



Zonotopes et zonoïdes : études et applications aux processus de la séparation

Otmane Daoudi

► To cite this version:

Otmane Daoudi. Zonotopes et zonoïdes : études et applications aux processus de la séparation. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1995. Français. NNT : . tel-00346057

HAL Id: tel-00346057

<https://theses.hal.science/tel-00346057>

Submitted on 11 Dec 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE

présentée par

Otmane DAOUDI

pour obtenir le titre de

DOCTEUR
DE L'UNIVERSITE JOSEPH FOURIER-GRENOBLE 1
(arrêtés ministériels du 5 Juillet 1984 et du 30 Mars 1992)

Spécialité : Mathématiques Appliquées

Zonotopes et Zonoïdes :
Etudes et Applications aux Processus
de la Séparation.

Date de soutenance: 19 Octobre 1995.

Composition du jury :

M. B. Van CUTSEM	Président
M. D. APPRATO	Rapporteurs
M. J.C. FIOROT	
M. B. LACOLLE	Examineurs
M. P. J. LAURENT	
M. P. VALENTIN	

Thèse préparée au sein du laboratoire LMC-IMAG

A ma mère,

à mon père.

Remerciements

Ce travail a eu lieu au sein du laboratoire **LMC** dans l'équipe Modélisation Géométrique et Approximation en collaboration avec le **C.R.E.S.**, Centre de Recherche Elf-Solaize.

La présente thèse a été effectuée sous la direction du Professeur Pierre-Jean Laurent, professeur à l'Université Joseph Fourier. Sa confiance, ses conseils, sa perspicacité, sa générosité et sa grande compétence ont guidé mes premiers pas dans la recherche et m'ont permis d'apprendre et de m'améliorer. Je lui en suis éternellement reconnaissant.

Monsieur Patrick Valentin a été à l'origine de ce travail. Je le remercie pour m'avoir accueilli au CRES, me permettant ainsi de découvrir des problèmes concrets. Son aide constante, ses conseils et son amitié ont été déterminants.

Je remercie Monsieur Bernard Van Cutsem, professeur à l'Université Joseph Fourier, qui me fait l'honneur de présider le jury, et pour l'intérêt qu'il a porté à ce travail. Ses conseils et ses encouragements m'ont été très précieux.

Je remercie Monsieur Dominique Apprato, professeur à l'Université de Pau, et Monsieur J. C. Fiorot, professeur à l'Université de Valenciennes, d'avoir accepté d'être les rapporteurs de cette thèse.

Monsieur Bernard Lacolle, professeur à l'Université Joseph Fourier, a porté un grand intérêt à ce travail, je le remercie pour la sympathie qu'il a eue à mon égard. Les discussions que nous avons eues ont été très fructueuses.

Je remercie mes amis (es) du laboratoire LMC (je ne nommerai personne, sinon il faut les citer tous) pour leurs soutiens et leurs encouragements. Le climat et l'ambiance, qu'ils ont pu créer, ont rendu le travail plus gai et agréable. Dans ma mémoire, je garderai plus particulièrement les soirées entre nous et les matchs de football...

Une pensée pour mes ex et présents camarades du bureau 58 : Raúl, Raj, Hamid, Malek, Elbar et Mireille mes coéquipiers et compagnons durant cette thèse.

Mireille m'a constamment soutenu et encouragé. Sa lecture et re-lecture de mon manuscrit m'ont permis de l'améliorer. Merci.

Mes parents et ma famille ont fait tant d'efforts pour que je puisse poursuivre mes études. Je leur en suis profondément reconnaissant.

Merci à tous ceux qui, de près ou de loin, ont contribué à l'aboutissement de ce travail.

Table des matières

Introduction	1
Chapitre 1. Zonotopes	5
1 Définitions et propriétés	6
1.1 Définitions et quelques propriétés des zonotopes	6
1.2 Sommets et faces d'un zonotope	8
1.3 Volume d'un zonotope	11
2 Premier diagramme projectif	13
2.1 Définitions et propriétés	13
2.2 Exemples	16
3 Zonoèdre et arrangement de droites dans le plan	20
4 Zonotopes et problème de groupement de vecteurs	23
4.1 Critère de groupement et problème d'optimisation	23
4.2 Méthode de recuit simulé	23
4.3 Application de l'algorithme au problème	25
Chapitre 2. Sensibilité du problème d'optimisation pour la fabrication des mélanges	33
1 Modèle géométrique de la fabrication de mélange	34
1.1 Modèle géométrique	34
1.2 Faisabilité	35
2 Critère de gestion de la fabrication d'un mélange	36
2.1 Zonotope résiduel	36
2.2 Fabrication à séparation résiduelle maximale	38
2.3 Algorithme de résolution	41
3 Influence des erreurs d'analyses sur la solution:	45
3.1 Régularité de la solution par rapport aux erreurs	45
3.1.1 Sensibilité dans le cas de contraintes linéaires de type égalité	45
3.1.2 Cas des contraintes de type égalité et inégalité	50
3.2 Région de confiance	54
3.3 Exemples :	58
3.3 Variations du mélange à réaliser :	62
Chapitre 3. Zonoïdes Associés à une Courbe	67
1 Zonoïde, rappels et Définitions	68
1.1 Définitions et propriétés	68

1. 2	Exemples	68
1. 3	Fonction support	69
2	Zonoïdes associés à une courbe paramétrique régulière	69
2. 1	Définitions et propriétés	69
2. 2	Zonoïdes associés à une courbe et mesures vectorielles	71
2. 3	Diagramme projectif	73
3	Surface Zonoïdale	75
3. 1	Définition d'une surface zonoïdale	75
3. 2	Paramétrisation	76
3. 3	Courbe cumulée	77
3. 4	Régularité de la paramétrisation	78
3. 5	Etude d'un exemple	82
4	Sélectivité totalement régulière (STR)	86
4. 1	Définition et propriétés	86
4. 2	Paramétrisation de la surface zonoïdale dans le cas STR	87
4. 3	Régularité de la paramétrisation	88
4. 4	Exemples	89
5	Coupe zonoïde-plan	93
5. 1	Problème d'intersection d'un plan avec un zonoïde	93
5. 2	Méthodes de résolution:	95
5. 3	Quelques exemples de visualisation par des coupes parallèles	99
Chapitre 4. Approximation de zonoïdes par des zonotopes		101
1	Rappels de quelques résultats sur l'approximation	102
1. 1	Notions de convergence	102
1. 2	Approximation d'un ensemble convexe compact par un polytope	102
1. 3	Approximation de zonoïdes par la somme de Minkowski de segments	103
2	Approximation d'un zonoïde associé à une fonction	105
2. 1	Construction du zonotope approximant	105
2. 2	Majoration de l'erreur d'approximation $ R_p(u, F) $ et ordre de convergence	107
2. 3	Exemples numériques:	112
2. 4	Approximation dans le cas de la sélectivité régulière dans $\mathbb{R}^3$	120
3	Quelques exemples d'applications	120
4	Zonoïde associé à une fonction de plusieurs variables	124
4. 1	Construction du zonotope approximant	124
4. 2	Majoration de $ R_q(u, F) $ et ordre de convergence	125

5	Exemples	128
---	--------------------	-----

Chapitre 5. Représentation des tables ASTM-IP 53/54 par B-splines régularisées

	B-splines régularisées	135
1	B-splines régularisées, problème variationnel:	137
1.1	Lissage unidimensionnel par B-splines régularisées:	137
1.2	Lissage bi-dimensionnel par produit tensoriel de B-splines régularisées	140
2	Traitement Numérique	143
2.1	Définitions et notations	143
2.2	Tables ASTM-IP 53/54	144
2.3	Partie I de la table 54:	147
2.4	Partie II de la table 54:	152
2.5	Partie I de la table 53	155
2.6	Partie II de la table 53:	158
3	Traitement global	161
4	Raccordement à $D_{15} = 0.600$ entre les parties I et II de la table 54 .	163
5	Relation entre les tables 53 et 54	164
6	Conclusions	166

Introduction

Dans l'industrie pétrolière, le problème de fabrication des mélanges est resté pendant longtemps traité d'une manière empirique. Les études sur les lois de mélange ont permis de développer des concepts théoriques qui ont ouvert le champs vers des modélisations physiques du problème.

Le principe de la modélisation géométrique de la fabrication des mélanges a été introduit par P. Valentin. Il est basé sur la représentation vectorielle des mélanges : à chaque produit correspond un vecteur ayant pour composantes les spécifications de ce produit. Les résultats sur les lois de mélange et les fonctions d'indice ont permis (modulo certaines transformations) de prendre le vecteur des spécifications de sorte que ces composantes obéissent à des lois additives. Ainsi, le problème de mélange de deux produits sera ramené à l'opération de sommation entre les vecteurs de spécifications.

Lorsqu'on dispose dans une unité de production d'un certain nombre de produits de base; on leur associe, par la représentation précédente, des vecteurs de spécifications. L'ensemble de tous les mélanges faisables à partir de ces bases est constitué de l'ensemble de toutes les fractions possibles qu'on peut prendre de chaque produit. En géométrie, cet ensemble constitue un zonotope (somme de Minkowski de segments). Cet ensemble s'appelle aussi "convexe de mélanges".

Plusieurs études ont été menées pour résoudre des problèmes de fabrication par les zonotopes. Les premiers résultats sont dus à Slaoui dans [Sal86], ils portaient essentiellement sur les mélanges binaires (mélanges déterminés par deux spécifications). Ces résultats s'appuient sur les outils de la géométrie algorithmique. L'extension à des mélanges non binaires a été faite par Girard dans [Gir86], dans ce cas, la modélisation a amené à un problème d'optimisation.

Le mot zonoèdre (zonotope dans $\mathbb{R}^3$) a été introduit pour la première fois par Fedorov dans [Fed53] en cristallographie. L'auteur s'intéressait aux problèmes de pavage de l'espace par ces polytopes. D'autres auteurs se sont intéressés aux zonotopes. Parmi ceux qui ont donné des impulsions décisives dans le domaine, nous citons Coxeter, Shephard, Bolker et McMullen qui ont donné plusieurs résultats de points de vue géométrique et combinatoire.

Quand le nombre de segments tend vers l'infinie, le zonotope associé tend vers un ensemble qui n'est pas un polytope. Cet ensemble est appelé zonoïde. Les études menées sur ces ensembles ont permis de dégager plusieurs autres caractérisations.

Ainsi, Bolker dans son article [Bol69] a lié ces ensembles à la théorie des mesures vectorielles, définissant un zonoïde comme l'image d'une mesure vectorielle. Weil et Goodey dans [Wei76] et [Goo77] ont caractérisé les zonoïdes par des ensembles vérifiant certaines propriétés sur les volumes mixtes.

Outre l'industrie pétrochimique, les zonotopes ont eu des applications dans plusieurs autres domaines, nous citons par exemple:

- En économie, Hildenbrand a modélisé la production d'un groupe d'entreprise par un zonotope.
- En automatique, l'ensemble de tous les états atteignables constitue un zonoïde.
- En architecture, J. Baracs dans son article "Polyhedral Habitat", a pris le zonoèdre comme module de base pour l'occupation de l'espace.

En pétrochimie, les zonoïdes apparaissent dans les processus continus comme la chromatographie ou la distillation (cf [Val91]). Si on note $F(t) \in \mathbb{R}^n$ le vecteur des spécifications d'un produit à un instant donné t , l'ensemble de tous les mélanges faisables à partir de la famille $F(t)$, quand t décrit un intervalle de temps $[0, \Delta T]$, est un zonoïde appelé zonoïde associé à la fonction F . Une bonne connaissance de ces ensembles permet la compréhension et la résolution de quelques problèmes de séparation dans le cas des processus continus.

Dans le chapitre 1, nous commencerons par donner quelques propriétés géométriques sur les zonotopes (expressions des sommets et des faces, volume, ...). La dernière partie de ce chapitre est consacré à l'application des zonotopes au problème de groupement de vecteurs dans l'espace par paquets.

Dans le chapitre 2, nous donnerons les éléments de la modélisation géométrique dans le cas de la fabrication de mélanges. Nous rappellerons le problème de la gestion de la fabrication, et nous traiterons le problème d'optimisation. Un mélange est représenté par un vecteur de spécifications qui sont des résultats de mesures donc sujets à des erreurs. Celles-ci affectent évidemment le problème d'optimisation. Dans la deuxième partie, nous étudierons la sensibilité de la solution de ce problème par rapport à ces erreurs et nous caractériserons, sous certaines hypothèses, la région de confiance que décrit la solution perturbée lorsque les erreurs sont gaussiennes.

Dans le chapitre 3, nous nous intéresserons aux zonoïdes associés à une fonction définie sur un intervalle de $\mathbb{R}$ à valeurs dans $\mathbb{R}^n$. Une bonne connaissance de ces ensembles passe par l'étude de la surface constituant leur bord. Nous donnerons une paramétrisation de cette surface, et nous démontrerons, sous certaines hypothèses, des résultats de régularité de cette paramétrisation connaissant la régularité de F . Ces résultats ont été améliorés quand la fonction F vérifie certaines propriétés.

Un deuxième problème qui a suscité la curiosité de quelques auteurs ces dernières années, est celui de l'approximation de zonoïdes par des zonotopes. Ils s'agit de déterminer le nombre de segments permettant de définir un zonotope qui approche, au sens de la métrique de Hausdorff, un zonoïde à ε près. C'est un problème qui a fait l'objet de diverses études émanant de différentes théories. Dans ce cadre, nous citons les articles de McMullen [McM84], Gordon [Gor85] ou Bourgain, Lindenstrauss et Milman [BLM89] et entres autres.

Dans le chapitre 4, nous avons étudié le problème de l'approximation de zonoïdes par des zonotopes dans le cas de zonoïdes associés à une fonction: nous avons donné des méthodes de construction de suites de zonotopes qui convergent vers un zonoïde donné; nous avons pu évaluer pour chaque zonotope approximant, l'erreur d'approximation, et l'ordre de convergence de ces suites (en fonction du nombre de segments). Comme conséquence, nous avons donné le nombre de segments suffisant pour définir un zonotope approchant le zonoïde à ε près.

Le dernier chapitre est indépendant des quatre premiers. Il s'agit d'une application industrielle qui est la représentation de certaines tables utilisées dans l'industrie du pétrole. Nous avons utilisé le produit tensoriel de B-splines régularisées pour la représentation des tables. Cette méthode a donné des résultats meilleurs que ceux donnés par des méthodes existantes. D'une manière annexe, cette méthode a permis de trouver quelques incohérences dans le raccordement entre certaines parties des tables.

