) GOTO 1
      RETURN
    1 CONTINUE
    2 DO 3 I=2,3
      CALL TRIGLE(IJERR,IERR,LR1,IMAUX(1),IMAUX(I),KR1)
      IF(IJERR.EQ.0) GOTO 5
      IF(IERR.EQ.0) GOTO 3
      RETURN
    3 CONTINUE
    5 CALL TRIGLE(IJERR,IERR,LR1,IMAUX(2),IMAUX(3),KR1)
      IF(IERR.EQ.0) GOTO 4
      RETURN
    4 KERR=1
      RETURN
      END
```

```
      SUBROUTINE F2EQSI(LR1,LR2,PAUX,MAUX,IMAUX,IPAUX,KERR)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /REF/XR(11000),YR(11000),NR
      COMMON /ALT/IA(10),NAL,NFB
      COMMON /PLUS/IP(10),PAL
      COMMON /MOINS/IM(10),MAL
      COMMON /RESULT/ISGN(10),ALPHA(10),RO
      COMMON /INTERM/AUX(10,10),KP
      DIMENSION IMAUX(10),IPAUX(10)
      INTEGER PAUX

C       LES DEUX POINTS ENTRANT DANS L'ALTERNANT ONT DES
C       ERREURS DE MEME SIGNE.

      KERR=0
      DO 2 K=1,PAUX
      DO 1 I=1,MAUX
      CALL TRIGLE(IJERR,IERR,LR1,LR2,IMAUX(I),IPAUX(K))
      IF(IJERR.EQ.0) GOTO 2
      IF(IERR.EQ.0) GOTO 1
      RETURN
    1 CONTINUE
    2 CONTINUE
      IF(PAUX.EQ.1) GOTO 3
      MU=PAUX-1
      DO 4 I=1,MU
      II=I+1
      DO 5 K=II,PAUX
      CALL CARRE(JIERR,JERR,LR1,LR2,IPAUX(I),IPAUX(K))
      IF(JIERR.EQ.0) GOTO 4
      IF(JERR.EQ.0) GOTO 5
      RETURN
    5 CONTINUE
    4 CONTINUE
    3 KERR=1
      RETURN
      END
```

```
      SUBROUTINE E2NESI(LR1,KR1,IPAUX,IMAUX,PAUX,MAUX,KERR)

C       DECLARATIONS.

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /REF/XR(11000),YR(11000),NR
      COMMON /ALT/IA(10),NAL,NFB
      COMMON /PLUS/IP(10),PAL
      COMMON /MOINS/IM(10),MAL
      COMMON /RESULT/ISGN(10),ALPHA(10),RO
      COMMON /INTERM/AUX(10,10),KP
      DIMENSION IMAUX(10),IPAUX(10)
      INTEGER PAUX

C       LES POINTS ENTRANT DANS L'ALTERNANT ONT DES
C       ERREURS DE SIGNE OPPOSE.

      KERR=0
      IAUX=MAUX-2
      IF(IAUX) 1,2,3
    1 CALL E2TYPE(KR1,LR1,IMAUX(1),IPAUX,MAUX,PAUX,KERR)
      RETURN
    2 CALL TRIGLE(IJERR,IERR,KR1,IPAUX(1),IPAUX(2),LR1)
      IF(IERR.EQ.0) GOTO 10
      RETURN
   10 CALL TRIGLE(IJERR,IERR,IMAUX(1),IMAUX(2),LR1,KR1)
      IF(IERR.EQ.0) GOTO 12
      RETURN
   12 DO 13 K=1,2
      DO 14 I=1,2
      CALL CARRE(JIERR,JERR,LR1,IMAUX(K),KR1,IPAUX(I))
      IF(JIERR.EQ.0) GOTO 13
      IF(JERR.EQ.0) GOTO 14
      RETURN
   14 CONTINUE
   13 CONTINUE
      KERR=1
      RETURN
    3 CALL E2TYPE(LR1,KR1,IPAUX(1),IMAUX,PAUX,MAUX,KERR)
      RETURN
      END
```

```
      SUBRCUTINE E1(IOLD,INEW,ISGNO)

C       DECLARATIONS

      IMPLICIT REAL*8 (A-H,O-Z)
      COMMON /REF/XR(11000),YR(11000),NR
      COMMON /ALT/IA(10),NAL,NFB
      COMMON /RESULT/ISGN(10),ALPHA(10),RO
      COMMON /DONNEE/MAXIT,IMP,TOLER,NUMF
      COMMON /INTERM/AUX(10,10),KP
      LOGICAL DG

C       ECHANGE DU POINT D'ERREUR MAXIMALE EN VALEUR ABSOLUE.

      K=KP-2
      IF(ISGNO.NE.1) GOTO 2
      IF(K) 4,10,3
    2 IF(K) 3,10,4
    3 CALL EXCHAO(INEW,ISGN,DG,IOLD,ISGNO)
      IF(DG) GOTO 11
      CALL ECHA(INEW,IOLD,AUX,ISGN,ISGNO,KP)
      RETURN
    4 CALL EXCHA1(INEW,ISGN,DG,IOLD,ISGNO)
      IF(DG) GOTO 11
      CALL ECHA(INEW,IOLD,AUX,ISGN,ISGNO,KP)
      RETURN
   10 CALL EXCHA2(INEW,ISGN,DG,IOLD,ISGNO)
      IF(DG) GOTO 11
      CALL ECHA(INEW,IOLD,AUX,ISGN,ISGNO,KP)
      RETURN
   11 PRINT 900
  900 FORMAT(///,1X,"ON NO PEUT ECHANGER LE POINT D''ERREUR MAX',///)
      STOP
      END
```

Annexe 2 : Figures.

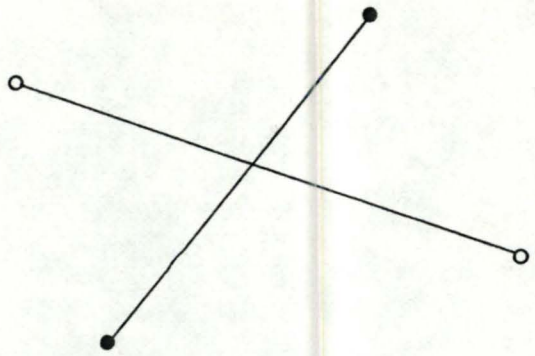


fig. 1.1. a.

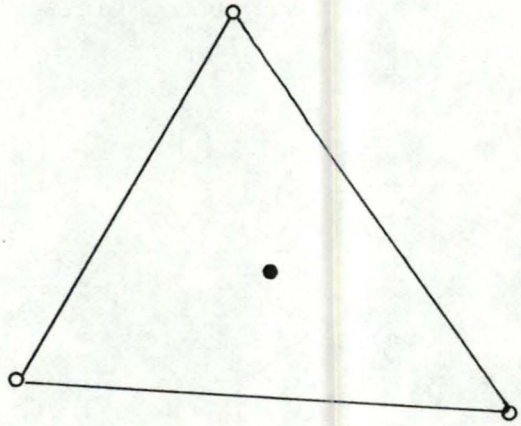


fig. 1.1. b.

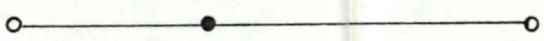


fig. 1.2.

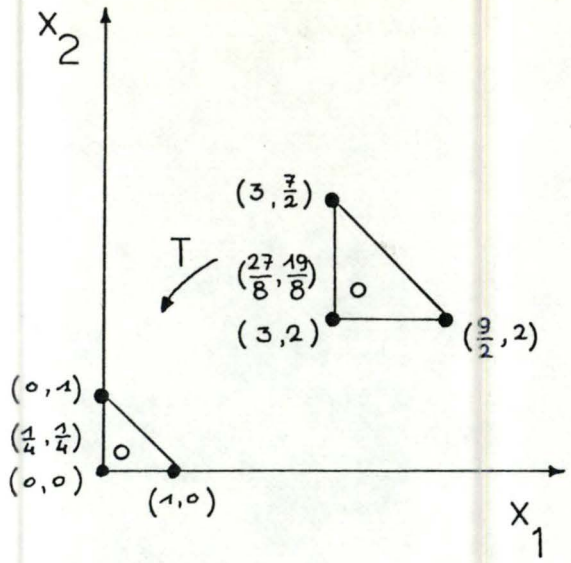


fig. 2.1.

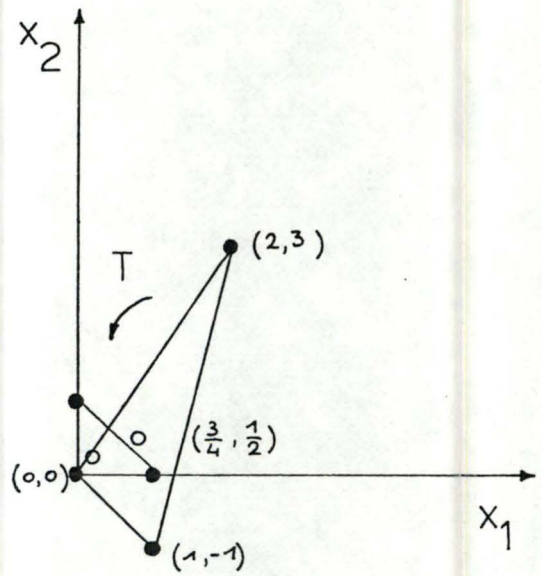


fig. 2.2.

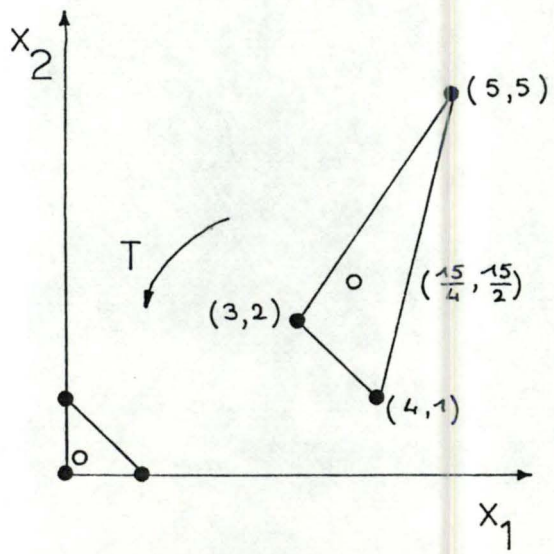


fig. 2.3.

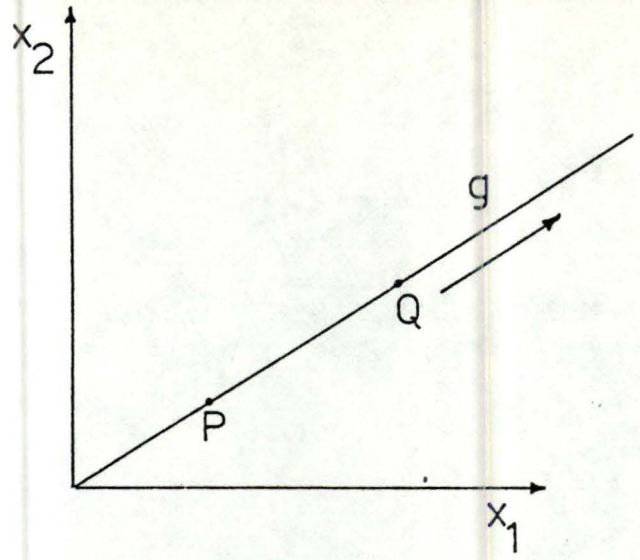


fig. 2.4. n= 2

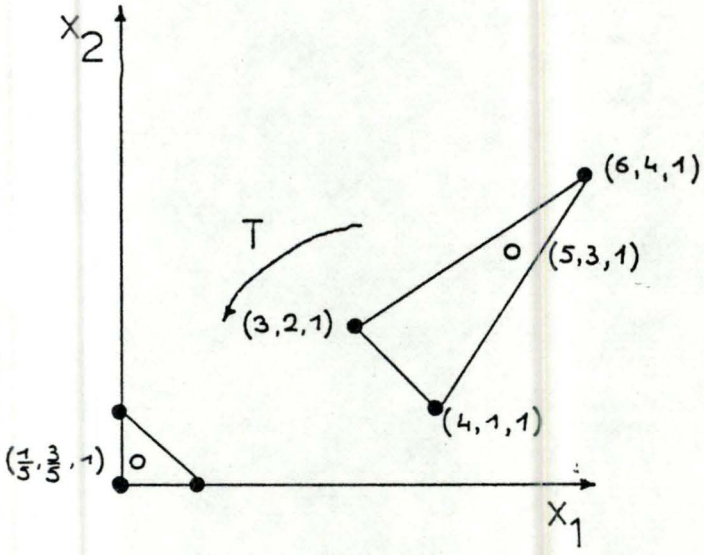


fig. 2.5.