Chapitre 1

Zonotopes

Introduction

Les zonotopes sont des polytopes ayant des propriétés géométriques qui font d'eux des objets "simples". Nous entendons par simple, le fait que ce sont des polytopes dont les faces et les sommets peuvent être caractérisés explicitement et dont les volumes peuvent être calculés. La notion de zonotope a fait l'objet de plusieurs études menées par différents auteurs (Coxeter, Shepard, Bolker, McMullen,...). Diverses approches ont été utilisées dans ces études: combinatoire des polytopes, théorie de la mesure,... Dans la pratique, les zonotopes ont été introduits dans plusieurs domaines; citons par exemple, l'économie, la chimie, l'automatique,....

Dans les parties 1, 2 et 3 nous présentons un rappel des principaux résultats: la première partie sera consacrée aux propriétés des zonotopes. Dans la deuxième partie, nous introduisons, dans le cas de dimension 3, la notion de diagramme projectif qui permet de faire le lien entre la distribution de points dans le plan et la structure du zonotope. Dans la troisième partie, nous donnons la relation existant entre la notion de zonotope et le problème d'arrangement d'hyperplans.

La quatrième partie sera consacrée à une application des zonotopes au problème du groupement par paquets de vecteurs dans l'espace.

Dans ce chapitre, nous considérons $\mathbb{R}^n$ l'espace vectoriel sur $\mathbb{R}$ de dimension n , muni du produit scalaire euclidien usuel noté $\langle ., . \rangle$, et la norme euclidienne $\|.\|$.

1. Définitions et propriétés

1. 1. Définitions et quelques propriétés des zonotopes

Pour cette partie, nous nous référons aux articles [Bol69],[She74] et [McM].

Définition 1. 1 (Somme de Minkowski) Soient A, B deux parties de $\mathbb{R}^n$. On appelle somme de Minkowski de A et B , l'opération définie par:

$$A + B = \{a + b \mid a \in A \text{ et } b \in B\}.$$

Définition 1. 2 Un zonotope est un polytope défini par la somme de Minkowski de segments de droite fermés.

On a la proposition suivante:

Proposition 1. 3 Un zonotope est un polytope convexe compact dans $\mathbb{R}^n$, à symétrie centrale.

Remarque 1: Le centre de symétrie d'un zonotope est donné par la somme des milieux de chaque segment.

Un zonotope dans $\mathbb{R}^2$ est appelé zonogone. De même, un zonotope dans $\mathbb{R}^3$ est appelé zonoèdre.

Soient $N_1, N_2, \dots, N_p$ des vecteurs de $\mathbb{R}^n$. On supposera que les p vecteurs N_i engendrent l'espace $\mathbb{R}^n$.

Définition 1. 4 On appelle zonotope associé aux vecteurs N_i , $i = 1, \dots, p$, l'ensemble:

$$Z = \{M = \sum_{i=1}^p \lambda_i N_i, \quad 0 \leq \lambda_i \leq 1, \quad i = 1, \dots, p\}.$$

Z est le zonotope défini par la somme des segments $[O, N_i]$.

Nous parlerons particulièrement de zonotopes donnés par la définition 1.4. Ceci n'est pas restrictif car tout zonotope dans $\mathbb{R}^n$ défini par la somme de segments quelconques peut être défini, moyennant une translation, par la somme de segments

ayant 0 comme extrémité. En effet si $[P_1, Q_1]$ et $[P_2, Q_2]$ sont deux segments; alors on a:

$$\begin{aligned} [P_1, Q_1] + [P_2, Q_2] &= \{(1 - \alpha)P_1 + \alpha Q_1 + (1 - \beta)P_2 + \beta Q_2 \mid \alpha, \beta \in [0, 1]\} \\ &= \{P_1 + P_2 + \alpha(Q_1 - P_1) + \beta(Q_2 - P_2), \mid \alpha, \beta \in [0, 1]\} \\ &= P_1 + P_2 + [O, Q_1 - P_1] + [O, Q_2 - P_2]. \end{aligned}$$

C'est à dire l'ensemble $[P_1, Q_1] + [P_2, Q_2]$ est le translaté par $P_1 + P_2$ de $[0, Q_1 - P_1] + [0, Q_2 - P_2]$.

De même on a:

$$[O, P] + [O, Q] = P + [O, -P] + [O, Q],$$

$[O, P] + [O, Q]$ est le translaté par P de $[O, -P] + [O, Q]$. Par conséquent nous supposons, dans toute la suite, qu'il existe un vecteur μ tel que

$$\langle N_i, \mu \rangle > 0 \quad \forall i \in \{1, \dots, p\}. \quad (1.1)$$

Si deux vecteurs quelconques N_i et N_j sont colinéaires alors:

$$[O, N_i] + [O, N_j] = [O, N_i + N_j].$$

Ainsi, s'il y a des vecteurs colinéaires parmi les N_i , ils peuvent être remplacés par le vecteur somme. Donc sans perte de généralité, on peut considérer que les vecteurs N_i , $i = 1, \dots, p$ sont deux à deux non colinéaires.

De la définition précédente, on peut déduire:

Proposition 1. 5 Soit A l'application linéaire définie par:

$$A : (\lambda_1, \lambda_2, \dots, \lambda_p) \in \mathbb{R}^p \longrightarrow \sum_{i=1}^p \lambda_i N_i.$$

Z est l'image du cube unité par l'application linéaire A .

Démonstration : immédiate.

Proposition 1. 6 Soient T une application linéaire et Z un zonotope défini par la somme de Minkowski des segments $[O, N_i]$, $i = 1, \dots, p$; alors $T(Z)$ est le zonotope défini par la somme des segments $[O, T(N_i)]$, $i = 1, \dots, p$.

Démonstration : Immédiate.

Un point M de Z s'écrit sous la forme:

$$M = \sum_{i=1}^p \lambda_i N_i, \quad \lambda_i \in [0, 1].$$

En particulier, les sommets de Z peuvent s'écrire sous la forme:

$$\sum_{i=1}^p \varepsilon_i N_i, \quad \varepsilon_i \in \{0, 1\}.$$

D'où la proposition:

Proposition 1. 7 [McM71] *Un zonotope est l'enveloppe convexe des points*

$$\sum_{i=1}^p \varepsilon_i N_i, \quad \varepsilon_i \in \{0, 1\}. \quad (1.2)$$

Remarque 2: Tous les sommets de Z sont de la forme (1.2) mais ces points ne sont évidemment pas tous des sommets. Parmi les points donnés par (1.2), on trouve les sommets, des points du bord du zonotope qui ne sont pas des sommets et aussi des points de l'intérieur du zonotope.

1. 2. Sommets et faces d'un zonotope

Les caractéristiques géométriques du zonotope font de lui un polytope pour lequel les faces et les sommets ont une expression simple et explicite; comme nous le montrent les propositions données dans cette section.

Définition 1. 8 (fonction support) *Soit K un ensemble convexe borné de $\mathbb{R}^n$; la fonction support H_K de K est la fonction définie sur $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ par:*

$$H_K(u) = \sup_{x \in K} \langle x, u \rangle, \quad \forall u \in \mathbb{R}^n.$$

Dans le cas d'un zonotope, l'expression de la fonction support devient plus simple. En effet, si on désigne par x_+ la fonction définie sur $\mathbb{R}$ par

$$x_+ = \begin{cases} x & \text{si } x > 0, \\ 0 & \text{sinon,} \end{cases}$$

on a la proposition suivante:

Proposition 1. 9 *Soit Z le zonotope associé aux p vecteurs N_i , $i = 1, \dots, p$; alors pour tout $u \in \mathbb{R}^n$*

$$H_Z(u) = \sum_{i=1}^p \langle N_i, u \rangle_+.$$

Démonstration : Pour $u \in \mathbb{R}^n$;

$$\begin{aligned} H_Z(u) &= \sup_{x \in Z} \langle x, u \rangle \\ &= \sup_{\lambda \in [0,1]^p} \sum_{i=1}^p \lambda_i \langle N_i, u \rangle. \end{aligned} \quad (1.3)$$

Les valeurs des λ_i , $i = 1, \dots, p$ sont comprises entre 0 et 1, donc le maximum est atteint lorsque

$$\lambda_i = \begin{cases} 1 & \text{si } \langle N_i, u \rangle > 0, \\ \in [0, 1] & \text{si } \langle N_i, u \rangle = 0, \\ 0 & \text{si } \langle N_i, u \rangle < 0, \end{cases}$$

d'où le résultat. ■

Une face d'un polytope est définie comme étant l'intersection de ce polytope avec un de ses plans support. De la proposition précédente, nous déduisons l'expression de l'ensemble des points d'appui, c'est à dire des faces du zonotope.

Proposition 1. 10 *Soit u un vecteur de $\mathbb{R}^n$. L'ensemble des points d'appui de l'hyperplan de normale u sur Z s'écrit:*

$$F = \sum_{\langle N_i, u \rangle > 0} N_i + \sum_{\langle N_i, u \rangle = 0} [O, N_i].$$

F est une face de dimension k si l'espace engendré par les vecteurs N_i tels que $\langle N_i, u \rangle = 0$, est de dimension k . Une facette désignera une face de dimension $n - 1$.

Proposition 1. 11 *Un point $S = \sum_{j=1}^r N_{i_j}$, ($r \leq p$) de Z est un sommet si et seulement si il existe un vecteur $\eta \in \mathbb{R}^n$ tel que:*

$$\langle N_{i_j}, \eta \rangle > 0 \quad \forall j = 1, \dots, r,$$

et

$$\langle N_{i_j}, \eta \rangle < 0 \quad \forall j = r + 1, \dots, p,$$

où i_j , $j = 1, \dots, p$ est une permutation de $\{1, \dots, p\}$.

Démonstration : D'après la proposition précédente, l'ensemble des points d'appui de l'hyperplan de normale η s'écrit:

$$\sum_{\langle N_i, \eta \rangle > 0} N_i + \sum_{\langle N_i, \eta \rangle = 0} [O, N_i].$$

Pour que l'ensemble des points d'appui soit réduit à un seul élément il faut et il suffit que pour tout $i \in \{1, \dots, p\}$,

$$\langle N_i, \eta \rangle \neq 0.$$

Le point d'appui s'écrit dans ce cas:

$$\sum_{\langle N_i, u \rangle > 0} N_i. \quad \blacksquare$$

Le sommet défini par $\sum_{i=1}^p N_i$ est appelé point distal du zonotope.

Remarque 3: : Tout segment $[O, N_i]$ détermine une "zone" sur le bord du zonotope. Une zone correspondant à un segment, est un ensemble de facettes du zonotope ayant toutes une arête qui se déduit de ce segment par translation. Deux facettes adjacentes de la zone ont une telle arête en commun.

Dans la figure 1, nous donnons un exemple de zonoèdre avec ses vecteurs de base. La partie hachurée représente la zone associée au vecteur N_{i_0} .

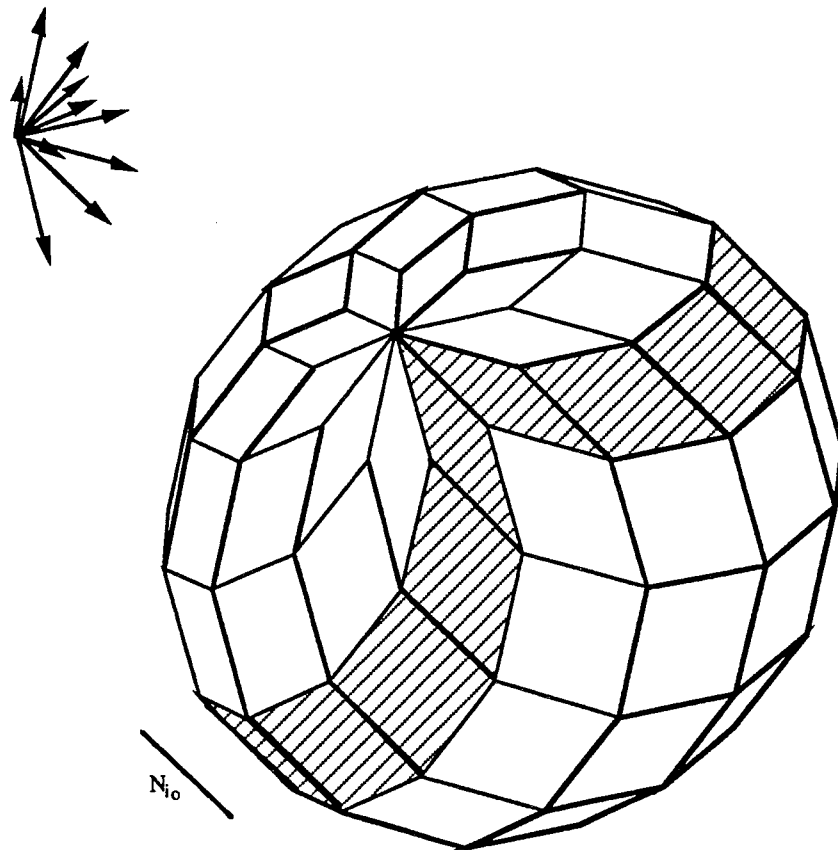


FIGURE .1. Exemple de zonoèdre et la zone associée à un segment.

D'après la proposition précédente, toute face F de Z est de la forme:

$$F = \sum_{\langle N_i, u \rangle = 0} [O, N_i] + \sum_{\langle N_i, u \rangle > 0} N_i.$$

Donc toute face de Z est la translatée d'une somme de Minkowski de segments, (i.e., d'un zonotope). D'où la proposition suivante:

Proposition 1. 12 *Si Z est un zonotope alors toutes ses faces sont aussi des zonotopes.*

La réciproque est aussi vraie; elle a été démontrée par Shephard dans [She67]. Dans cet article, l'auteur a donné la caractérisation suivante des zonotopes:

Proposition 1. 13 [She67] *Un polytope est un zonotope si et seulement si toutes ses faces de dimension 2 admettent une symétrie centrale.*

Remarques :

- (1) En dimension 2, tout polygone à symétrie centrale est un zonogone.
- (2) Dans le cas de dimension supérieure à 2, la symétrie n'est pas suffisante. Si on prend pour exemple l'ensemble

$$\{(x_1, x_2, x_3) \in \mathbb{R}^3 : |x_1| + |x_2| + |x_3| \leq 1\},$$

il est facile de voir qu'il admet 0 comme centre de symétrie mais toutes ses 2-faces sont des triangles. Donc il n'est pas un zonotope.

1. 3. Volume d'un zonotope

Soit $\mathcal{F}$ une famille finie de polytopes P_i , $i \in I$.

Définition 1. 14 *La famille de polytopes $\mathcal{F}$ est appelée complexe si*

- 1) $P_i \cap P_j$ est soit vide soit une face de P_i et de P_j .
- 2) toute face de P_i appartient à $\mathcal{F}$.

Soit $\mathcal{G}$ l'ensemble des éléments de la forme :

$$[O, N_{j_1}] + [O, N_{j_2}] + \cdots + [O, N_{j_n}], \quad 1 \leq j_1 < \cdots < j_n \leq p,$$

tels que les vecteurs $N_{j_1}, \dots, N_{j_n}$ sont linéairement indépendants.

On appelle *cellule* le polytope défini par

$$[O, N_{i_1}] + [O, N_{i_2}] + \cdots + [O, N_{i_k}] + \varepsilon_{i_{k+1}} N_{i_{k+1}} + \cdots + \varepsilon_{i_p} N_{i_p} \quad (1.4)$$

où $k \leq n$ et $\varepsilon_{i_r} \in \{0, 1\}$ pour $r = k + 1, \dots, p$ et les N_{i_j} , $j = 1, \dots, p$, étant linéairement indépendants.

Soit $\mathcal{F}$ un complexe tel que chaque P_i est une cellule et

$$\cup_{i \in I} P_i = Z. \quad (1.5)$$

L'existence d'un complexe vérifiant (1.5) a été démontrée par Shephard dans [She74]. Il est clair cependant qu'un tel complexe n'est pas unique.

Définition 1. 15 On appelle *dissection* de Z , l'ensemble $\mathcal{C}(Z) \subset \mathcal{F}$ défini par:

$$\mathcal{C}(Z) = \{P_i \in \mathcal{F} \mid \text{int}(P_i) \neq \emptyset\},$$

où $\text{int}(P)$ désigne l'intérieur de P .

Pour tout complexe $\mathcal{F}$ vérifiant (1.5), on a $\cup_{P_i \in \mathcal{C}(Z)} P_i = Z$ et on a le théorème suivant (cf [She74]):

Théorème 1.16 [She74] Il existe une correspondance bijective entre $\mathcal{C}(Z)$ et $\mathcal{G}$.

Or chaque élément de $\mathcal{G}$ a pour volume la valeur

$$|\det(N_{i_1}, N_{i_2}, \dots, N_{i_n})|.$$

Par conséquent, de la proposition précédente, on obtient l'expression du volume d'un zonotope:

Corollaire 1. 17 Soit Z le zonotope défini par la somme de Minkowski des segments $[O, N_i]$, $i = 1, \dots, p$; alors

$$\text{Vol}(Z) = \sum_{1 \leq i_1 < \dots < i_n \leq p} |\det(N_{i_1}, N_{i_2}, \dots, N_{i_n})|.$$

Comme autre conséquence de la proposition précédente, on a aussi

Corollaire 1. 18 Soit M un point du zonotope Z ; alors il existe J , un sous ensemble de n éléments de $\{1, \dots, p\}$, tel que

$$M = \sum_{i \notin J} \varepsilon_i N_i + \sum_{i \in J} \lambda_i N_i,$$

où $\varepsilon_i \in \{0, 1\}$ pour $i \notin J$ et $\lambda_i \in [0, 1]$ pour $i \in J$.

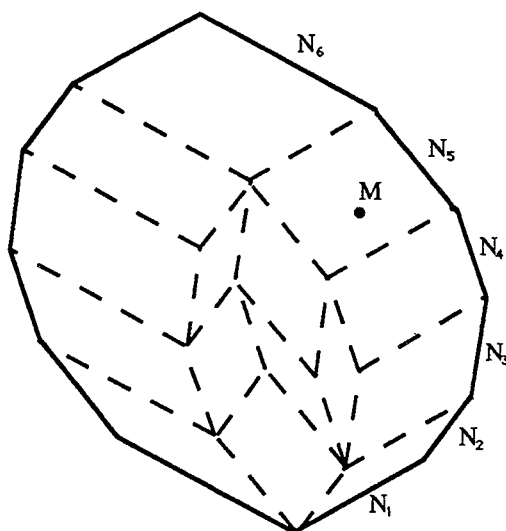


FIGURE .2. Un exemple de dissection d'un zonogone.

Dans la figure 2, nous donnons un exemple de dissection d'un zonogone. Le point M du zonogone peut être écrit sous la forme:

$$M = N_2 + N_3 + N_4 + \lambda_1 N_1 + \lambda_5 N_5, \quad \lambda_1, \lambda_5 \in [0, 1].$$

2. Premier diagramme projectif

Dans cette partie, nous nous placerons dans $\mathbb{R}^3$.

2. 1. Définitions et propriétés

Le premier diagramme projectif a été introduit par Coxeter, dans [Cox67], pour étudier des propriétés topologiques des zonoèdres et les classer.

Le premier diagramme projectif, noté P.D.P., est défini de la manière suivante:

Définition 1. 19 On appelle diagramme projectif associé aux vecteurs N_i , $i = 1, \dots, p$, le diagramme dans le plan construit de la manière suivante: soit P_μ le plan de $\mathbb{R}^3$ défini par:

$$P_\mu = \{x \in \mathbb{R}^3 \quad : \quad \langle x, \mu \rangle = 1\}.$$

- Pour chaque vecteur N_i , on calcule le point M_i : l'intersection de la demi droite $\mathbb{R}_+ N_i$ avec P_μ .
- On joint les p points M_i par toutes les droites possibles.

Pour les distinguer des autres droites du plan P_μ , les droites joignant les points M_i seront appelées droites du P.D.P.

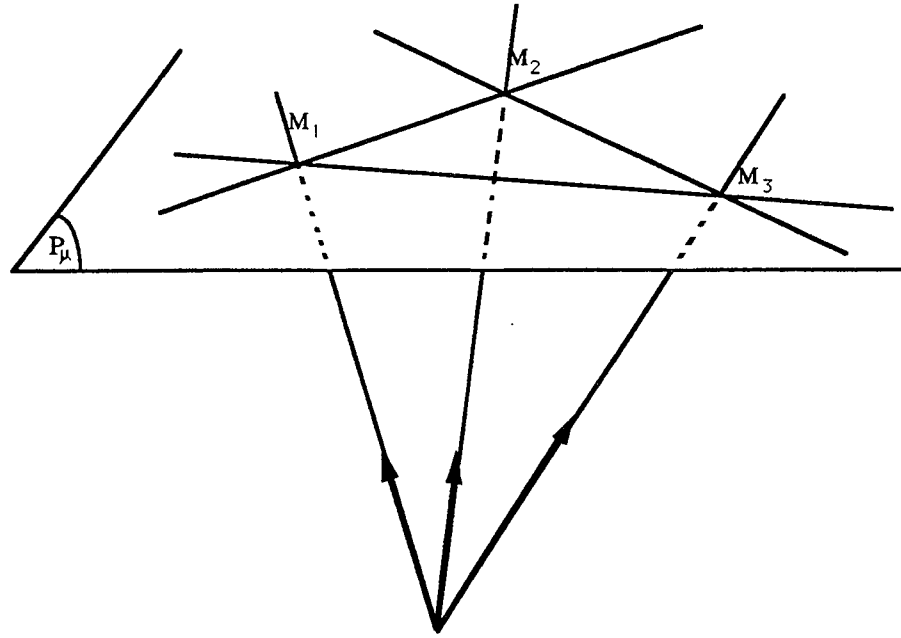


FIGURE .3. Construction du P.D.P..

Interprétation

Considérons un vecteur η de $\mathbb{R}^3$ ($\eta \neq \mu$) et Π_η le plan de normale η passant par l'origine. Soit Δ_η la droite intersection des plans Π_η et P_μ .

De façon triviale, si Δ_η contient M_i alors Π_η contient N_i , donc le type d'intersection de Z avec ses deux plans support dans les directions η et $-\eta$ est déterminé par le nombre de points M_i appartenant à la droite Δ_η . Ainsi:

- si Δ_η ne contient aucun point M_i , alors l'ensemble des points d'appui dans les directions η et $-\eta$ est réduit à un sommet.
- si Δ_η contient un unique point, alors l'intersection est réduite à une arête.
- si Δ_η contient au moins deux points, alors l'intersection est une facette.

A chaque droite du P.D.P. correspond une paire de facettes du zonotope qui sont opposées par rapport au centre de symétrie.

Pour lire la structure du zonotope dans le P.D.P., on utilise les trois propriétés suivantes:

Propriété 1. 20 *Si une droite Δ ne contenant aucun point M_i , partage l'ensemble des points du P.D.P. en deux ensembles E_1 et E_2 de points qui sont de part et d'autre de la droite; alors les deux points*

$$\sum_{M_i \in E_1} N_i \quad \text{et} \quad \sum_{M_i \in E_2} N_i,$$

sont deux sommets symétriques de Z .

Propriété 1. 21 Si une droite Δ contient m points M_i et partage les autres points en deux ensembles E_1 et E_2 alors

$$F_1 = \sum_{M_i \in E_1} N_i + \sum_{M_i \in \Delta} [O, N_i] \quad \text{et} \quad F_2 = \sum_{M_i \in E_2} N_i + \sum_{M_i \in \Delta} [O, N_i].$$

sont deux faces symétriques de Z .

Propriété 1. 22 Une zone correspondant à un segment $[O, N_i]$ est donnée par toutes les droites du P.D.P. qui passent par le point M_i .

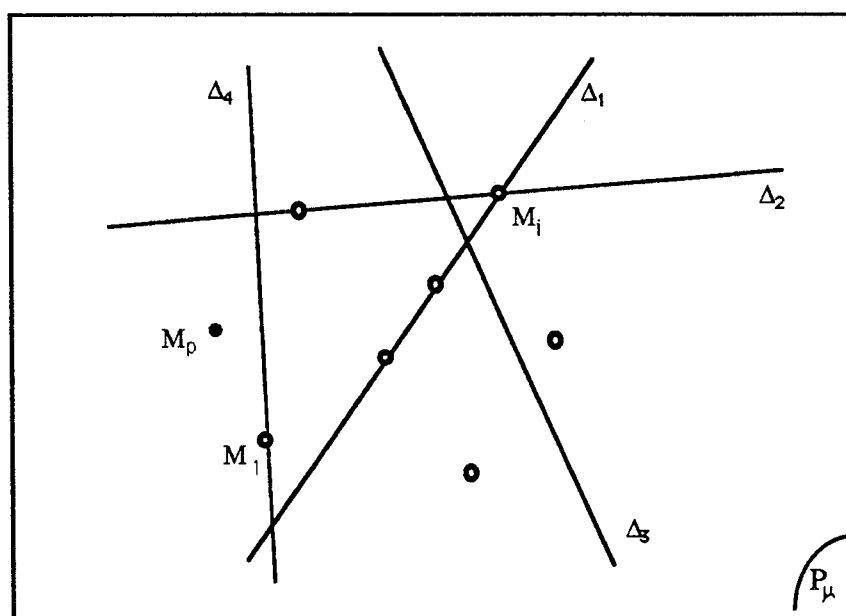


FIGURE 4. P.D.P. et faces du zonoèdres.

Dans la figure 4, Δ_1 correspond à un couple de facettes hexagonales (3 points sur Δ_1), Δ_2 correspond à un couple de facettes qui sont des parallélogrammes, Δ_3 ne contient aucun point donc correspond à un sommet et Δ_4 correspond à une arête.

Remarque 4: Quand μ est donné, à chaque zonoèdre correspond un unique diagramme projectif, mais pour chaque diagramme projectif on peut faire correspondre toute une famille de zonoèdres: zonoèdres définis par des segments de même directions mais de longueurs différentes. C'est ce que Coxeter appelle: famille de zonoèdres isomorphes.

Avec la notion de premier diagramme projectif, nous remarquons que la structure topologique du zonoèdre est liée à la situation des points dans le plan. De la propriété 1.21, on peut déduire la proposition suivante:

Proposition 1. 23 Si tous les points M_i du plan P_μ sont en position générale (3 points quelconques ne sont pas alignés) alors le zonoèdre Z admet $2C_p^2$ facettes qui sont toutes des parallélogrammes.

Dans le cas général, si on classe les facettes par le nombre de segments m qui les déterminent, Fedorov a démontré la relation suivante entre le nombre de chaque type de facettes et le nombre de segments:

Proposition 1. 24 (Formule de Fedorov 1953) *Soit Z un zonoèdre défini par la somme de p segments. Notons f_m le nombre de paires de facettes définies par la somme de m segments; alors*

$$2.(p-1)p = 1.2f_2 + 2.3f_3 + \dots + (m-1)mf_m + \dots + (p-2)(p-1)f_{p-1}.$$

• P.D.P. et construction de zonoèdres

On a vu qu'à chaque droite du P.D.P. joignant 2 points M_i correspond deux facettes translatées respectivement par les vecteurs correspondant aux points qui sont d'un côté ou de l'autre de cette ligne. Cela permet de concevoir un algorithme de détermination de toutes les facettes du zonoèdre. L'algorithme peut s'écrire sous la forme suivante:

Pour chaque droite L_i du P.D.P. contenant deux points, soit M_k, M_l :

- calculer $n_i = N_k \wedge N_l$,
- déterminer $S_i^1 = \sum_{\langle N_j, n_i \rangle > 0} N_j$ et $S_i^2 = \sum_{\langle N_j, n_i \rangle < 0} N_j$,
- la face F_i^1 (resp. F_i^2) est le zonogone défini par les vecteurs $[O, N_j]$ pour lesquels $M_j \in L_i$, et translaté par S_i^1 (resp. S_i^2).

Cet algorithme est intuitif, à partir du P.D.P., mais son coût est en $O(p^3)$. Pour remédier à cela, Szafran a proposé une méthode incrémentale basée sur les propriétés du bord du zonoèdre, particulièrement sur la notion de zone. Son coût est en $O(p^2)$. Une description détaillée se trouve dans [Sza91].

2. 2. Exemples

Dans cette partie, nous donnons quelques exemples de zonoèdres.

• Exemple 1

Un prisme est un zonoèdre dont $p - 1$ des vecteurs sont dans un hyperplan. Le premier diagramme projectif d'un prisme présente $(p - 1)$ points alignés; voir figure 5.

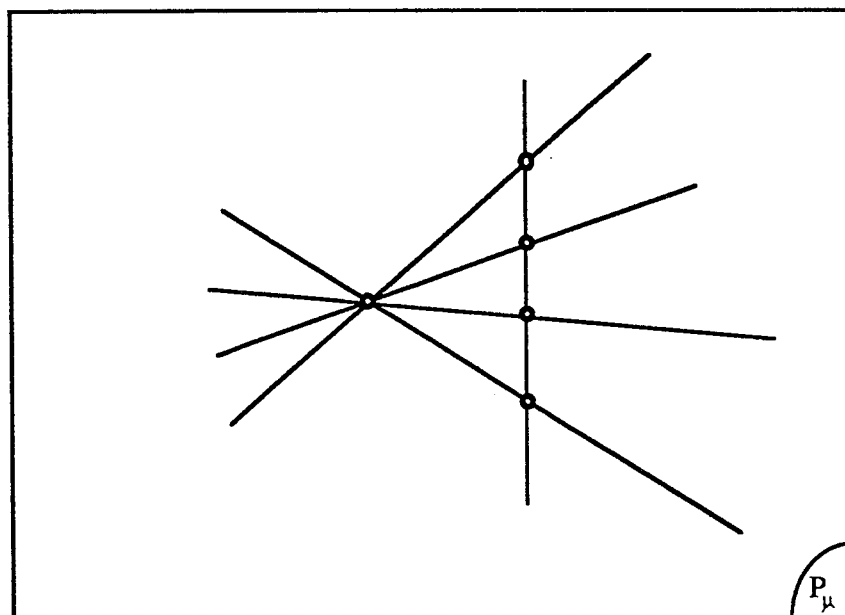


FIGURE .5. Exemple de P.D.P. d'un prisme.