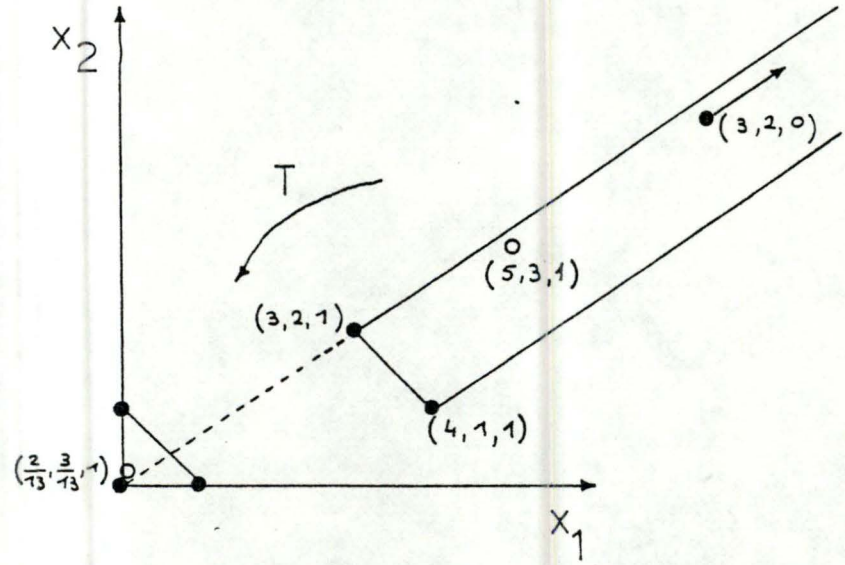


fig. 2.6.

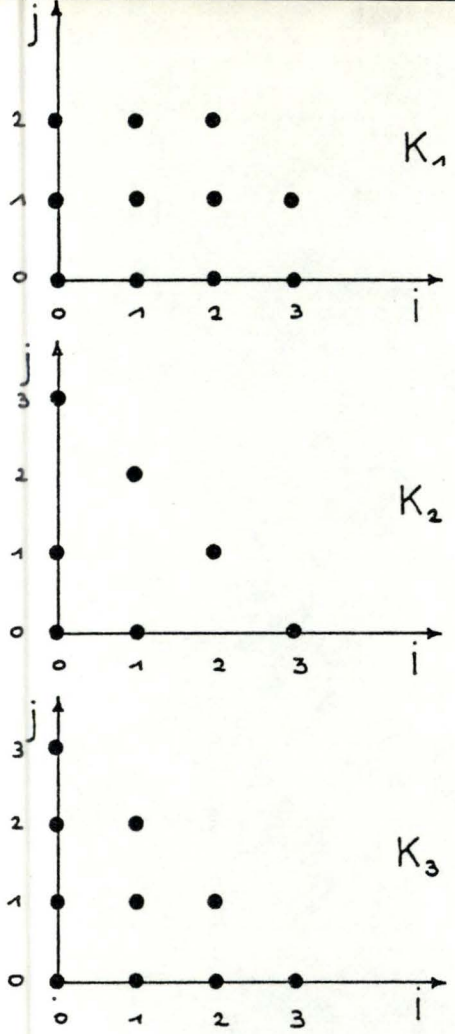


fig. 2.7.

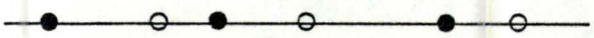


fig. 2.10.

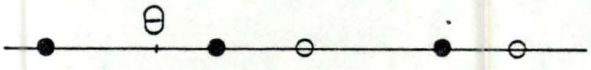


fig. 2.11.

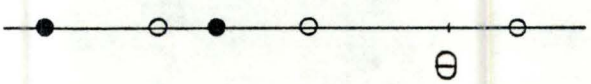


fig. 2.12.

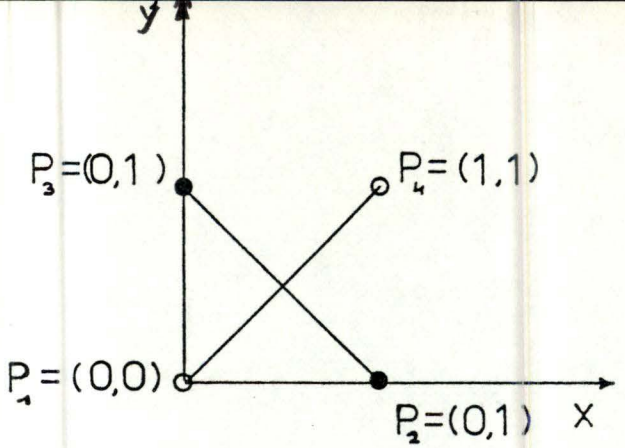
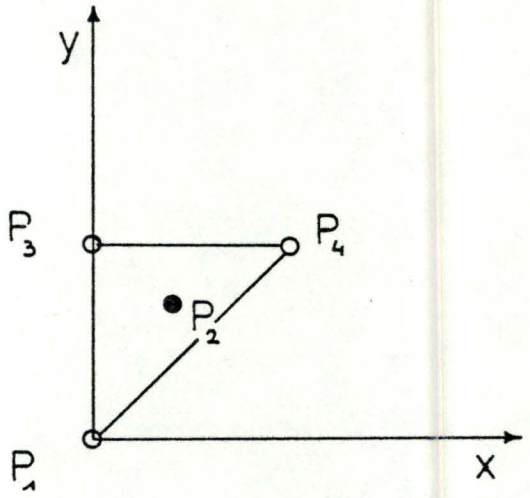


fig. 2.8



a.

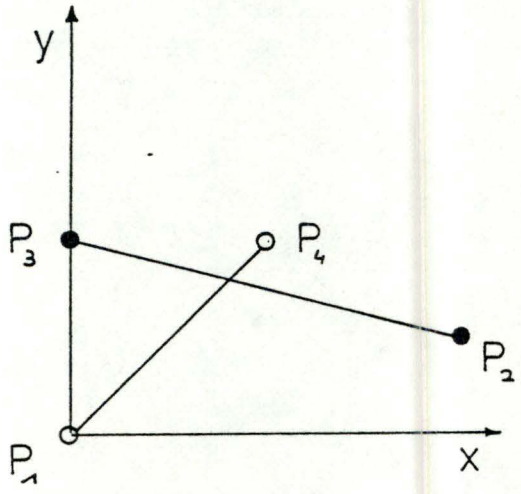
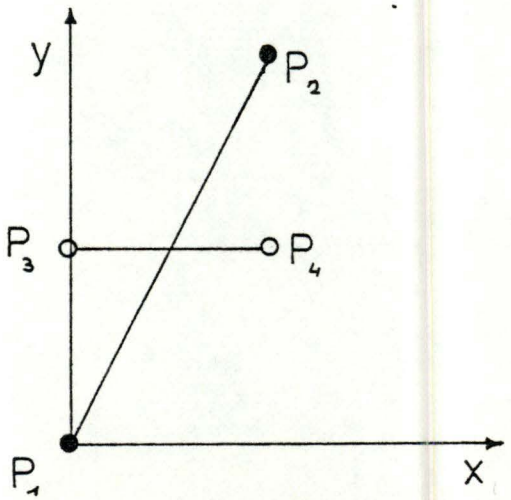


fig. 2.9.

b.



c.

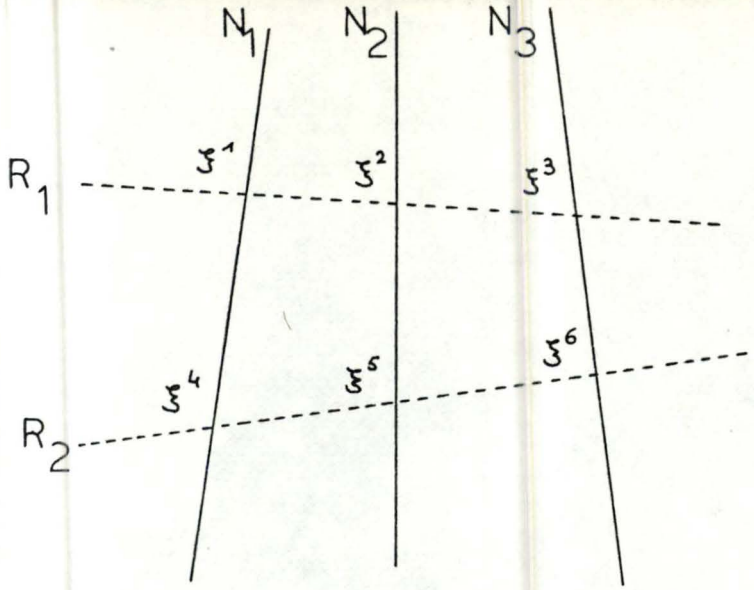


fig. 2.13.

$m_1 = 2 \quad m_2 = 3$

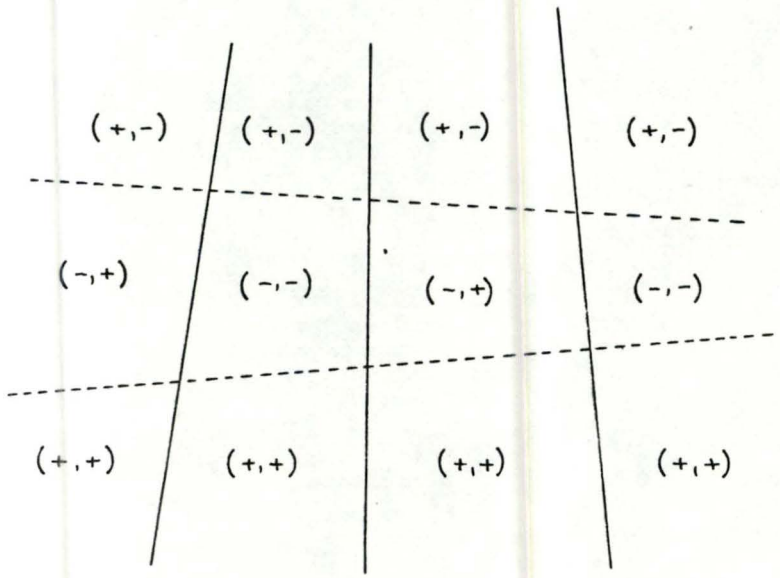


fig. 2.14.

$m_1 = 2 \quad m_2 = 3$

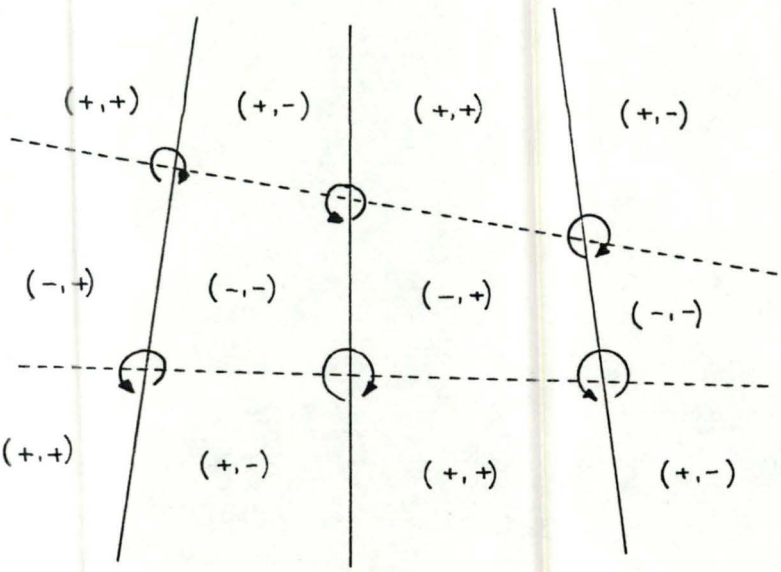


fig. 2.15.