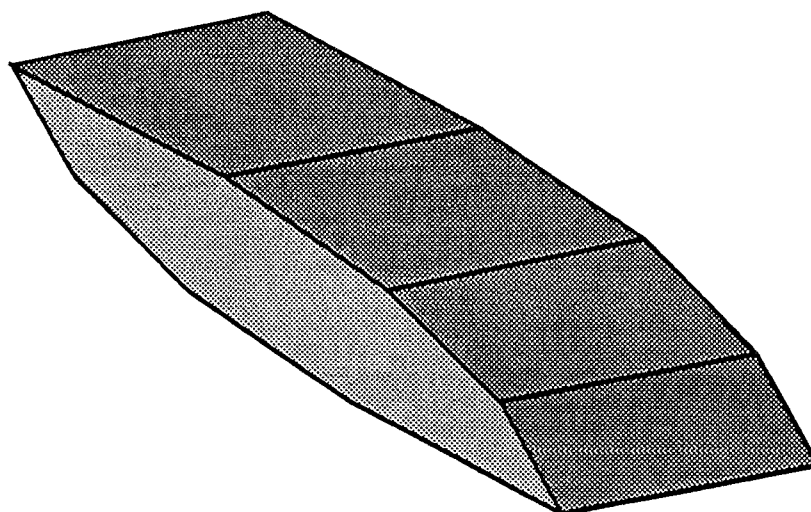


FIGURE .6. Un prisme.

Un prisme a 2 facettes définies par la somme de $p - 1$ segments, et $2(p - 1)$ facettes qui sont des parallélogrammes.

• **Exemple 2** (Dodécaèdre)

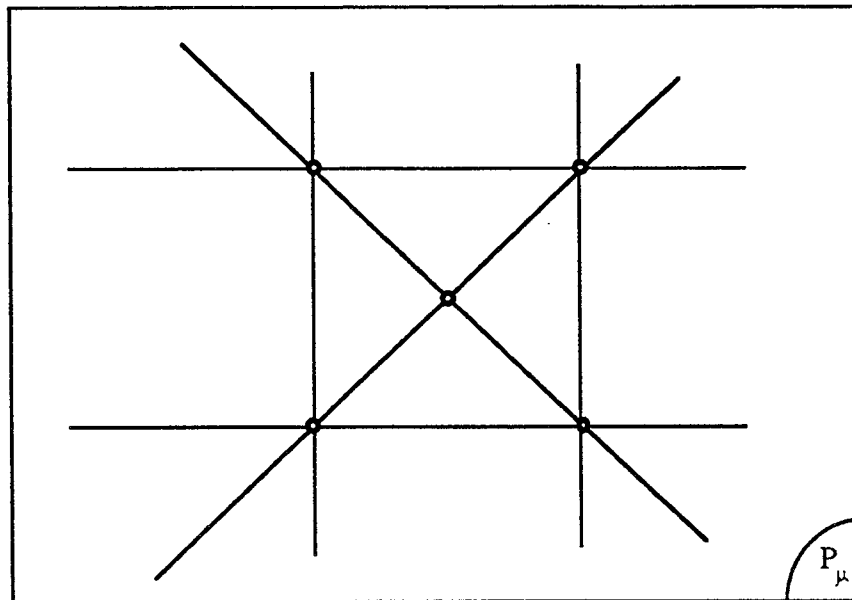


FIGURE .7. Exemple P.D.P. du dodécaèdre

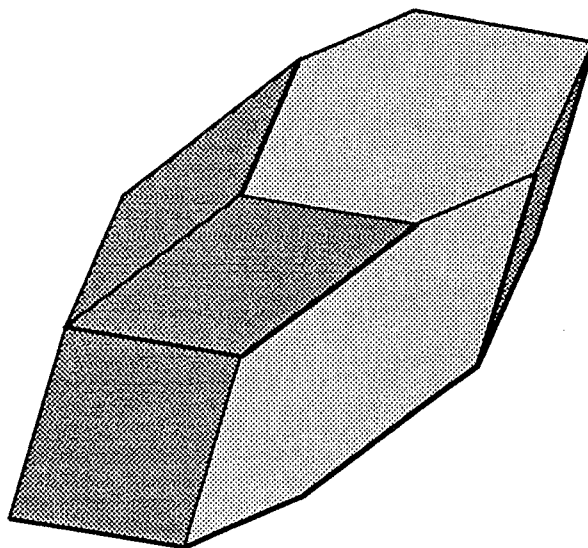


FIGURE .8. Dodécaèdre.

Le dodécaèdre est un zonoèdre qui a 4 facettes hexagonales et 8 facettes qui sont des parallélogrammes.

• Exemple 3

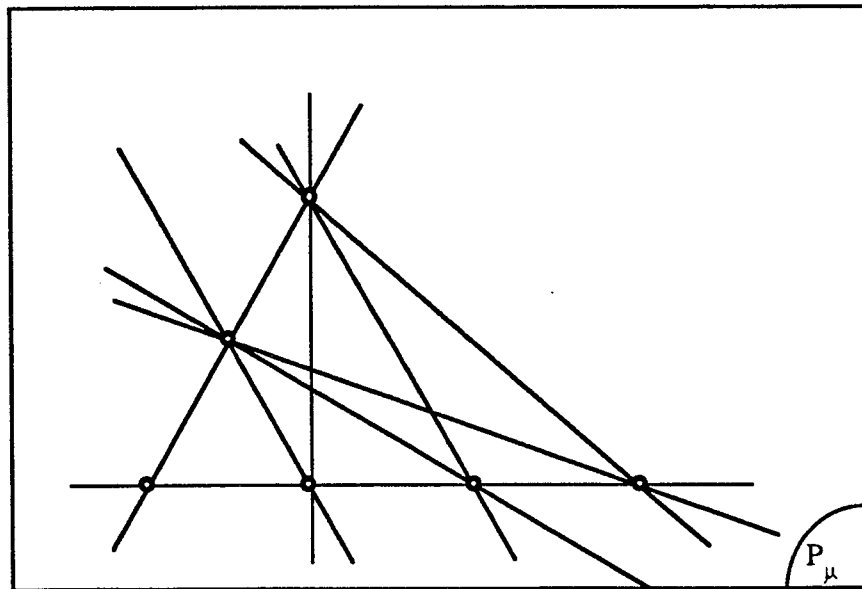


FIGURE .9. Exemple de P.D.P..

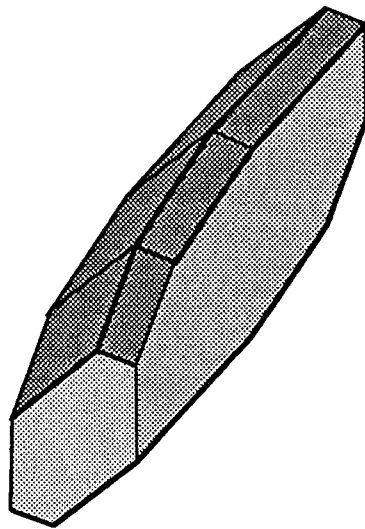


FIGURE .10. Exemple de zonoèdre.

Dans la figure 9, le P.D.P, représente:

- Un couple de facettes somme de 4 segments correspondant aux 4 points alignés.
- Un couple de facettes somme de 3 segments correspondant aux 3 points alignés.
- 6 couples de facettes qui sont des parallélogrammes.

3. Zonoèdre et arrangement de droites dans le plan

Pour $\mu \in \mathbb{R}^3$, on considère le plan P_μ défini dans la section 2. Soient N_i , $i = 1, \dots, p$ des vecteurs de $\mathbb{R}^3$ vérifiant pour tout $i < j$, les 3 vecteurs N_i , N_j et μ sont linéairement indépendants. Soit Z le zonotope associé aux vecteurs N_i .

• Construction de l'arrangement :

A chaque vecteur N_i , on associe le plan h_i de normale N_i passant par 0. On note h_i^+ (resp. h_i^-) le demi espace défini par $x \in \mathbb{R}^3$ tel que $\langle x, N_i \rangle > 0$ (resp. $\langle x, N_i \rangle < 0$). Soient g_i , $i = 1, \dots, p$, la droite d'intersection des plans h_i et P_μ (les droites g_i existent puisqu'on a supposé que pour tout $i = 1, \dots, p$, $N_i \neq \mu$). Les droites g_i , $i = 1, \dots, p$ partitionnent le plan P_μ .

• Interprétation :

La correspondance entre l'arrangement de droites et la structure du zonoèdre est déterminée de la manière suivante:

Dans toute la suite u désigne un vecteur de $\mathbb{R}^3$ appartenant à P_μ . Notons Π_u le plan de normale u passant par 0.

Remarquons que

$$u \in h_k \quad (i.e. \quad \langle N_k, u \rangle = 0) \quad \Longleftrightarrow \quad N_k \in \Pi_u (i.e. \quad \langle N_k, u \rangle = 0). \quad (1.6)$$

Soient F_u^1 et F_u^2 les deux faces d'appui des plans de normales respectives u et $-u$ sur la surface de Z . D'après la proposition 1.10, F_u^1 et F_u^2 s'écrivent:

$$\begin{aligned} F_u^1 &= \sum_{\langle N_i, u \rangle > 0} N_i + \sum_{\langle N_i, u \rangle = 0} [O, N_i], \\ F_u^2 &= \sum_{\langle N_i, u \rangle < 0} N_i + \sum_{\langle N_i, u \rangle = 0} [O, N_i]. \end{aligned}$$

Donc en utilisant la relation (1.6) et la proposition 1.10, on a :

- $u \in g_k$ pour un certain k est équivalent au fait que F_u^1 et de F_u^2 ont une arête équipolente à $[O, N_k]$.
- $u \in g_j \cap g_k$ est équivalent au fait que les faces F_u^1 et de F_u^2 ont deux arêtes équipolentes à $[O, N_j]$ et $[O, N_k]$ respectivement.

Plus généralement, on a la propriété:

Si exactement m droites g_j passent par le même point u alors le plan de normale u s'appuie sur une face qui est à une translation près la somme des m segments $[O, N_j]$ qui correspondent à ces droites.

- u n'appartenant à aucune des droites g_i est équivalent, d'après (1.6), au fait que Π_u ne contient aucun des vecteurs N_i , et donc, d'après la proposition 1.11, au fait que les faces d'appui F_u^1 et F_u^2 sont réduites à deux sommets.

On désigne par sommet du partitionnement l'intersection d'au moins deux droites non confondues et par une arête du partitionnement, tout segment joignant deux sommets (ceux ci étant exclus).

On a donc une relation duale entre le partitionnement du plan par les droites g_i et le zonoèdre Z . En effet:

- A tout sommet du partitionnement correspond une facette de Z .
- A toute arête du partitionnement correspond une arête de Z .
- A toute droite g_i correspond la zone associée au segment $[O, N_i]$.
- A toute cellule du partitionnement correspond un sommet de Z .

Remarque 5: Pour chaque point P de $\mathbb{R}^2 \setminus \{0\}$, on associe la droite (duale) définie par

$$Dual(P) = \{x \in \mathbb{R}^2 \mid \langle x, P \rangle = 1\}.$$

On suppose que $N_i \neq \mu$, $i = 1, \dots, p$. En identifiant le plan P_μ avec $\mathbb{R}^2$, on a une relation duale entre le premier diagramme projectif et l'arrangement de droites défini ci-dessus.

Dans la figure 11, par le point p_1 passe 4 droites; donc il correspond à un couple de facettes du zonoèdre. Ces facettes sont définies par la translatée de la somme de 4 segments. Aux points p_2, p_3, p_4 et p_5 correspondent des facettes qui sont des parallélogrammes. On remarque que le premier diagramme projectif et l'arrangement de droites associé à un prisme ont la même allure. Dans la figure 12, nous donnons l'arrangement de droites associé à un dodécaèdre.

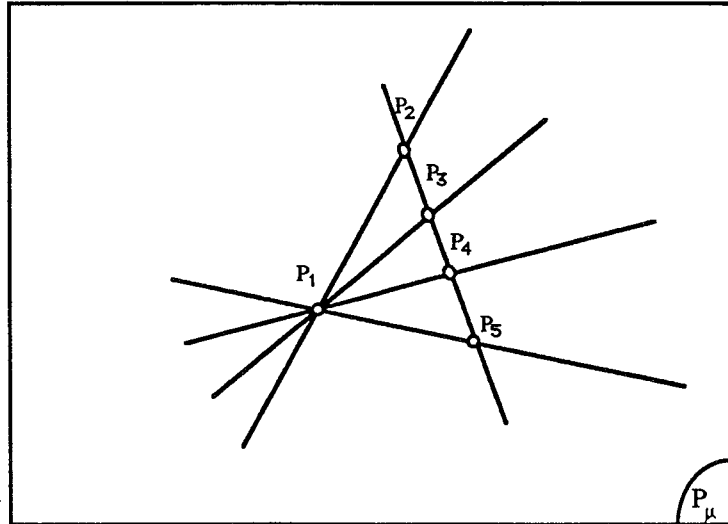


FIGURE .11. Arrangement de droites associé à un prisme.

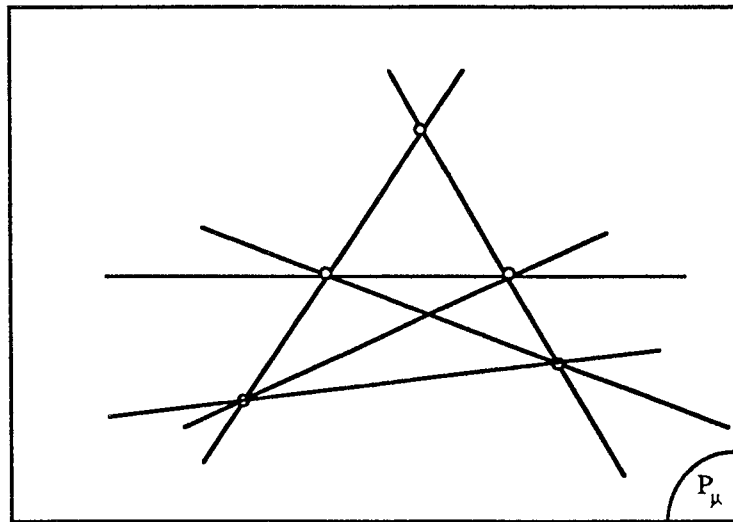


FIGURE .12. Arrangement de droites associé à un dodécaèdre.

4. Zonotopes et problème de groupement de vecteurs

4. 1. Critère de groupement et problème d'optimisation

Soient $\tilde{N}_i, i = 1, \dots, q$ (q grand) des vecteurs de $\mathbb{R}^n$, que nous voulons grouper en un nombre p de paquets. Nous avons choisi le critère suivant:

- grouper les vecteurs $\tilde{N}_i$ en p paquets ($p < q$) de sorte que le zonotope associé aux vecteurs résultant de chaque paquet, soit de volume maximum.

Notons $N_i, i = 1, \dots, p$ le vecteur résultant du paquet i . Il est défini par:

$$N_i = \sum_{j=1}^q \alpha_{ij} \tilde{N}_j, \quad (1.7)$$

où

$$\alpha_{ij} = \begin{cases} 1 & \text{si } \tilde{N}_j \text{ appartient au paquet } i, \\ 0 & \text{sinon.} \end{cases} \quad (1.8)$$

Le fait que chaque vecteur $\tilde{N}_j$ n'apparaît que dans un seul paquet, entraîne

$$\sum_{i=1}^p \alpha_{ij} = 1, \quad j = 1, \dots, q. \quad (1.9)$$

Soit $Z(\alpha)$ le zonotope associé aux vecteurs $N_i, i = 1, \dots, p$ (dépend de α). Le problème peut être écrit sous la forme:

$$\left\{ \begin{array}{l} \max \quad \text{Vol}(Z(\alpha)) \\ \sum_{j=1}^q \alpha_{ij} \tilde{N}_j = N_i \quad i = 1, \dots, p, \\ \alpha_{ij} \in \{0, 1\} \quad i = 1, \dots, p, \quad j = 1, \dots, q, \\ \sum_{i=1}^p \alpha_{ij} = 1, \quad j = 1, \dots, q. \end{array} \right. \quad (P)$$

On a donc un problème d'optimisation combinatoire. Une des méthodes pour le résoudre est celle du recuit simulé.

4. 2. Méthode de recuit simulé

L'algorithme de recuit simulé permet d'atteindre le minimum d'une fonction sur un ensemble fini pouvant être de grande taille. La méthode utilisée, adopte

une idée issue de la mécanique statistique: il s'agit de simuler le refroidissement d'un système pour atteindre les états d'énergie minimum.

Cet algorithme a l'ambition d'éviter la convergence vers le premier minimum local rencontré et une trop forte dépendance de l'état initial. Pour cela, au cours des itérations, alors qu'on accepte toujours que la fonction décroisse, on accepte qu'elle ne croisse qu'avec une probabilité non nulle, contrôlée par un paramètre appelé "température". L'algorithme commence avec une valeur de température grande et décroissant lentement vers 0, simulant ainsi le processus de refroidissement.

Le cadre du recuit simulé est défini de la manière suivante: on considère un ensemble fini d'états E sur lequel on définit une structure de voisinage (un changement élémentaire, bien défini, permettant de passer d'un état à un autre).

Soit $U : E \rightarrow \mathbb{R}$ une fonction à minimiser (appelée l'énergie du système) et $(X_i), i = 1, \dots, N$ les éléments de E . La quantité $U(X_i)$ s'appelle l'énergie à l'état X_i . L'algorithme dit de Metropolis s'écrit de la manière suivante:

On suppose que la notion de voisinage est bien définie sur E et que tout élément de E a, au moins, un voisin.

Algorithme

- Soit T_0 la température initiale donnée (grande), élément de $\mathbb{R}_+^*$.
- Partant d'un état initial e_0 , on tire au hasard suivant une loi uniforme l'un des éléments du voisinage $V(e_0)$ de e_0 . Si le nouvel état fait diminuer la fonction alors on accepte le changement. Sinon, (i.e., s'il fait augmenter la fonction), on l'accepte avec la probabilité $e^{-\Delta U/T_0}$, où ΔU représente l'augmentation.
- Une fois e_1 déterminé, on diminue la température: T_0 devient T_1 ($T_1 < T_0$) suivant un schéma prédéfini et on recommence le processus.
- L'algorithme est arrêté, lorsque son état d'équilibre est atteint, ce qui se traduit pratiquement par le fait qu'il n'est pas possible d'améliorer l'optimum.

* Une variante de l'algorithme consiste à faire tourner l'algorithme à température constante jusqu'à l'obtention d'une position d'équilibre. Puis, on baisse T et on recommence. C'est ce qui est appelé l'algorithme homogène.

Remarques :

- A température élevée, presque tous les candidats voisins sont admis. Si $T = +\infty$, alors l'algorithme est celui de promenade aléatoire sur les états de E .
- A température faible, T proche de 0, presque seulement les candidats qui font décroître la fonction sont admis. Si $T = 0$, l'algorithme est celui de la descente systématique de U sur E (dans ce cas, le résultat peut être un minimum local).
- Supposons qu'à l'étape k , l'algorithme soit à l'état e_k . Soit e un voisin de e_k . Plus l'accroissement $\Delta U = U(e) - U(e_k)$ est grand moins il aura de chance d'être

admis; sauf quand l'accroissement de U est compensé par T (la probabilité d'admission est $e^{-\Delta U/T}$).

La convergence de l'itération e_k , définie dans cet algorithme, vers un état réalisant le minimum d'énergie, a été démontrée par (Hajek 85) sous les hypothèses suivantes:

- pour tout couple d'états (e_i, e_j) , il existe un chemin de e_i à e_j (connexité).
- pour tout réel r et pour tout couple d'états e_i, e_j , s'il existe un chemin de e_i à e_j tel que tous les états intermédiaires ont une énergie inférieure à r , alors il existe un chemin de e_j vers e_i ayant la même propriété.

Notons:

$$E^* = \{e \in E \mid U(e) = \min_{f \in E} U(f)\},$$

si T_n décroît vers 0 quand $n \rightarrow \infty$, alors

$$\lim_{n \rightarrow \infty} \text{Prob}(e_n \in E^*) = 1,$$

si et seulement si

$$\sum_{k=1}^{\infty} \exp\left(-\frac{D}{T_k}\right) = \infty.$$

D est une constante qui dépend de la fonction à minimiser et qui, en pratique, n'est pas connue.

Remarque 6: Si T_n décroît selon le schéma

$$T_{n+1} = \frac{D}{\log(n+1)},$$

alors l'algorithme converge. Mais ce schéma n'est pas utilisable parce que la convergence est trop lente.

En pratique, d'autres schémas ont donné de bonnes performances. Par exemple,

$$T_{n+1} = \alpha T_n \quad \text{où} \quad 0.95 \leq \alpha \leq 0.99. \quad \rightarrow$$

4. 3. Application de l'algorithme au problème (P)

Dans notre cas, un état est donné par la matrice α de taille $p \times q$ dont les composantes appartiennent à $\{0, 1\}$ et vérifiant (1.9).

On dira que deux états sont voisins s'ils ont $(q - 1)$ colonnes identiques. Puisque chaque colonne de α comporte une seule fois la valeur 1, les autres valeurs étant des 0, pour passer d'un état à un état voisin, il suffit de changer dans une colonne la place de la valeur 1.

L'énergie dans notre cas est la fonction définie de E vers $\mathbb{R}$ et qui à $\alpha \rightarrow -\text{Vol}(Z(\alpha))$.

4.3.1. Exemple dans $\mathbb{R}^2$

• Exemple 1

On dispose des 6 vecteurs suivants.

$$\begin{pmatrix} 1.00 & 1.00 & 1.00 & 0.00 & 1.00 & 2.00 \\ 1.00 & 2.00 & 0.00 & 1.00 & 3.00 & 1.00 \end{pmatrix}.$$

On veut les partager en 3 groupes. Chaque vecteur est représenté par son numéro de colonne.

Etat initial:

Paquet	vecteurs
1	2 4 6
2	
3	1 3 5

volume à l'état initial: 0.000

Etat final :

Paquet	vecteurs
1	1 2
2	4 5
3	3 6

volume maximum trouvé est 23.00.

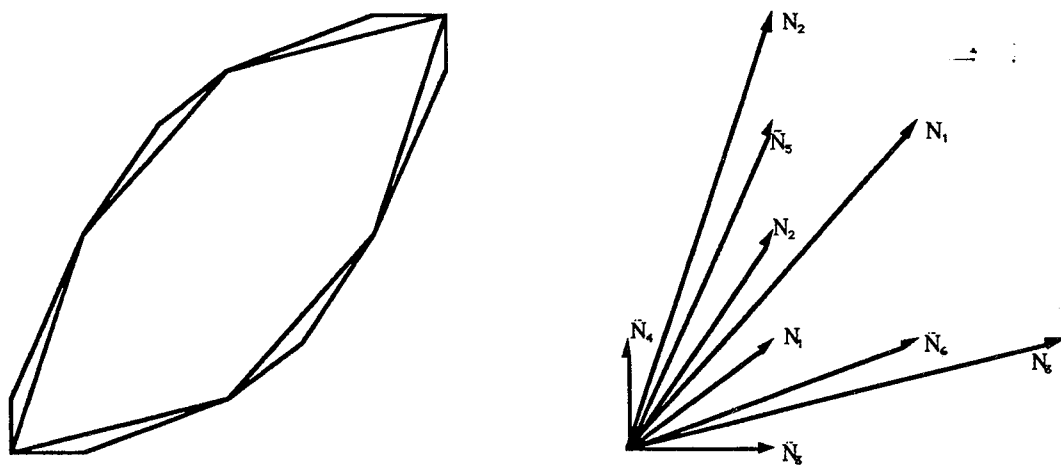


FIGURE .13. Zonogones associés à $\tilde{N}_i$ et N_i . $T_0 = 1$.

• Exemple 2

Dans cet exemple, nous avons tiré aléatoirement 60 vecteurs dans $\mathbb{R}^2$, et le but est de les regrouper en 6 paquets.

Nous avons remarqué que pour une valeur initiale du paramètre $T = 0.1$, la convergence n'a pas eu lieu vers un maximum global de la fonction volume. Celui-ci a pu être atteint pour des paramètres $T \geq 1$. Dans les deux figures suivantes, nous donnons les deux résultats. Dans chaque cas, nous avons superposé les deux zonogones, celui associé à la famille $\tilde{N}_i$, $i = 1, \dots, q$ et celui associé à la famille N_i , $i = 1, \dots, p$.

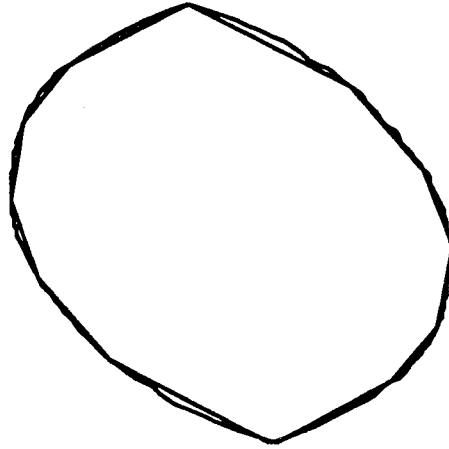


FIGURE .14. Zonogones associés à $\tilde{N}_i$ et N_i . $T_0 = 0.1$, $Vol(Z(\alpha)) = 3208$

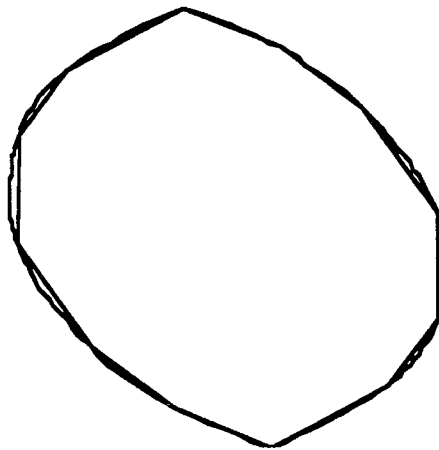


FIGURE .15. Zonogones associés à $\tilde{N}_i$ et N_i . $T_0 = 1$, $Vol(Z(\alpha)) = 3213$

• **Exemple 3** Nous avons choisi 6 vecteurs dans $\mathbb{R}^3$. Chaque vecteur a été perturbé 10 fois, nous avons ainsi obtenu 60 vecteurs dans $\mathbb{R}^3$. En résolvant le problème (P), nous avons obtenu le résultat illustré sur le diagramme projectif ($z = 1$) associé aux vecteurs; voir figure 16 (chaque groupe correspond à un motif). Dans la figure 17, on représente à gauche le zonoèdre associé aux vecteurs $\tilde{N}_i$, $i = 1, \dots, q$ ($q = 60$); et à droite le zonoèdre associé aux vecteurs N_j , $j = 1, \dots, p$ ($p = 6$).

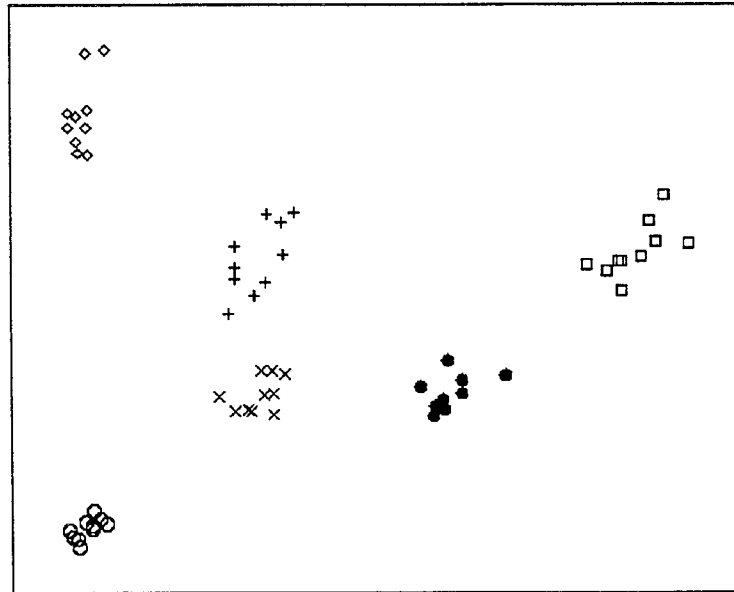


Figure 16 Vue des vecteurs dans le plan $z = 1$; paramètre $T_0 = 0.1$, $Vol(Z) = 53231.41$.