$m_1 = 2 \quad m_2 = 3$

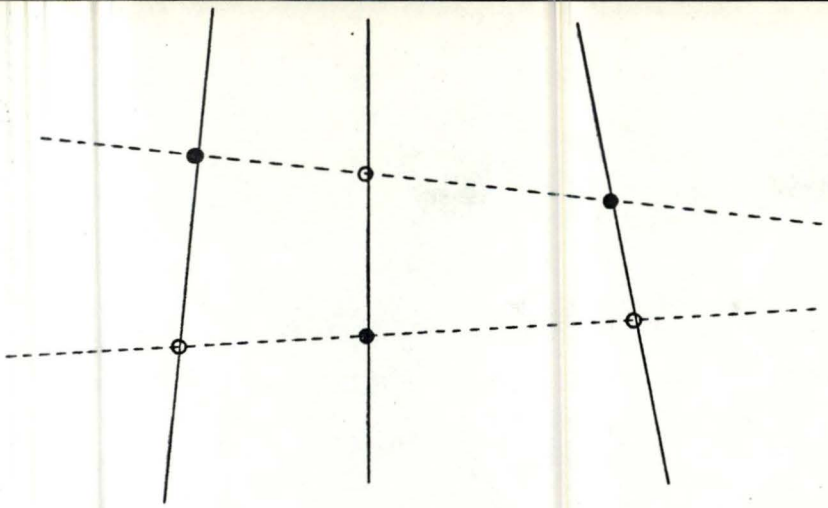


fig. 2.16.

$m_1 = 2 \quad m_2 = 3$

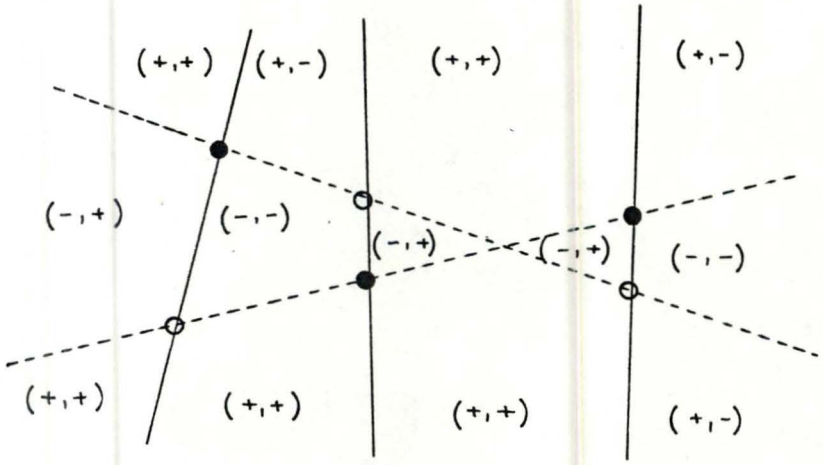


fig. 2.17.

$m_1 = 2 \quad m_2 = 3$

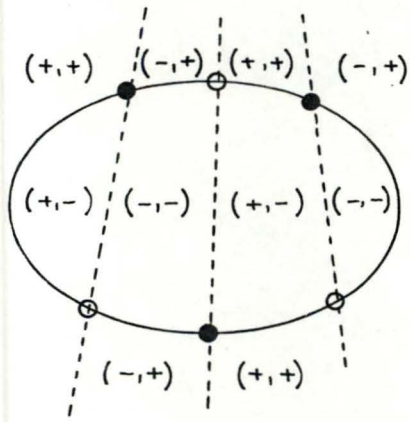


fig. 2.18

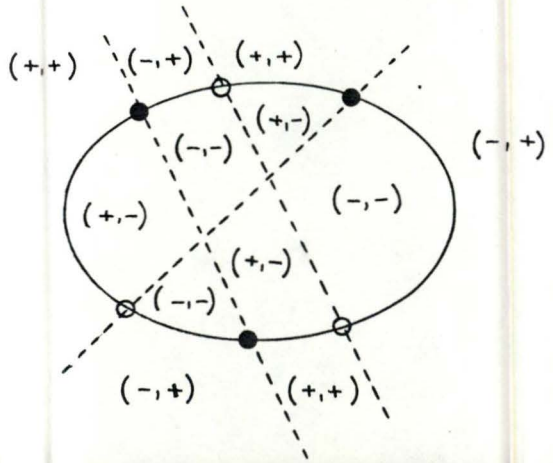


fig. 2.19.

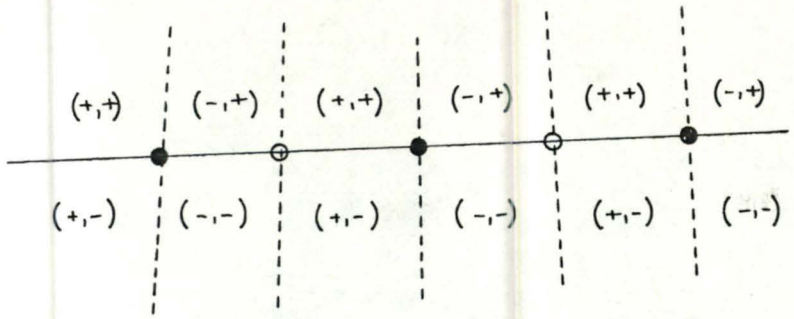


fig. 2.20.

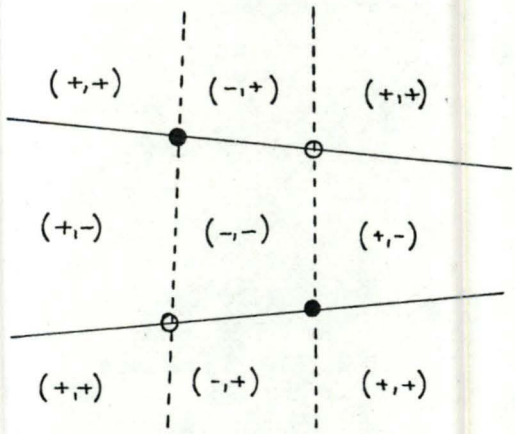


fig. 2.21.

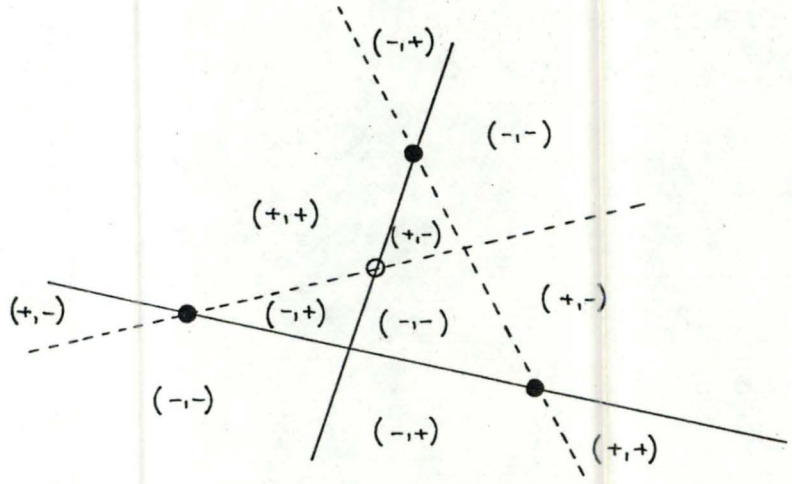


fig. 2.22.

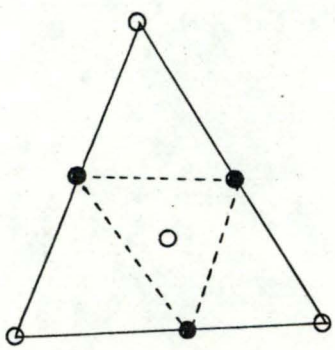


fig. 2.23.

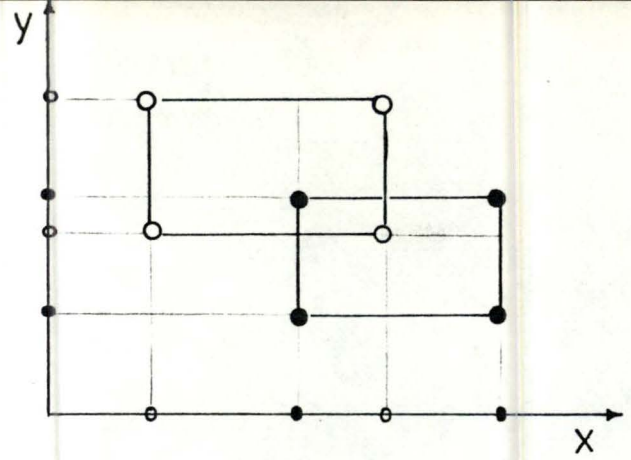


fig. 2.24.

n=2

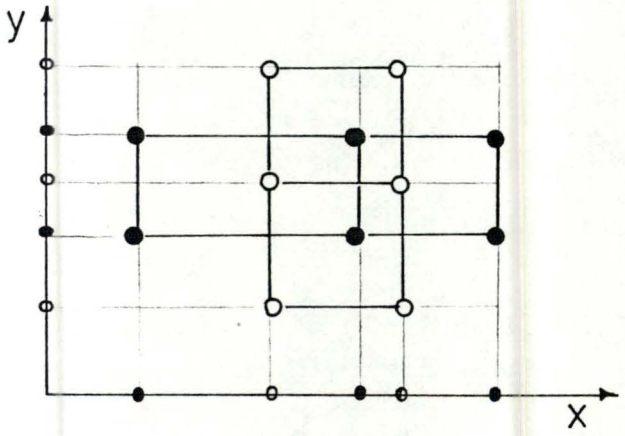


fig. 2.25.

n= 3

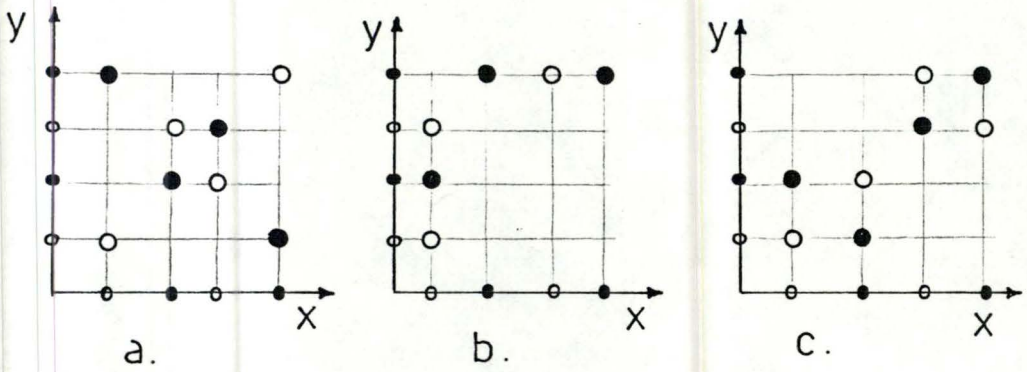


fig. 2.26.

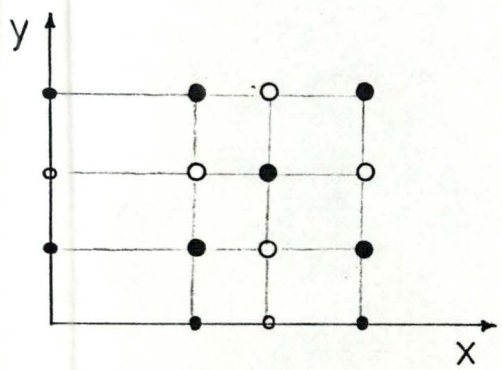


fig. 2.27.

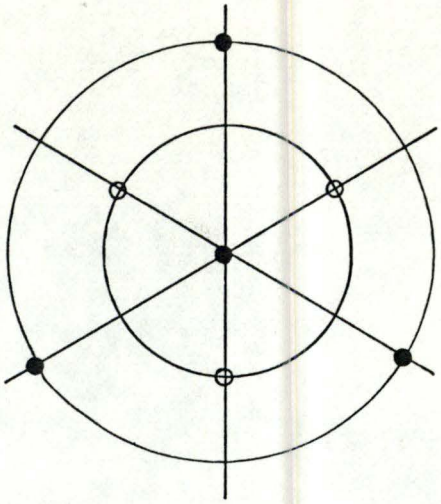


fig. 2.28.

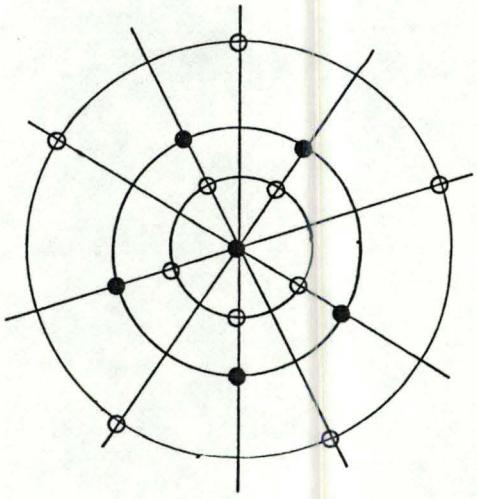


fig. 2.29.

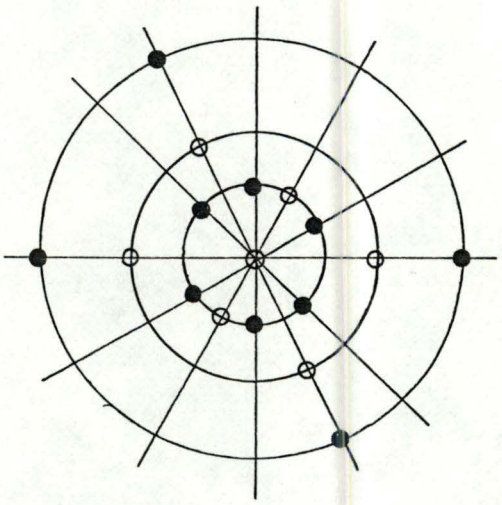


fig. 2.30.

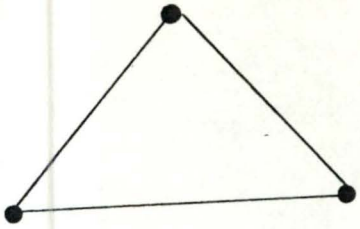


fig. 3.1.

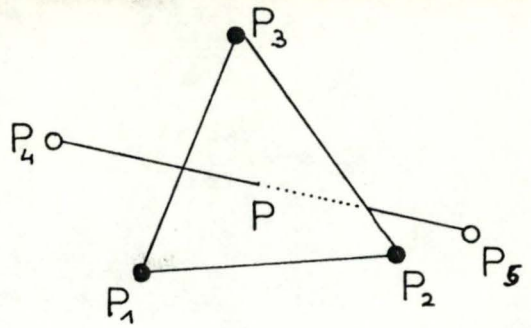


fig. 3.2

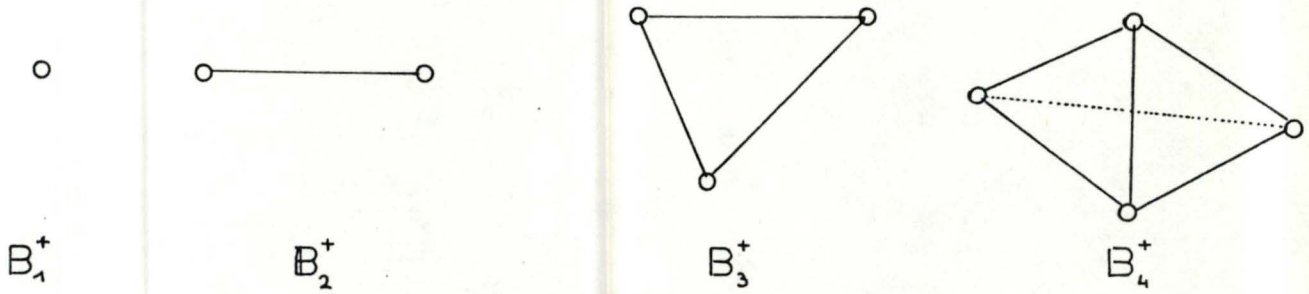
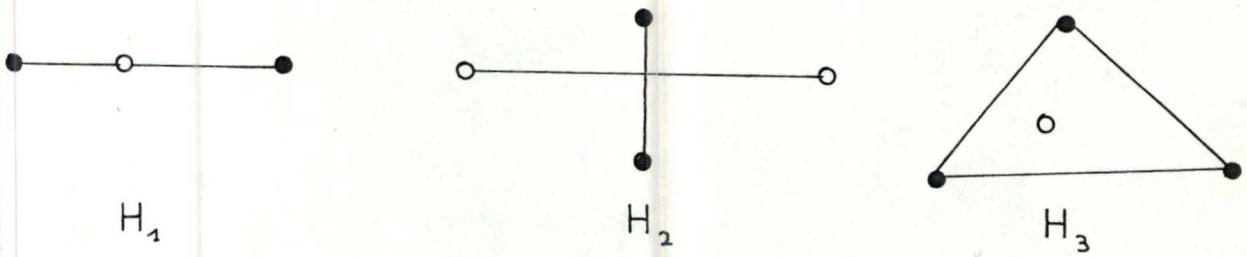


fig. 3.3.



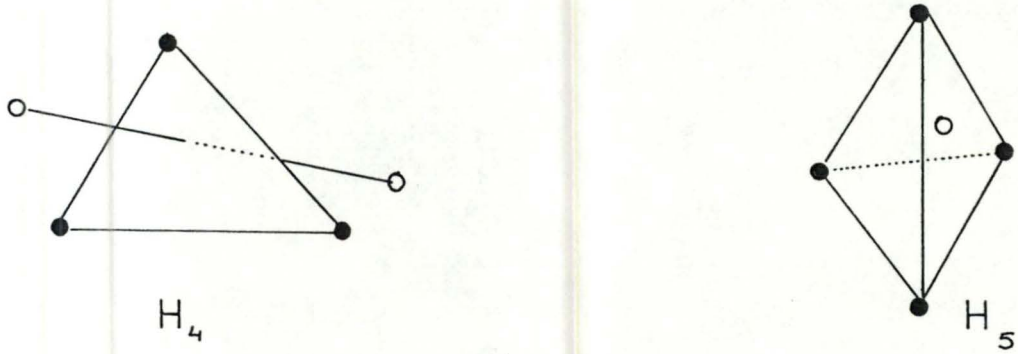
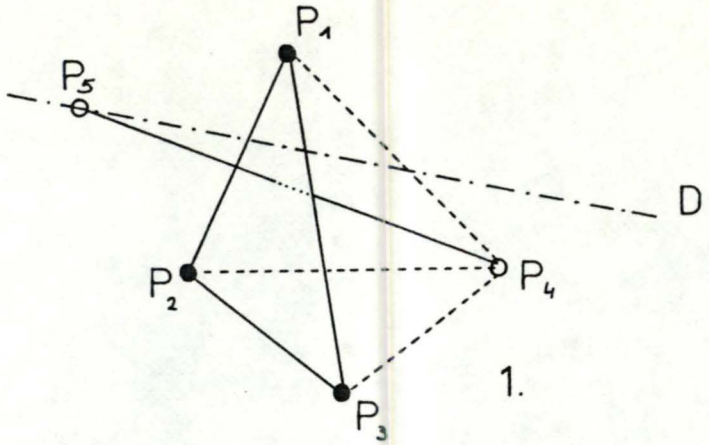


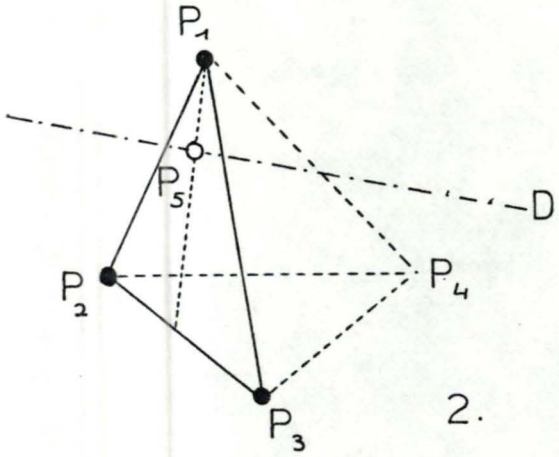
fig. 3.4.



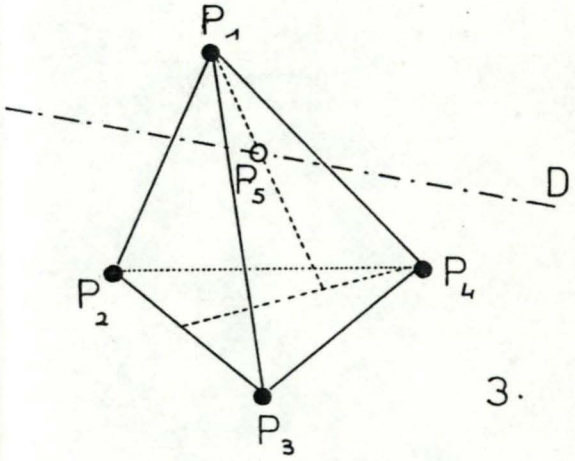
fig. 3.5.



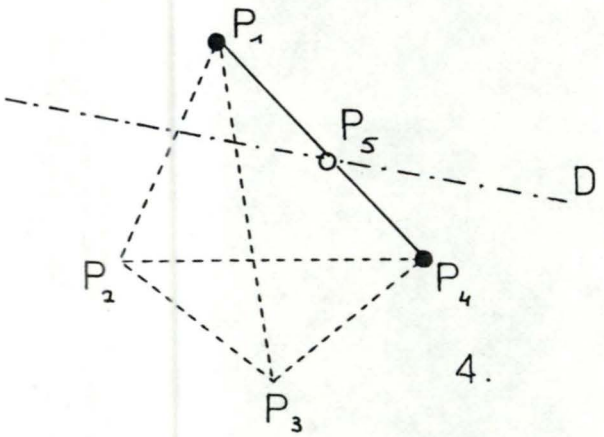
1.



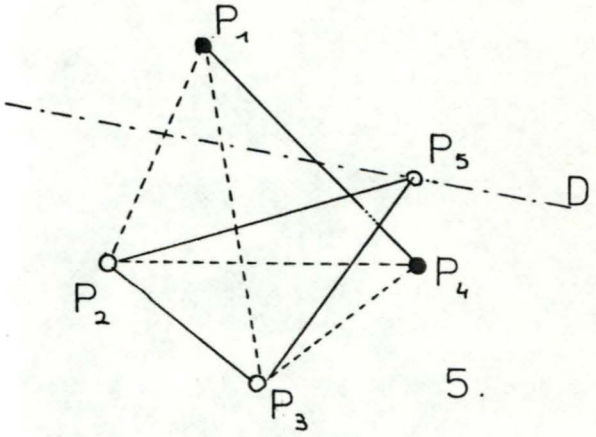
2.



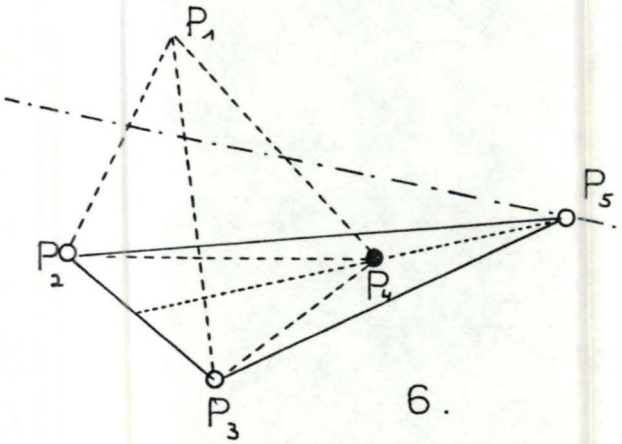
3.



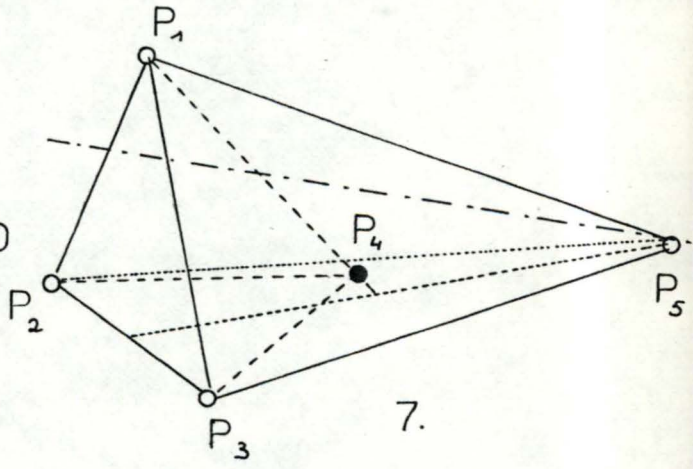
4.



5.



6.



7.

fig. 3.6.