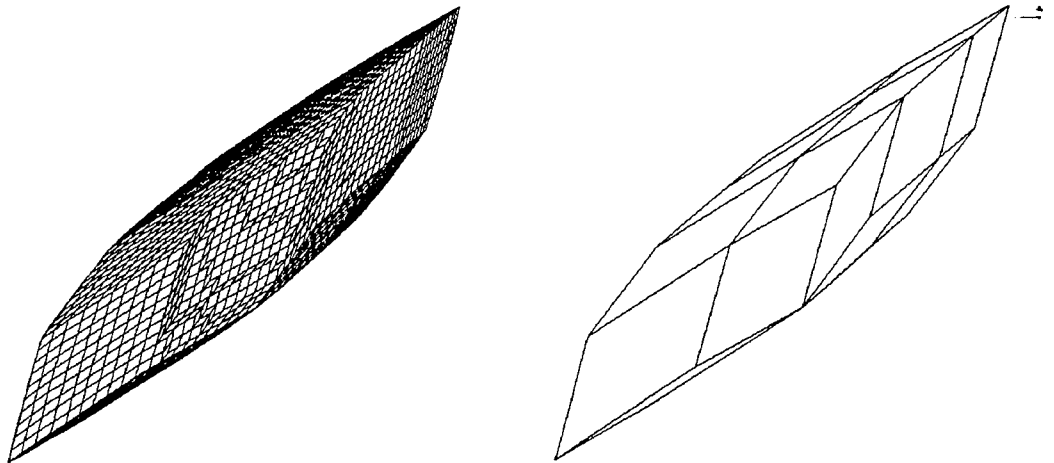


Figure 17 A gauche zonoèdre associé aux $\tilde{N}_i$ et à droite celui associé aux N_j .

- **Exemple 4** Nous avons 9 vecteurs de $\mathbb{R}^3$ que nous avons perturbé 10 fois chacun. Après résolution du problème (P) avec l'algorithme du recuit simulé, on a les résultats, représentés dans les figures 18 et 19.

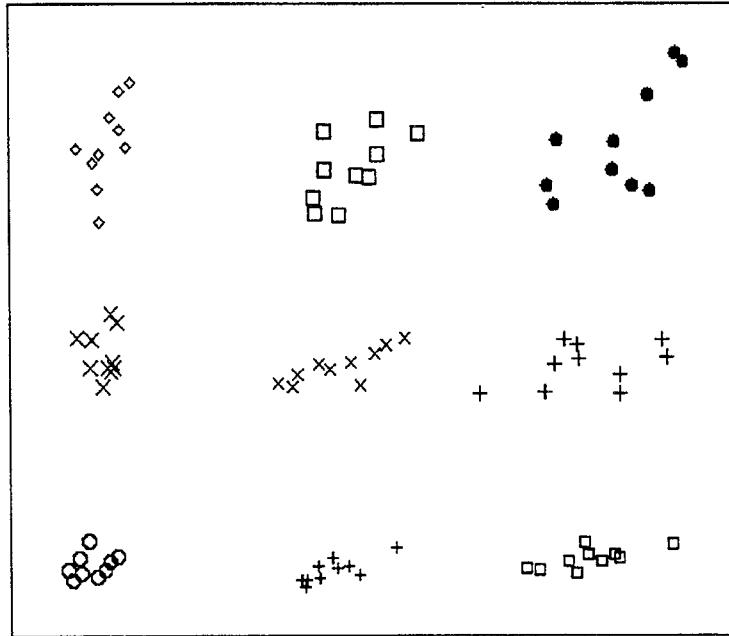


Figure 18 Vu dans le plan $z = 1$ des vecteurs, paramètre $T_0 = 0.1$, $Vol(Z) = 149077.60$.

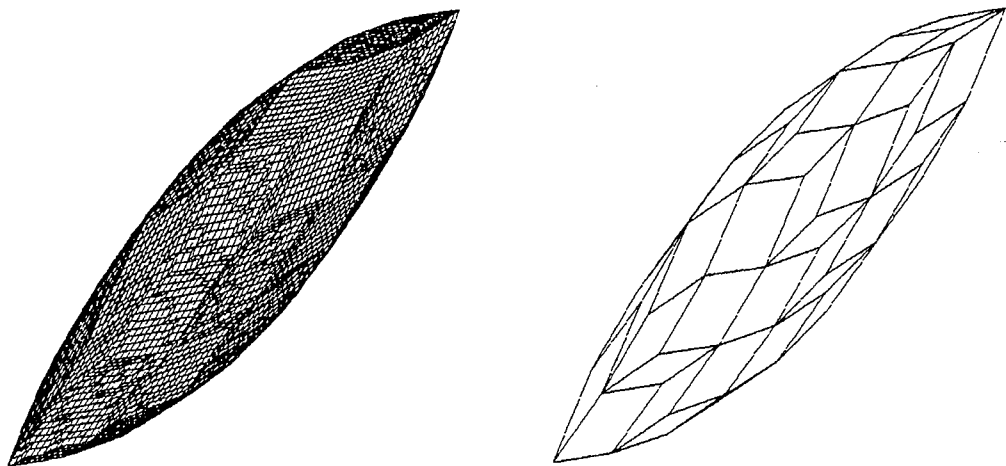


Figure 19 A gauche zonoèdre associé aux $\tilde{N}_i$ et à droite celui associé aux N_j .

• **Exemple 5** Dans cet exemple, nous avons tiré alatoirement 60 vecteurs dans $\mathbb{R}^3$. Nous avons demandé à ce qu'ils soient séparés en 7 groupes. Dans la figure 20, est donné une visualisation de la disposition des vecteurs dans le plan $z = 1$. Dans la figure 21, nous donnons le zonoèdre associé aux vecteurs $\tilde{N}_i$ et celui correspondant aux vecteurs N_i .

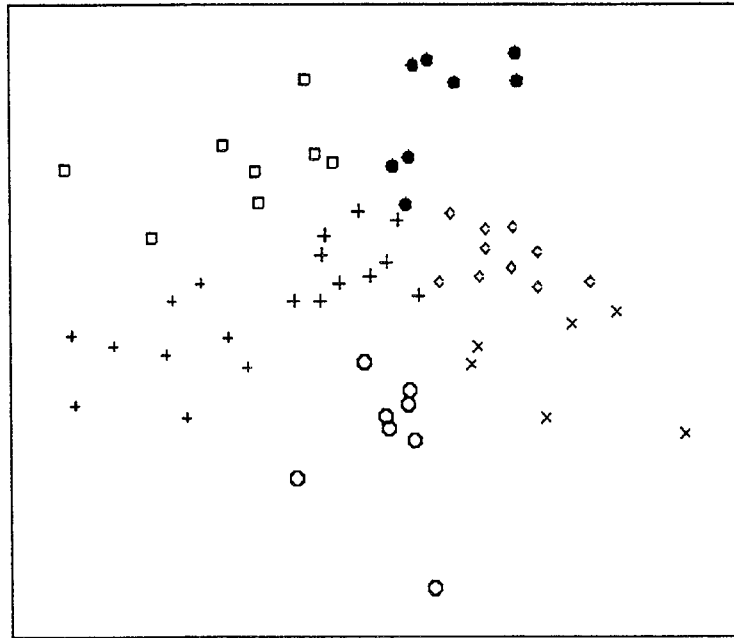


Figure 20 Distribution des vecteurs dans les groupes, paramètre $T_0 = 100$, $Vol(Z_p) = 52268.02$.

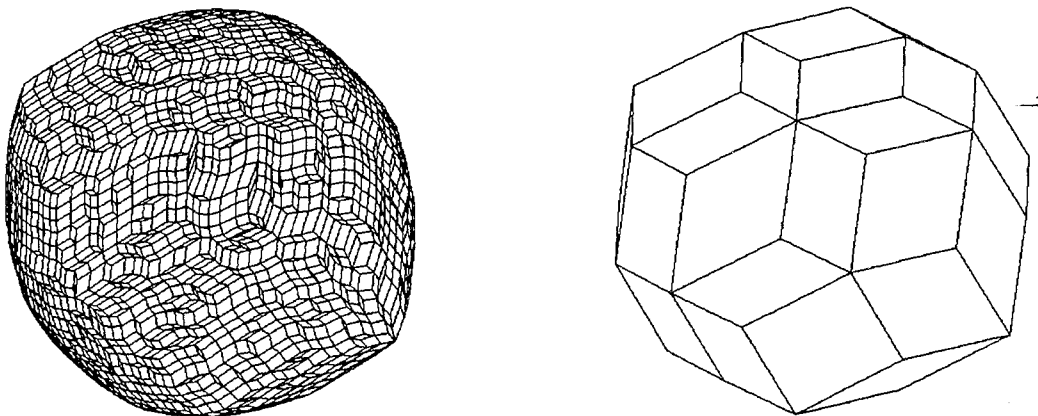


Figure 21 A gauche zonoèdre associé aux $\tilde{N}_i$ et à droite celui associé aux N_i .

Remarques sur la convergence, et le choix du paramètre température

Au cours des tests sur les exemples, nous avons remarqué que le choix de la température initiale est lié à la nature des données. En effet, lorsque les vecteurs traités sont en forme de groupes bien distincts, l'algorithme converge vers l'état optimal même pour des températures initiales très faibles qui obligent l'algorithme à une descente systématique, comme dans les exemples 3 et 4. Par contre, dans les exemples où les vecteurs sont distribués uniformément, comme dans l'exemple 5, le choix d'une température initiale faible mène, en général, à la convergence vers un état non optimal. Les figures suivantes donnent l'état final pour deux températures initiales différentes.

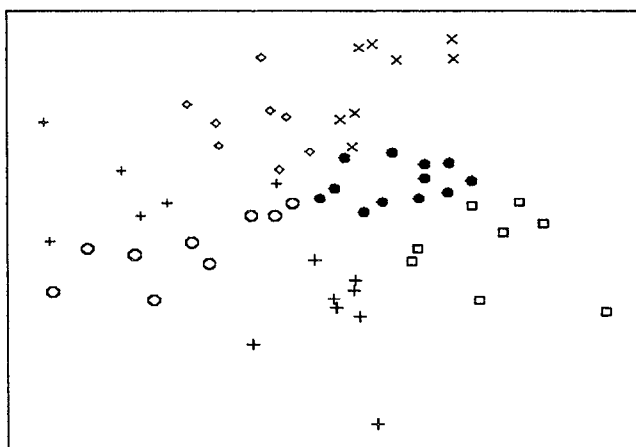


Figure 22 Convergence quand le paramètre $T_0 = 0.1$, $Vol(Z) = 50826.40$.

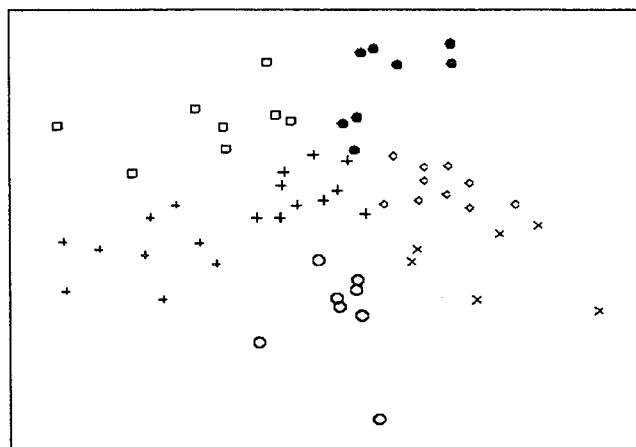


Figure 23 Convergence quand le paramètre $T_0 = 100$, $Vol(Z) = 52268.02$.

Chapitre 2

Sensibilité du problème d'optimisation pour la fabrication des mélanges

Introduction

Dans une raffinerie, on se sert de certains produits issus du raffinage comme bases à la fabrication d'autres produits. La gestion de cette fabrication pose quelques problèmes; citons par exemple:

- le problème de la faisabilité d'un mélange à partir de ces bases,
- le problème de la gestion du stock des produits de bases.

Le concept de fabrication de mélange développé par P. Valentin a introduit la représentation vectorielle d'un produit, laquelle a amené à considérer des entités géométriques appelées zonotopes. Le premier problème se ramène à déterminer l'appartenance d'un point à un zonotope, le second donne lieu au problème d'optimisation qui fait l'objet de cette étude. Plusieurs travaux ont été menés dans ce sens. Les premiers résultats ont été obtenus par K. Slaoui qui a étudié les problèmes de faisabilité d'un mélange et un critère d'optimalité dans le cadre des mélanges binaires (dans $\mathbb{R}^2$). D. Girard a généralisé ce critère aux mélanges non binaires, ce qui a conduit à un problème d'optimisation. N. Odeh a fait l'étude de quelques lois de mélanges ainsi que la visualisation par des coupes parallèles des zonotopes. N. Szafran a développé les méthodes de construction, visualisation et manipulation des zonoèdres.

Les propriétés des mélanges de base sont des résultats de mesures; elles sont donc entachées d'erreurs. Ces erreurs engendrent des variations de la solution du problème d'optimisation. Nous étudierons ces variations et nous caractériserons, sous certaines hypothèses, le domaine de variation de la solution du problème d'optimisation quand les erreurs varient dans un ellipsoïde.

Nous commencerons par rappeler la modélisation géométrique pour la fabrication des mélanges, puis nous parlerons du critère d'optimalité et du problème d'optimisation. Dans la troisième partie nous étudions la variation de la solution par rapport aux erreurs. Quelques exemples seront donnés à la fin.

1. Modèle géométrique de la fabrication de mélange

1.1. Modèle géométrique

On dispose de plusieurs produits P_i aux propriétés différentes. Chaque produit est représenté par n spécifications (ou propriétés) s_j . Ainsi, le produit P_i est représenté par $(s_1^i, s_2^i, \dots, s_n^i)$.

Définition 2.1 On appelle Ω_j , l'espace de représentation de la spécification s_j . Ω_j est en général $\mathbb{R}$ ou un intervalle de $\mathbb{R}$. $\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n$ est appelé espace des phases.

Définition 2.2 (Propriété additive) Une propriété F d'un mélange est dite additive, si étant donnés deux produits A et B et leurs masses m_A et m_B alors:

$$P_1 + P_2 \longrightarrow C,$$

$$m_C F_C = (m_{P_1} + m_{P_2}) F_C = m_{P_1} F_{P_1} + m_{P_2} F_{P_2}.$$

Exemple: un produit gasoil est caractérisé notamment par deux propriétés:

- son pourcentage en soufre s qui caractérise la propriété du gasoil,
- sa température de trouble T : température en dessous de laquelle le gasoil commence à cristalliser.

D'après la définition d'une loi additive, la première propriété est additive. Cela n'est pas le cas pour la deuxième spécification. Mais l'expérience a montré, qu'avec une certaine fonction biunivoque découverte empiriquement, on peut associer à la température de trouble T la valeur $fi(T)$ appelée indice de température de trouble. La grandeur $fi(T)$ est "presque" additive.