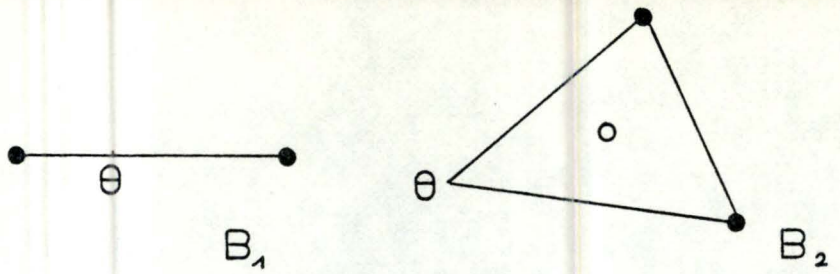
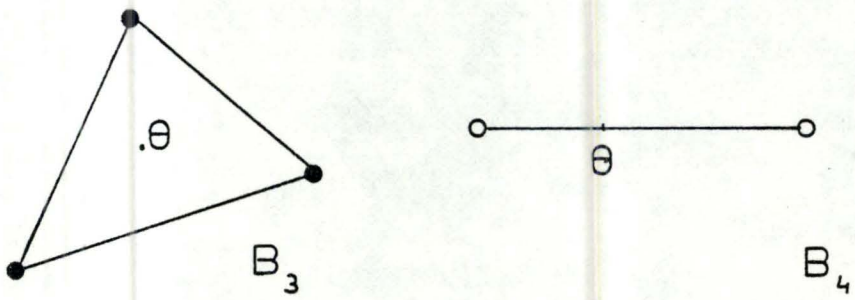


fig. 3.7.



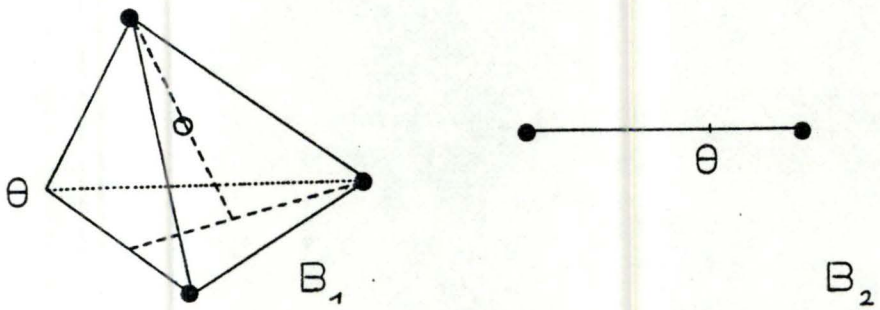
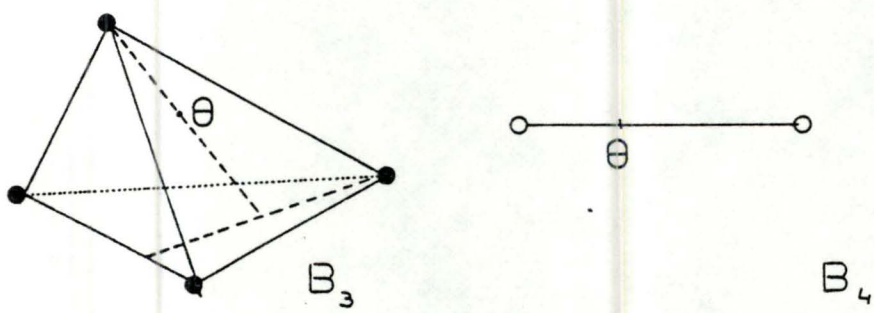
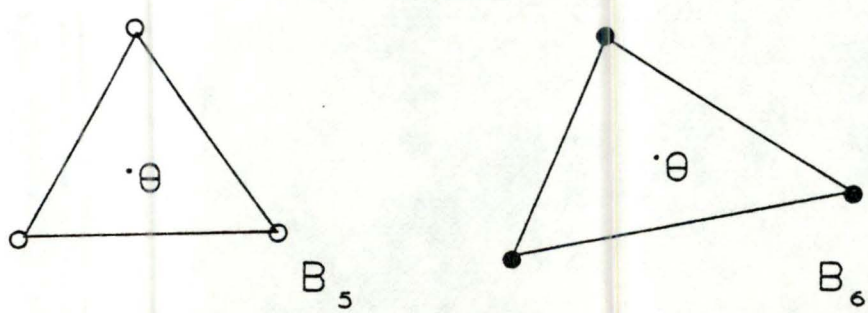


fig. 3.8.





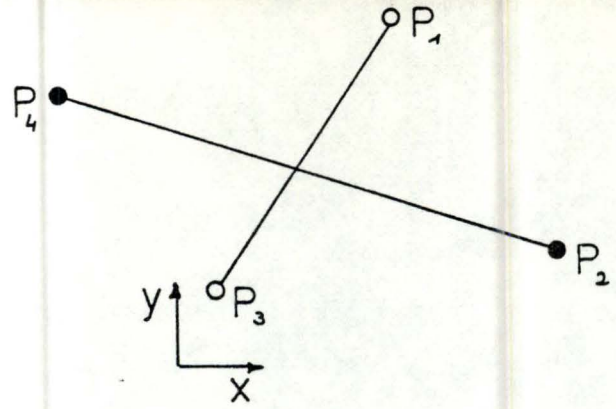


fig. 4.1.

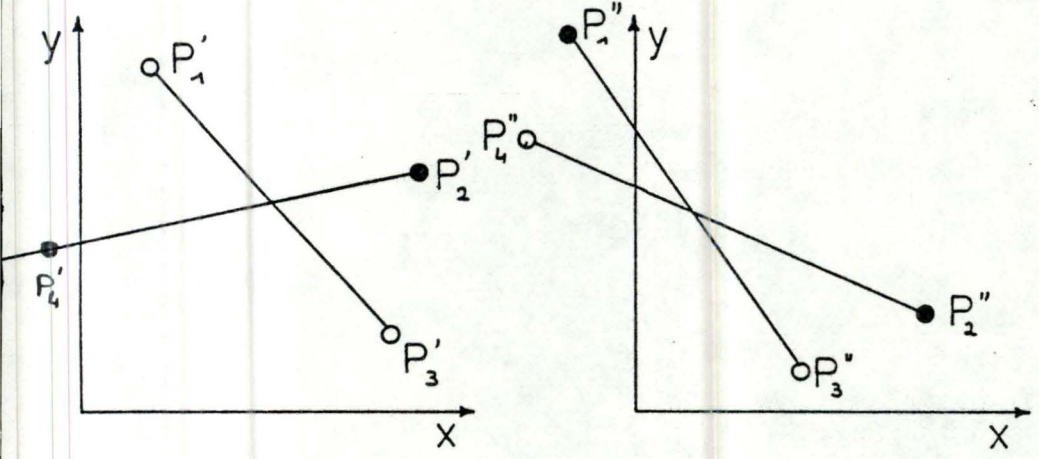


fig. 4.2.

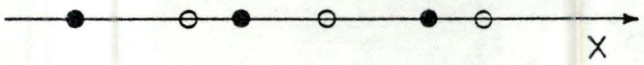


fig. 4.3.

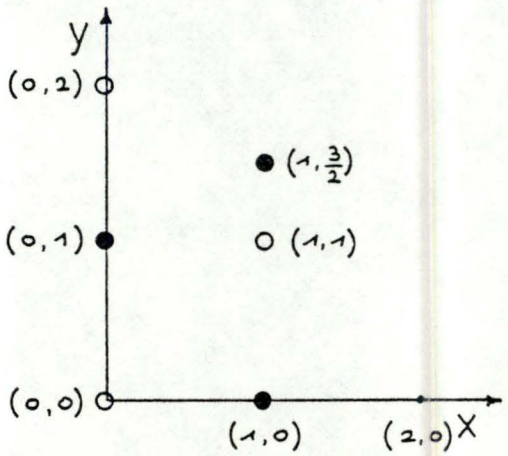


fig. 4.4.

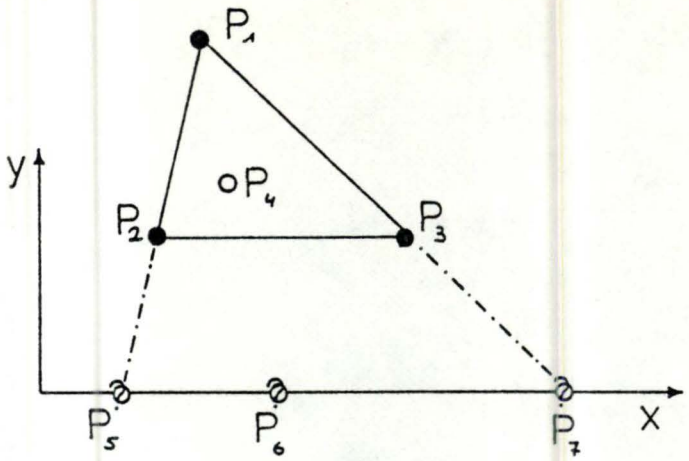


fig. 4.5.

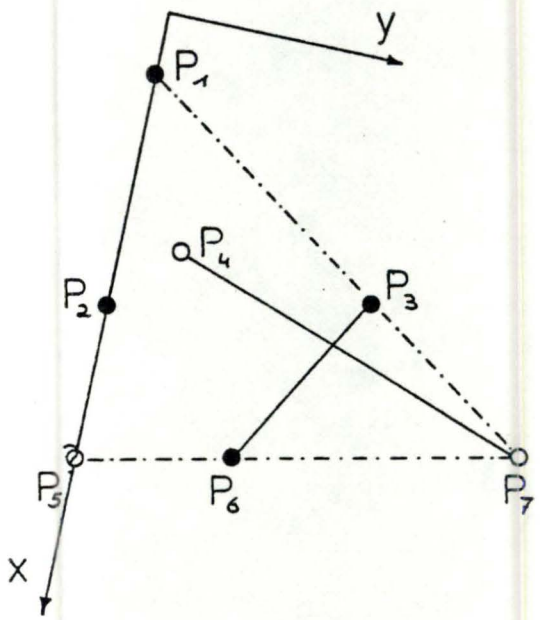


fig. 4.6.

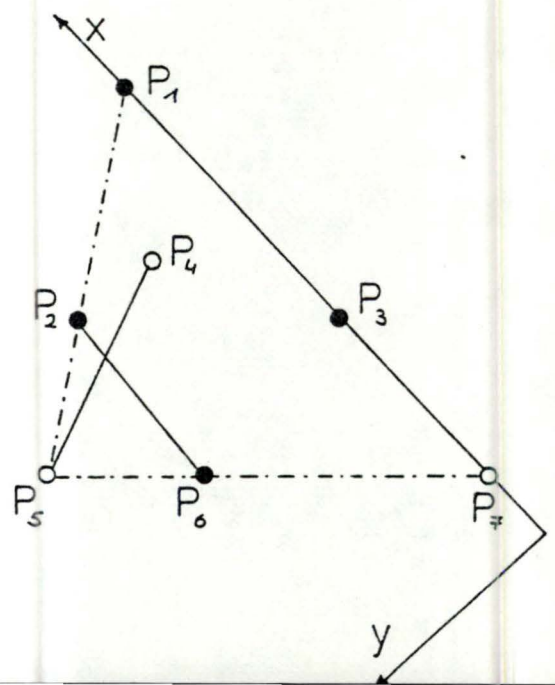
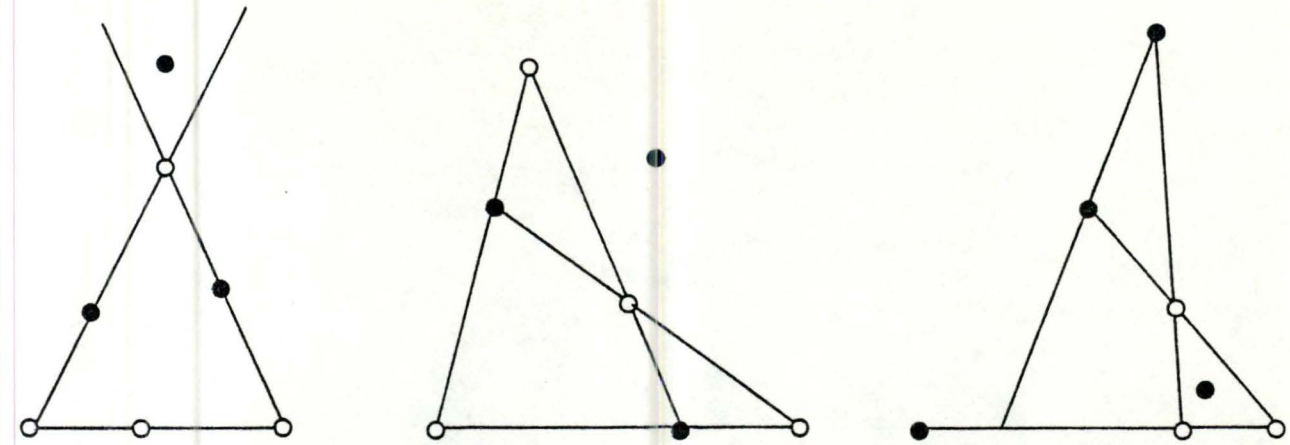


fig. 4.7.



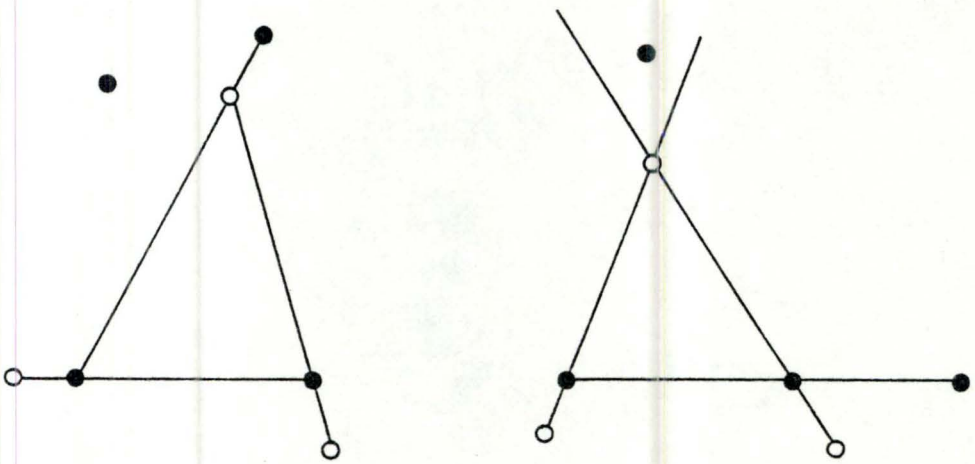


fig. 4.8

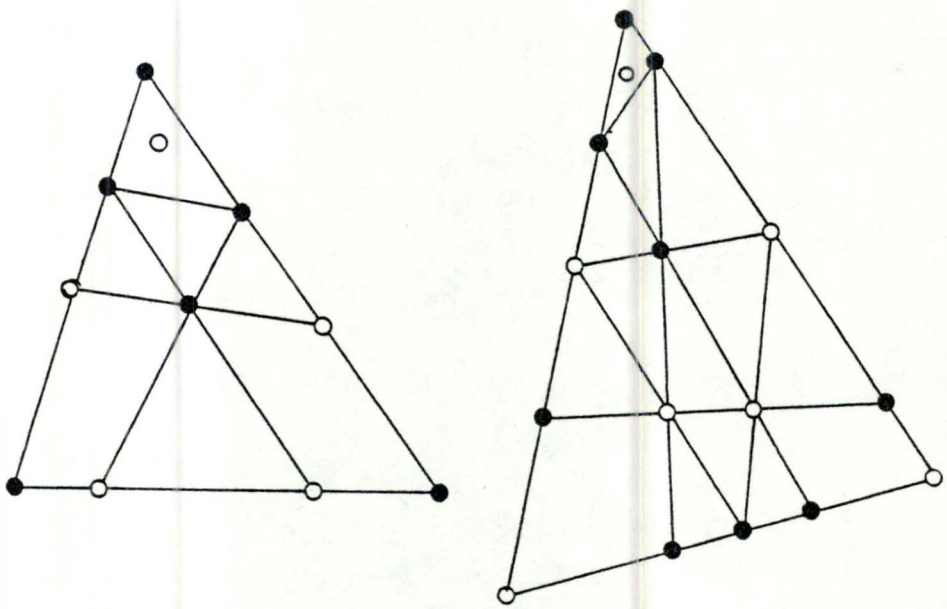


fig. 4.9.

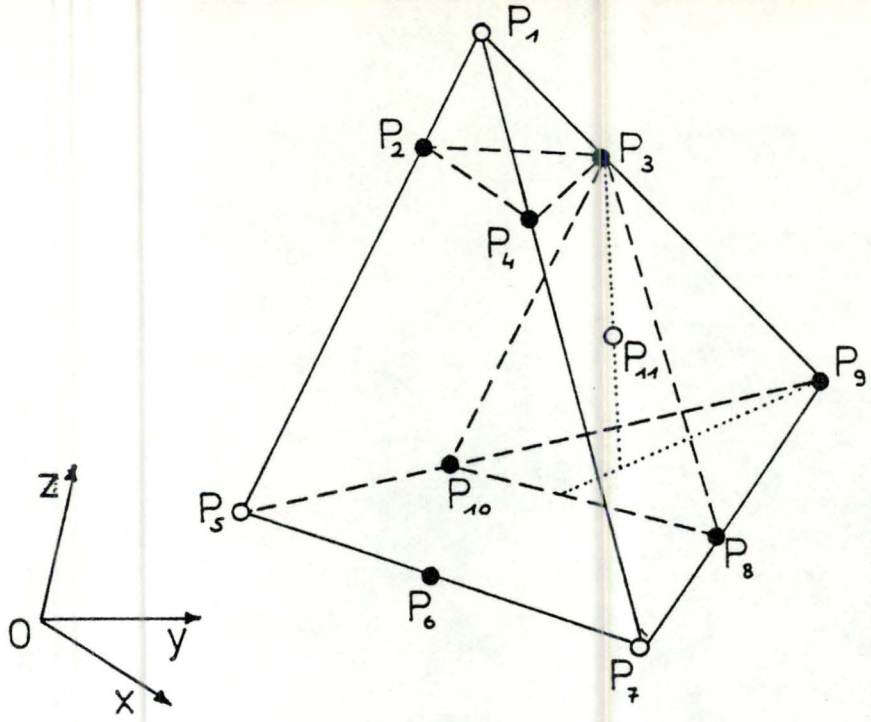
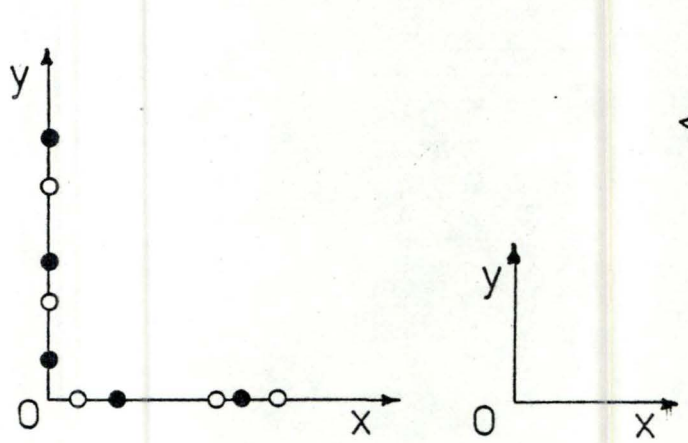


fig. 4.10.



a. b.

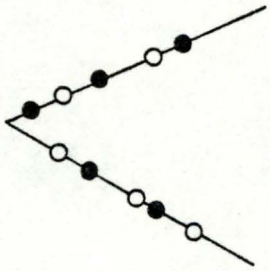
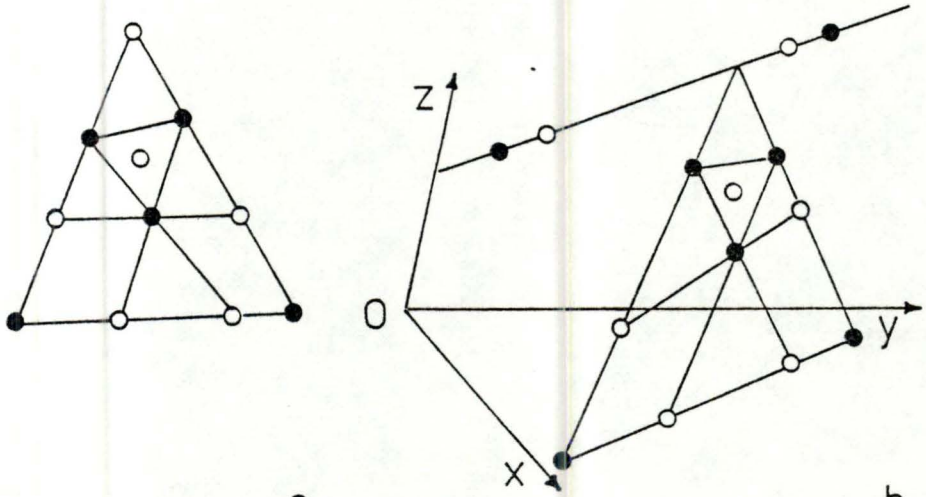


fig. 4.11.



a. b.

fig. 4.12.

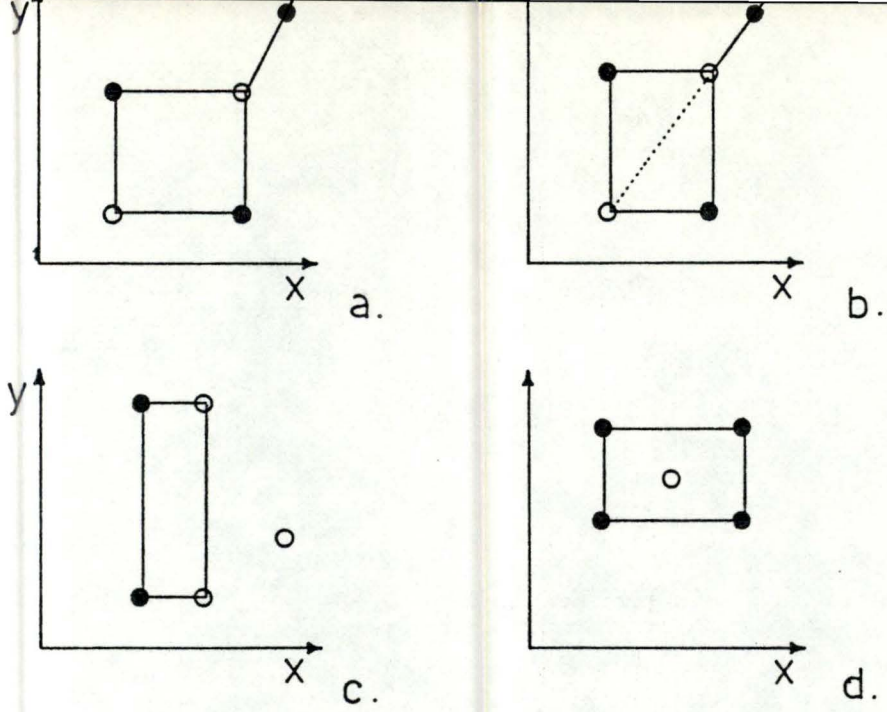


fig. 4.13.

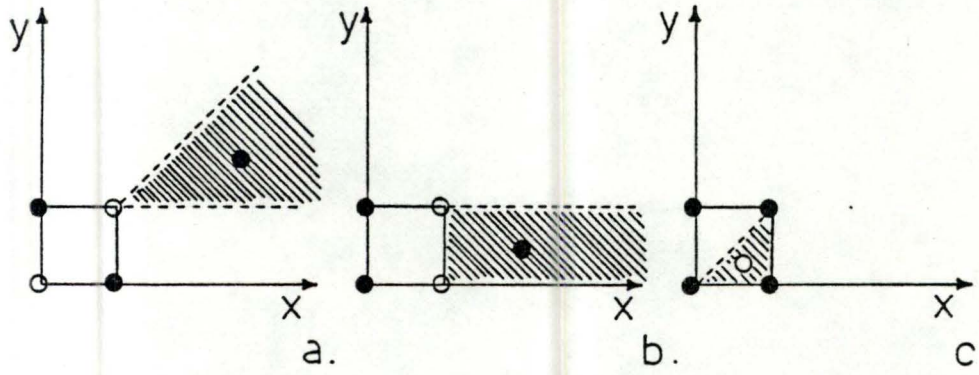


fig. 4.14.

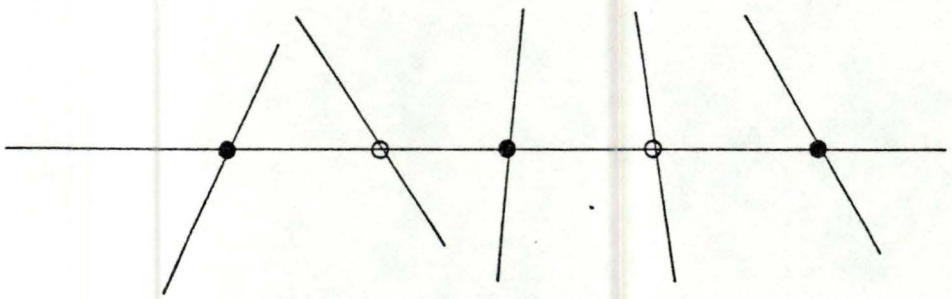


fig. 4.15.