Si on a deux produits P_1 et P_2 tels que chacun a les spécifications:

$$P_1 = \begin{pmatrix} m_{P_1} \\ s_{P_1} \\ T_{P_1} \end{pmatrix} \quad \text{et} \quad P_2 = \begin{pmatrix} m_{P_2} \\ s_{P_2} \\ T_{P_2} \end{pmatrix}.$$

En les mélangeant, on obtient le produit $C = \begin{pmatrix} m_C \\ s_C \\ T_C \end{pmatrix}$ tel que:

$$m_C = m_{P_1} + m_{P_2},$$

$$m_C s_C = m_{P_1} s_{P_1} + m_{P_2} s_{P_2},$$

$$m_C fi(T_C) = m_{P_1} fi(T_{P_1}) + m_{P_2} fi(T_{P_2}).$$

Ainsi, mélanger P_1 et P_2 revient à faire la somme de vecteurs.

Un produit P ayant les spécifications (m, s, T) a pour représentation vectorielle le vecteur:

$$\begin{pmatrix} m \\ ms \\ mfi(T) \end{pmatrix}.$$

Dans la suite, on identifiera un produit et sa représentation vectorielle.

1. 2. Faisabilité

Soient N_i $i = 1, \dots, p$ des vecteurs représentant les produits de base dont on dispose. En mélangeant une fraction λ_i de chaque produit N_i , on obtient un mélange $M = \sum_{i=1}^p \lambda_i N_i$. Quand les λ_i , $i = 1, \dots, p$, décrivent $[0,1]$, M décrit le zonotope Z associé aux vecteurs N_i , $i = 1, \dots, p$. Le zonotope Z représente ainsi l'ensemble de tous les mélanges faisables à partir des produits N_i . Z est appelé aussi l'état de séparation de la famille $\{N_i\}$.

Définition 2. 3 Un produit M est faisable (ou réalisable) à partir des N_i , s'il existe $(\lambda_1, \lambda_2, \dots, \lambda_p)$, $0 \leq \lambda_i \leq 1$ tel que $\sum_{i=1}^p \lambda_i N_i = M$ (i.e. si $M \in Z$).

Un point M est partiellement faisable s'il est dans le cône formé par l'enveloppe positive des vecteurs N_i . Si M est partiellement faisable alors il existe $\alpha \in]0, 1[$ tel que $\alpha M \in Z$. (αM est un produit de mêmes caractéristiques physiques que M , mais de masse différente). Par opposition à "partiellement faisable" on dira que M est totalement faisable s'il est faisable au sens de la définition 2.3.

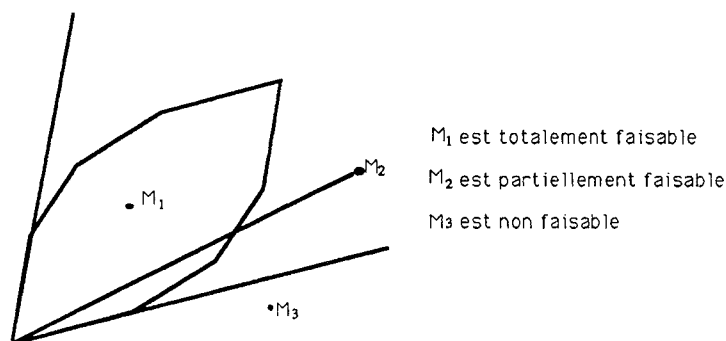


FIGURE .1. Faisabilité.

Notons A la matrice dont les colonnes sont formées par les vecteurs $N_i, i = 1, \dots, p$. On pourra déterminer la faisabilité partielle ou totale d'un mélange M , en résolvant le problème d'optimisation linéaire PL suivant:

$$(PL) \quad \begin{cases} \max & \alpha \\ [A, -M] \begin{pmatrix} \lambda \\ \alpha \end{pmatrix} &= 0, \\ 0 \leq \lambda_i \leq 1, \\ 0 \leq \alpha \leq 1. \end{cases}$$

Soit (λ^0, α^0) la solution de (PL) .

On a la relation suivante:

$$A\lambda^0 = \alpha^0 M.$$

Si M est faisable alors $\alpha^0 = 1$ et dans ce cas λ^0 est une solution faisable. Si $\alpha^0 = 0$ alors M n'est ni faisable ni partiellement faisable.

On appelle domaine de faisabilité d'un mélange M , l'ensemble:

$$U_M = \{ \lambda \in [0, 1]^p \mid \sum_{i=1}^p \lambda_i N_i = M \}.$$

C'est l'intersection de l'hypercube $[0, 1]^p$ de $\mathbb{R}^p$ avec le sous espace affine :

$$\{ \lambda \in \mathbb{R}^p \mid A\lambda = M \}.$$

C'est un fermé borné. Si M est réalisable alors U_M est non vide.

Remarque 1: Pour déterminer la proportion maximale faisable d'un produit M , il suffit de résoudre le problème (PL) sans prendre en compte la contrainte $\alpha \leq 1$.

2. Critère de gestion de la fabrication d'un mélange

Si U_M n'est pas réduit à un point, la solution n'est pas unique, d'où l'importance du choix d'un critère de gestion de la fabrication.

2.1. Zonotope résiduel

Le zonotope Z ,

$$Z = \{ \sum_{i=1}^p \lambda_i N_i \mid 0 \leq \lambda_i \leq 1 \},$$

est l'ensemble de tous les mélanges réalisables.

En consommant une quantité λ_i de chaque N_i pour fabriquer M , le résidu dans chaque bac est le produit représenté par:

$$(1 - \lambda_i)N_i.$$

Définition 2. 4 *L'ensemble des mélanges faisables après la fabrication de M , au moyen de $\lambda_i N_i$, $i = 1, \dots, p$, est donc le zonotope:*

$$Z_R(\lambda) = \left\{ \sum \mu_i (1 - \lambda_i) N_i \mid 0 \leq \mu_i \leq 1, \ i = 1, \dots, p \right\}.$$

$Z_R(\lambda)$ est appelé *zonotope résiduel après fabrication de M au moyen de λ* .

N.B : En toute logique, la notion indique que Z_R n'est pas un zonotope. $Z_R(\lambda)$ est un zonotope pour chaque λ . Z_R est une application de $[0, 1]^p$ dans l'espace des zonotopes.

L'ensemble des produits faisables après fabrication de M est entièrement défini par le convexe de mélanges $Z_R(\lambda)$. Remarquons que c'est un ensemble qui dépend du choix de λ ; pour illustrer ceci, prenons l'exemple suivant:

$$N_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad N_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad N_3 = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \text{ et } M = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}.$$

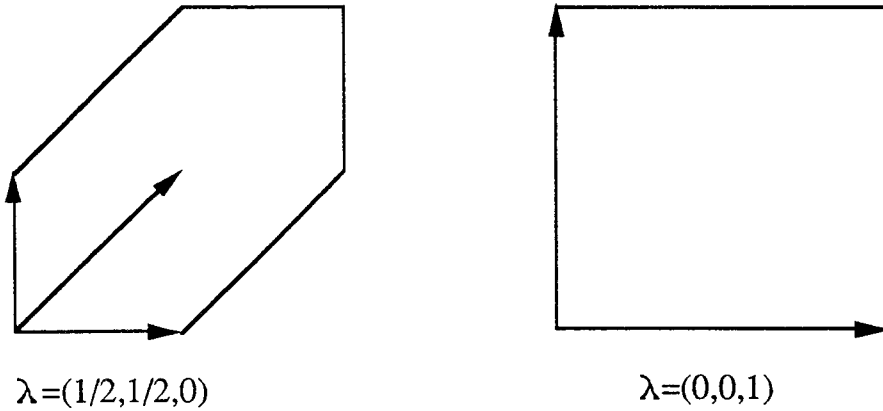


FIGURE .2. Exemple de zonotopes résiduels (dépendance par rapport à λ)

Remarque 2: Pour tout $\lambda \in U_M$, $\sum_{i=1}^p (1 - \lambda_i) N_i = \sum_{i=1}^p N_i - M$; tous les zonotopes résiduels ont donc au moins deux points en commun, qui sont O et $\sum_{i=1}^p N_i - M$, ce dernier étant le point distal de $Z_R(\lambda)$.

Notons $Z - M$ le translaté, par $-M$, du zonotope Z . Comme majoration, au sens de l'inclusion, du convexe résiduel, on a:

Proposition 2. 5 [Sla86] *Pour tout $\lambda \in U_M$, on a*

$$Z_R(\lambda) \subset Z \cap (Z - M).$$

L'ensemble $Z \cap (Z - M)$ admet la propriété suivante:

Proposition 2. 6 *Pour tout $M \in Z$, $Z \cap (Z - M)$ est un polytope ayant une symétrie centrale, le centre de symétrie étant $C - \frac{1}{2}M$; $C = \frac{1}{2} \sum_{i=1}^p N_i$ désigne le centre de symétrie de Z .*

Démonstration :

$x \in Z \cap (Z - M)$. Montrons que l'on a:

$$2[C - \frac{1}{2}M] - x = 2C - M - x \in Z \cap (Z - M).$$

Comme $x \in Z$, on a $2C - x \in Z$, donc $2C - M - x \in (Z - M)$.

Comme $x \in (Z - M)$, on a $x + M \in Z$, donc $2C - x - M \in Z$. ■

Corollaire 2. 7 *Dans $\mathbb{R}^2$, $Z \cap (Z - M)$ est un zonogone.*

Démonstration : D'après la proposition précédente et le fait que dans $\mathbb{R}^2$ tout polygone à symétrie centrale est un zonogone (cf chapitre 1). ■

2. 2. Fabrication à séparation résiduelle maximale

Le critère adopté pour choisir les fractions λ pour fabriquer un produit M , est celui de la gestion de la fabrication à séparation résiduelle maximale. C'est à dire, fabriquer M en augmentant l'état de séparation de la famille résiduelle et avoir un large éventail de produits faisables après la fabrication de M . Puisque l'état de séparation est donné par le zonotope associé aux vecteurs $(1 - \lambda_i)N_i$, alors il faut fabriquer M tout en augmentant la "taille" du zonotope résiduel Z_R .

• Cas $n = 2$

Dans le cas de la dimension 2, $Z \cap (Z - M)$ est un zonogone qui contient tous les zonogones résiduels. De plus, Slaoui a donné un algorithme constructif de troncature par la corde, pour trouver la décomposition λ vérifiant

$$Z_R(\lambda) = Z \cap (Z - M). \quad (2.1)$$

Donc $Z \cap (Z - M)$ est un convexe résiduel et c'est le plus grand d'après la proposition 2.5. L'idée de l'algorithme est la suivante: on suppose que les vecteurs N_i sont préalablement triés (par rapport à l'angle que fait chaque vecteur avec la droite Ox). Si M est faisable, on peut placer un segment $[M_1, M_2]$, équipotent à

OM , en position de corde sur Z (c'est à dire M_1 et M_2 sont sur la frontière de Z , voir figure 3). M_1 et M_2 s'écrivent:

$$M_1 = \sum_{k=1}^{i-1} N_k + \lambda_i N_i, \quad \lambda_i \in [0, 1],$$

$$M_2 = \sum_{k=1}^{j-1} N_k + \lambda_j N_j, \quad \lambda_j \in [0, 1].$$

Le vecteur λ optimal (i.e., vérifiant (2.1)) est de la forme

$$\lambda = (0, \dots, 0, \lambda_i, 1, \dots, 1, \lambda_j, 0, \dots, 0),$$

où i, j ($i < j$), $\lambda_i, \lambda_j \in [0, 1]$ vérifient:

$$M = (1 - \lambda_i)N_i + \sum_{k=i+1}^{j-1} N_k + \lambda_j N_j, \quad \lambda_i, \lambda_j \in [0, 1].$$

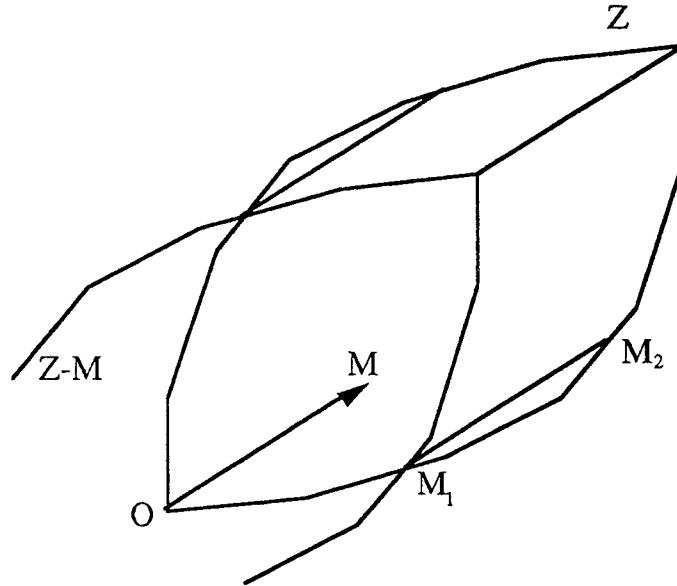


FIGURE 3. Troncature par la corde.

• Cas $n \geq 3$

Dans le cas où $n \geq 3$, $Z \cap (Z - M)$ n'est pas, en général, un zonotope comme le montre le contre exemple dû à P. Valentin: dans $\mathbb{R}^3$; si M vérifie:

$$M = N_1 + N_3 = N_2 + N_4,$$

alors $Z \cap (Z - M)$ est un octaèdre avec des faces triangulaires, donc ce n'est pas un zonoèdre d'après la proposition 1.13.

Si on prend le volume comme mesure de la taille de $Z_R(\lambda)$, on se ramène donc au problème:

$$\begin{cases} \max_{\lambda \in \mathbb{R}^p} \text{Vol}(Z_R(\lambda)) \\ A\lambda = M, \\ 0 \leq \lambda_i \leq 1, \end{cases} \quad i = 1, \dots, p. \quad (P)$$

Problème d'optimisation

D'après le corollaire 1.17,

$$\text{Vol}(Z_R(\lambda)) = \sum_{|I|=n} \prod_{i \in I} (1 - \lambda_i) |\Delta_I|, \quad (2.2)$$

où Δ_I est le déterminant des vecteurs dont l'indice est dans I ($I = \{i_j \mid j = 1, \dots, n \mid 1 \leq i_1 < \dots < i_n \leq p\}$).

Notons:

$$J(\lambda) = \text{Vol}(Z_R(\lambda)).$$

$J(\lambda)$ est un polynôme homogène de degré n par rapport à $\lambda \in \mathbb{R}^p$, donc continu.