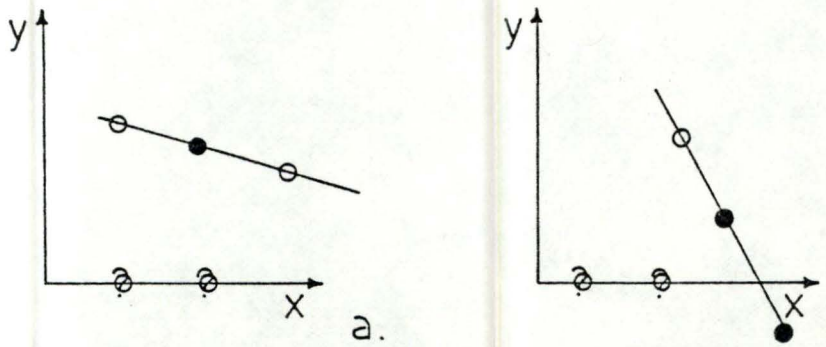


fig. 4.16.

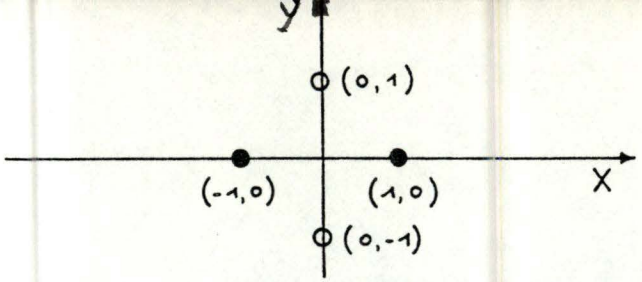


fig. 5.1.

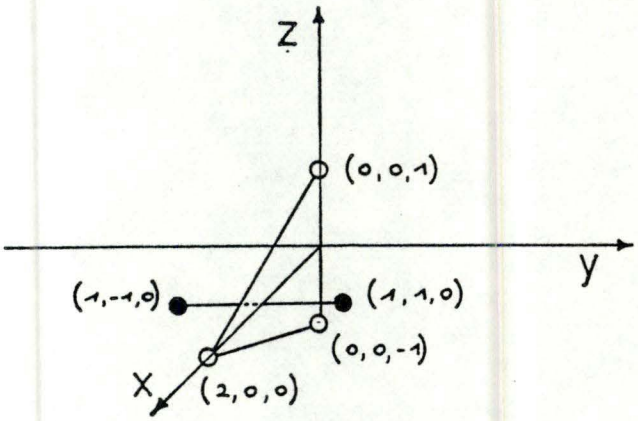


fig. 5.2.

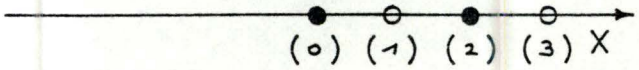


fig. 5.3.

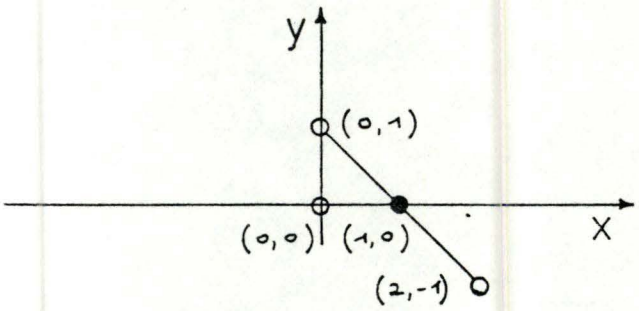


fig. 5.4.

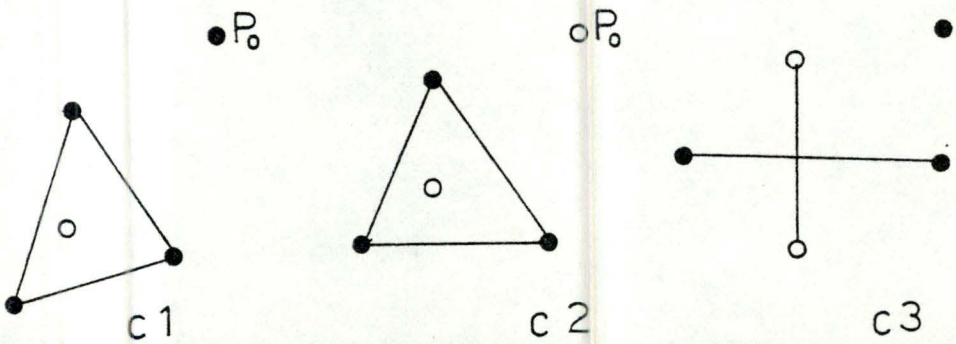


fig. 5.5.

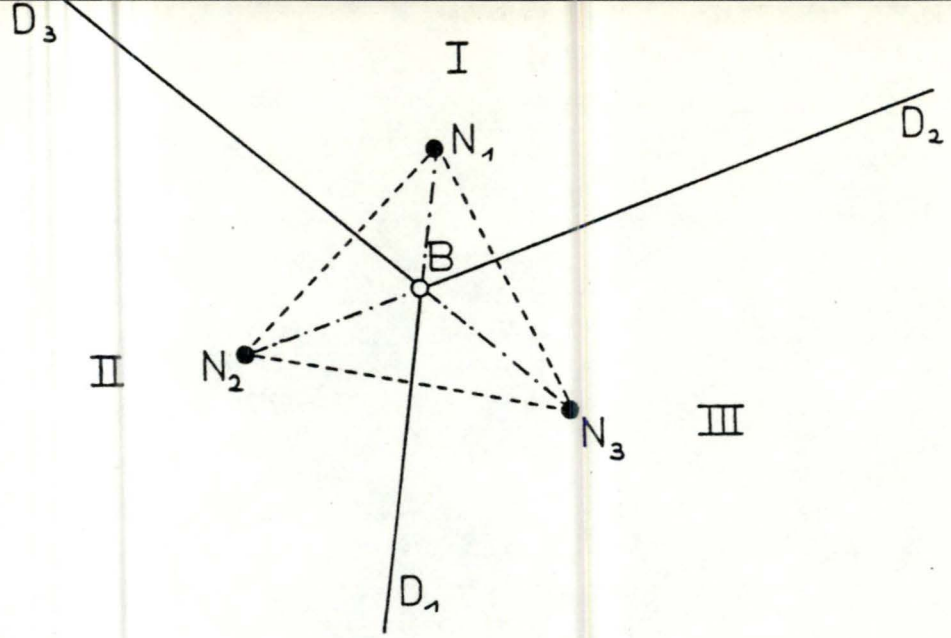


fig. 5.6

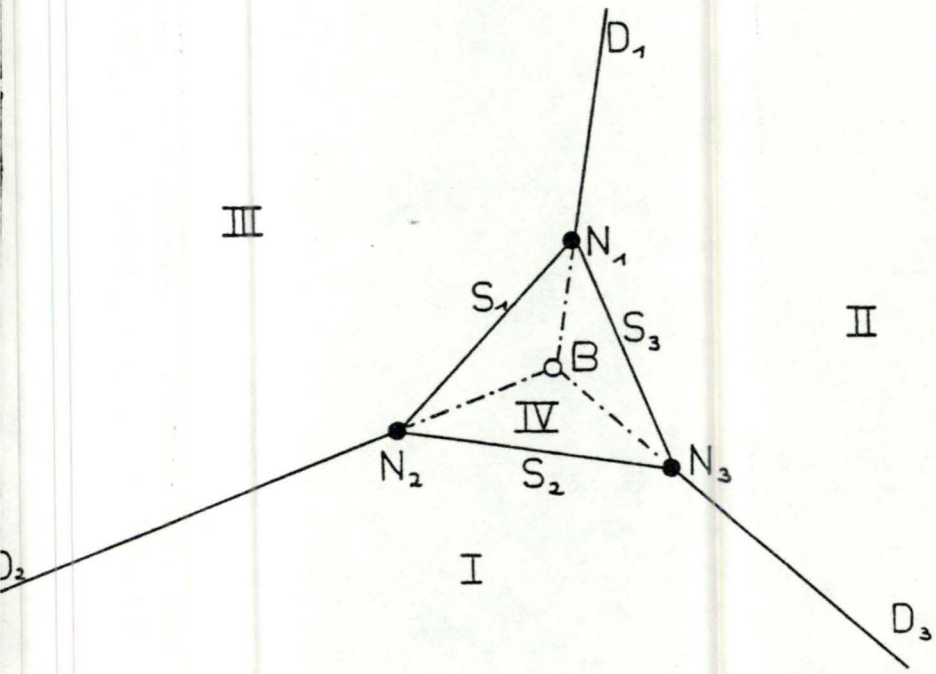


fig. 5.7

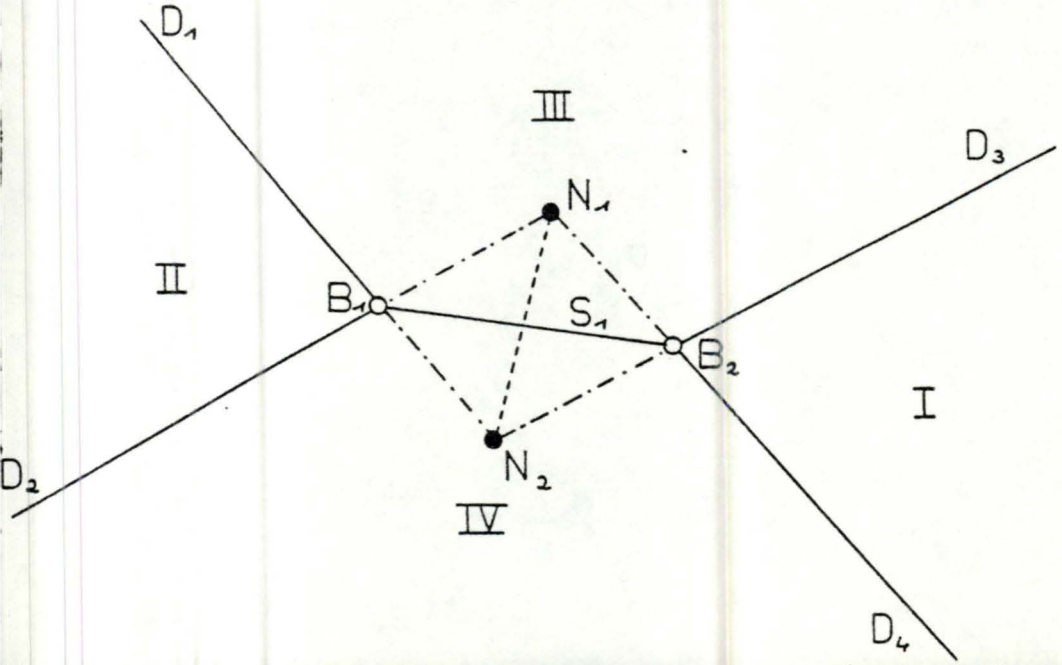
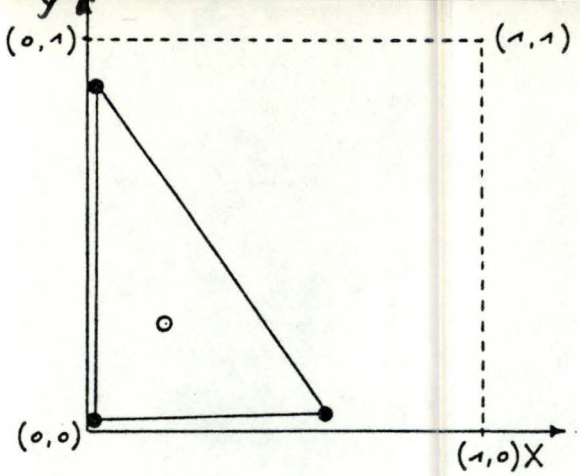
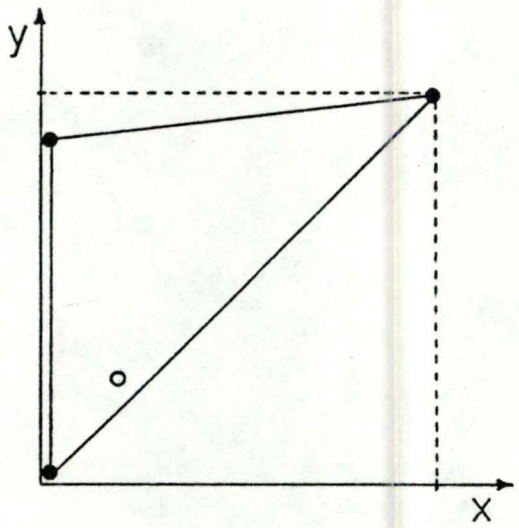


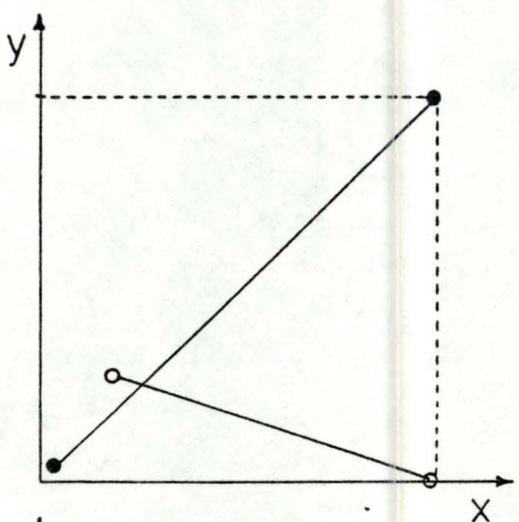
fig. 5.8



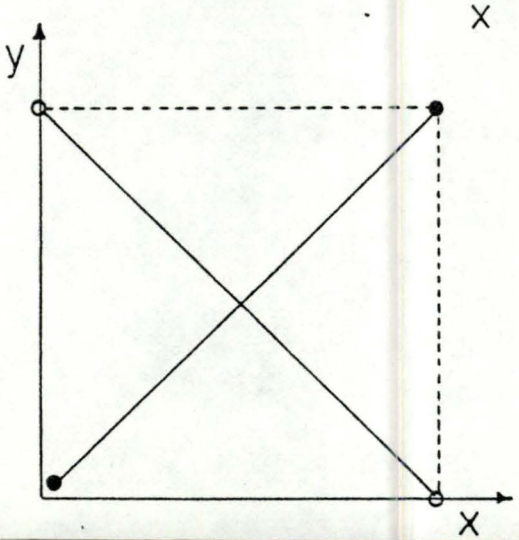
M^1



M^2

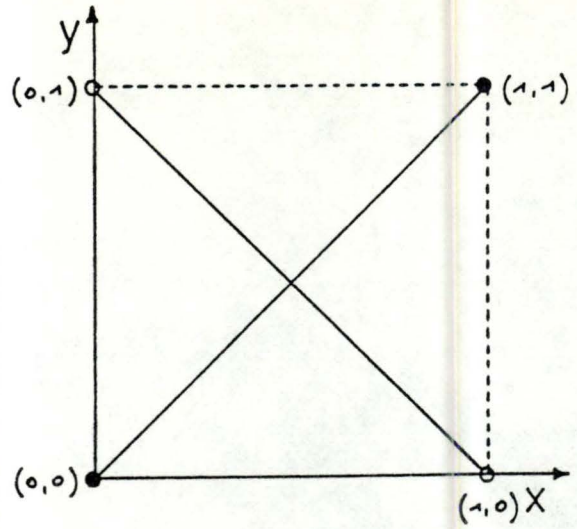


M^3

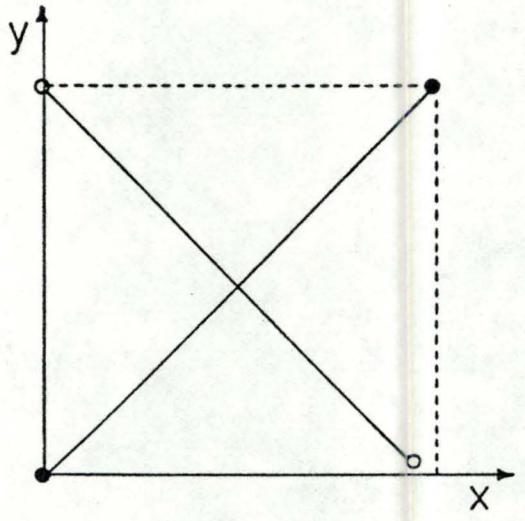


M^4

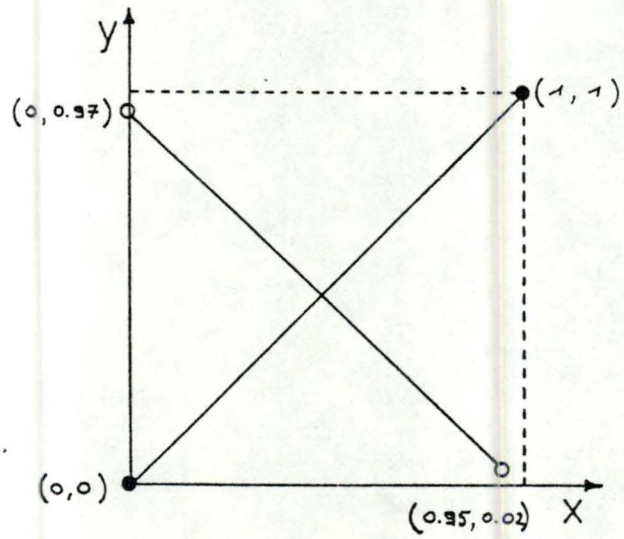
fig. 5.9.



M^5

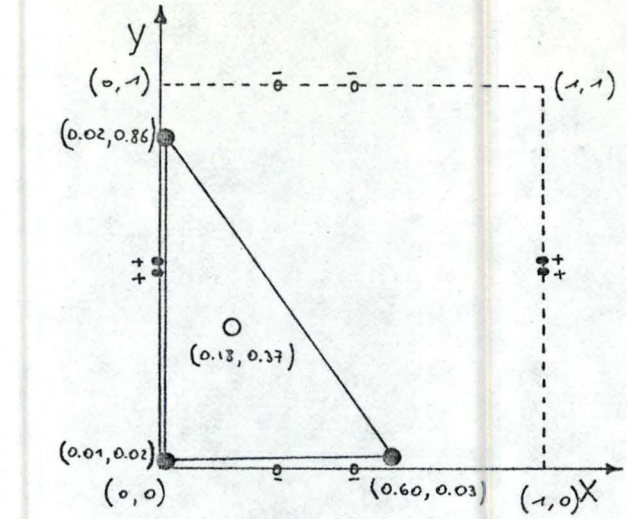


M^6

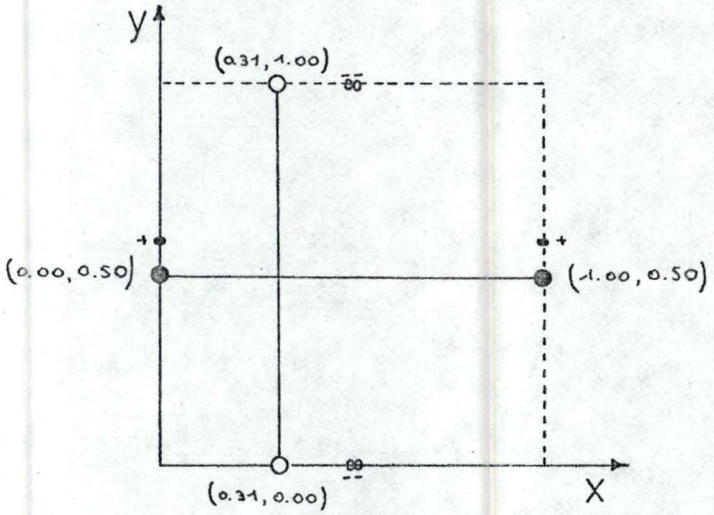


M^7

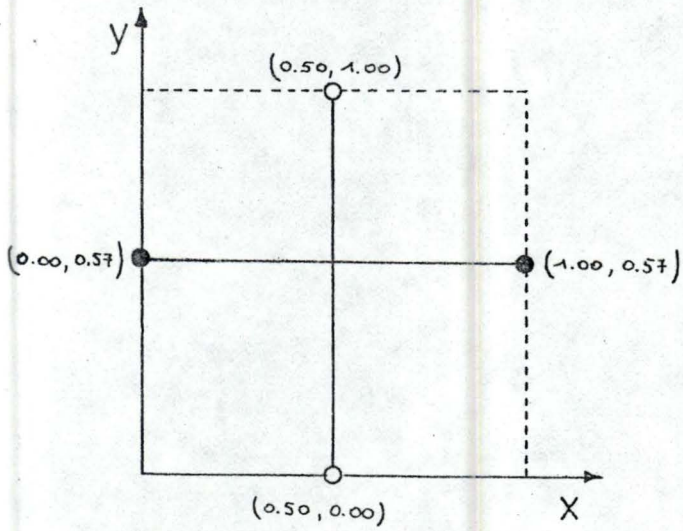
fig. 5.9.



M_1



M_2



M_3

fig. 5.10.

Références.

REFERENCES :

[1] BROSOWSKI, Bruno : Uber Extremalsignaturen linearer Polynome in n Veränderlichen, Numerische Mathematik 7 (1965), 396-405.

[2] CHENEY, E.W. : Introduction to Approximation Theory, International Series in Pure and Applied Mathematics, Mc Graw Hill, New York, 1966.

[3] COLLATZ, L. : Approximation von Funktionen bei einer und bei mehreren unabhängigen Veränderlichen, Z. Angew. Math. Mech. 36 (1956), 198-211.

[4] COLLATZ, L. : Inclusion Theorems for the Minimal Distance in Rational Tchebycheff Approximation with Several Variables, Approximation of Functions (H.L. Garabedian, ed.), 43-56, Elsevier, Amsterdam, 1965.

[5] COLLATZ, L. et KRABS, W. : Approximationstheorie, 1973, Teubner, Stuttgart.

[6] DIERIECK, C. : Contributions to Approximation Theory, Thèse de doctorat, Université Catholique de Louvain, 1977.

[7] GEARHART, W.B. : Some Extremal Signatures for Polynomials, Journal of Approximation Theory, 7 (1973), 8-20.

[8] KRONECKER, L. : "Werke", Teubner Verlag, Leipzig, 1895, Band I.

[9] LAURENT, P.J. : Approximation et Optimisation, p. 74.

[10] LAURENT, P.J. : Théorèmes de Caractérisation d'une Meilleure Approximation dans un Espace Normé et Généralisation de l'Algorithme de Rémès, Normes. Math. 10 (1967), 190-208.

[11] MANGAZARIAN : Non Linear Programming, pp. 16-37, Mc Graw-Hill Book C°, New York.

[12] NEWMAN, D.J. and SHAPIRO, H.S. : Some Theorems on TCHEBYCHEFF Approximation, Duke Math. J. 30 (1963), 673-682.

[13] ROYDEN, H.L. : Real Analysis, Mac Millan, N.Y., 1968.

[14] SHAPIRO, H.S. : Some Theorems on TCHEBYCHEFF Approximation II, Journal of Mathematical Analysis and Applications 17 (1967), 262-268.

[15] TAYLOR, G.D. : On Minimal H-Sets, Journal of Approximation Theory 5 (1972), 113-117.

Table des matières :

Chapitre 1 : Introduction à la théorie des ensembles-H.

1.1. Position du problème d'approximation linéaire au sens de TCHEBYCHEFF
1.2. Définition des ensembles-H et théorème d'inclusion
1.3. Formulation duale de la définition d'ensemble-H
1.4. Définitions apparentées
1.5. Propriétés des ensembles-H

Chapitre 2 : Invariance et méthodes de construction

2.1. Invariance d'ensembles-H
2.2. Méthodes de construction d'ensembles-H
- Principes de décomposition
- Méthode utilisant le signe du Jacobien
- Méthode basée sur le produit cartésien
- Méthode basée sur le passage aux coordonnées polaires

Chapitre 3 : Les ensembles-H pour une approximation multivariable de degré 1

3.1. Théorèmes de base
3.2. Choix d'une méthode de construction
3.3. Nombre d'ensembles-H dans $\mathbf{R}^n$
3.4. Construction systématique des ensembles-H minimaux
3.5. Classification de BROSOWSKY

Chapitre 4 : Ensembles-H minimaux relatifs à une approximation par des polynomes à plusieurs variables de degré supérieur ou égal à deux

4.1. Préliminaires
4.2. Construction d'ensembles-H à l'aide des ensembles-B
4.3. Construction d'ensembles-H à l'aide des blocs
4.4. Premières conclusions
4.5. Ensembles-H minimaux relatifs à V = eng {1, x, y, xy}

Chapitre 5 : Algorithmes et résultats numériques

5.1. Détermination d'ensembles-H

5.2. Algorithme de REMES

5.3. Algorithme d'échange point par point

5.4. Algorithme d'échange global

Annexe 1 : Programmes

Annexe 2 : Figures

Références