On a le problème d'optimisation:

$$\begin{cases} \max_{\lambda \in \mathbb{R}^p} J(\lambda) \\ A\lambda = M \\ 0 \leq \lambda_i \leq 1 \end{cases} \quad i = 1, \dots, p. \quad (P)$$

L'existence de la solution est assurée par le fait que U_M est fermé borné et que J est continue.

Proposition 2.8 [Gir86] *$\text{Vol}^{1/n}$ est une fonction strictement concave sur U_M .*

Ce résultat est une conséquence du théorème de Brunn Minkowski suivant:

Théorème 2.9 *Si A et B sont deux ensembles convexes, alors on a:*

$$\text{Vol}^{1/n}(\alpha A + (1 - \alpha)B) \geq \alpha \text{Vol}^{1/n}(A) + (1 - \alpha) \text{Vol}^{1/n}(B).$$

L'inégalité est stricte sauf si A et B sont homothétiques ou reposent dans un même hyperplan.

Corollaire 2.10 *Le problème:*

$$\begin{cases} \max J^{1/n}(\lambda) \\ A\lambda = M, \\ 0 \leq \lambda_i \leq 1, \end{cases} \quad i = 1, \dots, p, \quad (P')$$

admet une solution λ^* unique qui est aussi l'unique solution du problème (P).

Démonstration :

Cela découle du fait que la fonctionnelle $J^{\frac{1}{n}}$ est strictement concave et que la fonction $(\cdot)^n$ est croissante sur $\mathbb{R}^+$. ■

Remarque 3: Si $J(\lambda) \neq 0$ alors on a la relation:

$$\nabla J^{1/n}(\lambda) = 1/n \cdot J^{\frac{1}{n}-1}(\lambda) \cdot \nabla J(\lambda). \quad (2.3)$$

Donc les gradients de $J^{1/n}$ et J sont proportionnels et ainsi $J^{1/n}$ et J ont les mêmes points stationnaires.

Ceci nous permet d'utiliser un algorithme itératif de descente pour résoudre le problème (P).

Le Lagrangien associé au problème (P) s'écrit:

$$\begin{aligned} \mathcal{L}(\lambda, \alpha, \beta, \gamma) = & J(\lambda) + \langle \alpha, A\lambda - M \rangle \\ & + \langle \beta, \lambda \rangle - \langle \gamma, \lambda - 1 \rangle, \end{aligned}$$

où $\alpha \in \mathbb{R}^n$ est le multiplicateur de Lagrange associé à la contrainte $A\lambda = M$ et

β (resp. γ) $\in \mathbb{R}_+^p$ est le multiplicateur de Kuhn et Tucker associé à la contrainte $0 \leq \lambda$ (resp. $\lambda \leq 1$).

Les multiplicateurs de Kuhn et Tucker β et γ doivent être à composantes toutes positives.

Si λ^* est la solution de (P) alors il existe $\alpha^* \in \mathbb{R}^n$, $\beta^* \in \mathbb{R}_+^p$ et $\gamma^* \in \mathbb{R}_+^p$, tels qu'on ait les relations:

$$\begin{aligned} \nabla J(\lambda^*) + A^t \alpha^* + \beta^* - \gamma^* &= 0, \\ A\lambda^* &= M, \\ \beta_i^* \lambda_i^* &= 0, \quad i = 1, \dots, p, \\ \gamma_i^* (\lambda_i^* - 1) &= 0, \quad i = 1, \dots, p. \end{aligned}$$

En un point $\lambda \in U_M$, on définit les ensembles des indices de contraintes saturées suivants:

$$\mathcal{I}_0 = \{i \mid \lambda_i = 0\},$$

$$\mathcal{I}_1 = \{i \mid \lambda_i = 1\}.$$

A l'optimum, on a:

$$\beta_i^* = 0 \text{ si } i \notin \mathcal{I}_0 \text{ et } \gamma_i^* = 0 \text{ si } i \notin \mathcal{I}_1.$$

2. 3. Algorithme de résolution

Le problème qui se pose est un problème d'optimisation non linéaire avec contraintes linéaires du type égalité et inégalité de la forme

$$\begin{cases} \max J(\lambda) \\ A\lambda = M, \\ C\lambda \geq b. \end{cases} \quad (P)$$

Pour le résoudre, on a choisi la méthode du gradient projeté proposé dans [Ros60]. Cette méthode consiste à projeter le gradient sur la frontière du domaine défini par les contraintes, ce qui permet un cheminement le long de la frontière dans la direction de la plus forte pente "relative" (autorisée par les contraintes). En un point y , soit

$$I(y) = \{ i \mid \langle C_i, y \rangle = b_i \},$$

$I(y)$ est l'ensemble des indices des contraintes actives au point y .

Notons C_F la sous matrice de C constituée par les lignes dont les indices sont dans $I(y)$. Soit S le sous espace affine:

$$S = \left\{ x \mid \begin{pmatrix} A \\ C_F \end{pmatrix} x = \begin{pmatrix} M \\ b_F \end{pmatrix} \right\}.$$

L'algorithme se déroule selon le schéma suivant:

Soit y_0 une solution faisable. y_0 peut être obtenue en résolvant le problème (PL). (Voir faisabilité).

A l'étape k on est en y_k .

On cherche une direction de déplacement admissible de descente δ^k ($\|\delta^k\| = 1$), qui permet d'augmenter le plus possible la fonctionnelle J et qui, pour un petit déplacement, permet de rester dans l'intersection de S avec le domaine défini par les contraintes. Ceci suppose que δ_k vérifie:

$$\begin{cases} A\delta^k = 0, \\ C_F\delta^k = 0. \end{cases} \quad (2.4)$$

On suppose que le rang de la matrice $\begin{pmatrix} A \\ C_F \end{pmatrix}$ est plein: $\text{rang} \begin{pmatrix} A \\ C_F \end{pmatrix} = \text{rang}(A) + |I(y_k)| = n + |I(y_k)|$; c'est à dire, on suppose qu'il n'y a pas dégénérescence au point y_k .

Soit S^0 le sous espace vectoriel parallèle S c'est à dire vérifiant (2.4).

On montre (voir Minoux 83) que la direction de déplacement $\delta^k \in S^0$ telle que $\langle \nabla J(y_k), \delta^k \rangle$ soit maximal avec la contrainte de normalisation $\|\delta^k\| = 1$, est donnée par $\delta^k = d_k / \|d_k\|$ où d_k est la projection orthogonale de $\nabla J(y_k)$ sur S^0 . Notons P_{S^0} l'opérateur de projection sur S^0 . On a alors :

$$\begin{aligned} d_k &= P_{S^0}(\nabla J(y^k)), \\ P_{S^0} &= I - \tilde{A}^t(\tilde{A}\tilde{A}^t)^{-1}\tilde{A}, \end{aligned} \quad (2.5)$$

où $\tilde{A}$ est la matrice $\begin{pmatrix} A \\ C_F \end{pmatrix}$.

Le pas de déplacement ρ_k dans la direction δ^k est déterminé de manière à maximiser la fonction $\rho \longrightarrow J(y_k + \rho\delta^k)$ par rapport à $\rho > 0$ tel que $y_k + \rho\delta^k$ appartient au domaine défini par les contraintes.

Soit:

$$\rho_{max} = \max\{\rho > 0 \mid y_k + \rho\delta^k \in U_M\},$$

ρ_k est la solution du problème unidimensionnel:

$$\max_{0 \leq \rho \leq \rho_{max}} J(y_k + \rho\delta^k).$$

Le nouveau point est:

$$y_{k+1} = y_k + \rho_k\delta^k.$$

L'algorithme se poursuit tant que d_k , la projection sur S^0 du gradient, n'est pas nulle.

Lorsque $d_k = 0$ au point courant y_k , alors $\nabla J(y_k)$ est orthogonale à S . Dans ce cas y_k est l'optimum sur le sous espace S , mais cela ne signifie pas que c'est l'optimum sur tout le domaine des contraintes. Pour vérifier ceci, on calcule le vecteur μ_F des multiplicateurs de Kuhn et Tucker correspondants aux contraintes actives. Les vecteurs α et μ_F sont solutions de l'équation:

$$[A^t, C_F^t] \begin{pmatrix} \alpha \\ \mu_F \end{pmatrix} = -\nabla J(y_k).$$

Si le vecteur μ_F est à composantes non négatives (condition de Kuhn et Tucker), alors le point courant est l'optimum $y_k = y^*$. Sinon, on réduit l'ensemble des contraintes actives. Le fait qu'il existe j tel que $\mu_j < 0$ exprime le fait qu'on a obligé la solution à être sur la face définie par la $j^{\text{ème}}$ contrainte sans que cela ne soit nécessaire. On supprime cette contrainte de C_F et on continue le cheminement.

Mise en oeuvre pratique

L'opérateur de projection sur l'espace S^0 est donné par l'expression (2.5). Cependant, la matrice A peut comporter des lignes très voisines puisqu'il s'agit de la quantité d'une espèce physique dans les bases. Le calcul de la projection peut être instable. Pour cela, il est préférable d'opérer par des transformations orthogonales. On a utilisé la factorisation QR avec orthogonalisation de Householder de la matrice $\tilde{A}^t$.

$$\tilde{A}^t = [Q_1, Q_2] \begin{bmatrix} R \\ 0 \end{bmatrix},$$

où $Q = [Q_1, Q_2]$ est une matrice orthogonale et R est une matrice $n \times n$ triangulaire supérieure. La matrice Q_1 représente une base de $Im(\tilde{A}^t)$ et Q_2 est une base du sous-espace $N(\tilde{A})$ (noyau de $\tilde{A}$). Dans ce cas, la projection sur S^0 s'écrit sous la forme simple:

$$P_{S^0} = Q_2 Q_2^t.$$

La résolution de l'équation vérifiée par les multiplicateurs se traduit par la résolution du problème:

$$R \mu = - \begin{pmatrix} Q_1^t \\ Q_2^t \end{pmatrix} \nabla J(y_k),$$

avec R triangulaire supérieure.

3. Influence des erreurs d'analyses sur la solution:

Chaque composante N_i^j , $j = 1, \dots, n$, de N_i , $i = 1, \dots, p$, est le résultat d'une mesure, donc entachée d'une erreur ε_{ij} .

Soit ε la matrice $(n \times p)$ des variables ε_{ij} , $1 \leq i \leq n$ et $1 \leq j \leq p$.

Remarque 4: Dans la suite, ε désignera soit la matrice de $\mathbb{R}^n \times \mathbb{R}^p$, soit le vecteur de $\mathbb{R}^{np}$.

Au cours du processus de la fabrication, la matrice de donnée A est remplacée par la matrice $A + \varepsilon$. Donc le problème d'optimisation (P) devient un problème perturbé de la forme:

$$\begin{cases} \max J(\lambda, \varepsilon), \\ (A + \varepsilon)\lambda = M, \\ 0 \leq \lambda_i \leq 1, \end{cases} \quad (P_\varepsilon)$$

où:

$$J(\lambda, \varepsilon) = V(\lambda, \varepsilon) = \sum_{|I|=n} \prod_{i \in I} (1 - \lambda_i) |\Delta_I^\varepsilon|. \quad (2.6)$$

Le but est d'étudier la variation de la solution λ_ε du problème (P_ε) . (λ_0 étant la solution du problème initial correspondant à $\varepsilon = 0$).

3. 1. Régularité de la solution par rapport aux erreurs

On se placera dans le cas d'un problème d'optimisation avec contraintes linéaires de type égalité puis on étendra l'étude au cas de contraintes linéaires de type inégalité.

3.1.1. Sensibilité dans le cas de contraintes linéaires de type égalité

Supposons que l'on ait le problème d'optimisation sous contraintes de type égalité suivant:

$$\begin{cases} \max J(\lambda, \varepsilon), \\ (A + \varepsilon)\lambda = M. \end{cases} \quad (PE_\varepsilon)$$

Le Lagrangien associé dans ce cas est de la forme:

$$\mathcal{L}(\lambda, \alpha, \varepsilon) = J(\lambda, \varepsilon) + \langle \alpha, (A + \varepsilon)\lambda - M \rangle.$$

Soit λ_ε la solution du problème (PE_ε) ; il existe $\alpha_\varepsilon \in \mathbb{R}^n$, vecteur des multiplicateurs de Lagrange, tel que l'on ait les relations:

$$\begin{aligned}\nabla_\lambda \mathcal{L}(\lambda_\varepsilon, \alpha_\varepsilon, \varepsilon) &= 0, \\ \nabla_\alpha \mathcal{L}(\lambda_\varepsilon, \alpha_\varepsilon, \varepsilon) &= 0,\end{aligned}\tag{2.7}$$

On a le résultat suivant:

Théorème 2.11 *Soit λ_ε la solution du problème (PE_ε) . Si la fonction J est de classe C^∞ par rapport à (λ, ε) quand ε varie dans un voisinage U_0 de 0, alors il existe un voisinage de 0 tel que λ_ε est une fonction de classe C^∞ par rapport à ε .*

Remarque 5: Pour que J soit C^∞ par rapport à ε dans U_0 , il faut et il suffit que les vecteurs N_i et ε soient tels que les erreurs n'entraînent pas un changement de signe des déterminants Δ_f^ε sur U_0 .

démonstration du théorème:

Pour démontrer le théorème, on essayera de se placer dans les conditions du théorème des fonctions implicites.

Notons $H = \nabla_\lambda^2 J(\lambda_0, 0)$, le Hessien de la fonction J au point $(\lambda_0, 0)$. (On utilisera la notation H ou $\nabla_\lambda^2 J((\lambda_0, 0))$ selon le besoin).

Puisque la fonction J est de classe C^∞ , la fonction $\mathcal{L}$ est de classe C^∞ par rapport aux variables $(\lambda, \alpha, \varepsilon)$. Donc la fonction $\begin{pmatrix} \nabla_\lambda \mathcal{L} \\ \nabla_\alpha \mathcal{L} \end{pmatrix}$ est de classe C^∞ .

Les relations à l'optimum pour $\varepsilon = 0$ s'écrivent:

$$\begin{aligned}\nabla_\lambda \mathcal{L}(\lambda_0, \alpha_0, 0) &= 0, \\ \nabla_\alpha \mathcal{L}(\lambda_0, \alpha_0, 0) &= 0.\end{aligned}$$

La Jacobienne partielle par rapport à (λ, α) à l'optimum s'écrit:

$$\begin{pmatrix} \nabla_{\lambda^2}^2 \mathcal{L}(\lambda_0, \alpha_0, 0) & \nabla_{\lambda\alpha}^2 \mathcal{L}(\lambda_0, \alpha_0, 0) \\ \nabla_{\alpha\lambda}^2 \mathcal{L}(\lambda_0, \alpha_0, 0) & \nabla_{\alpha^2}^2 \mathcal{L}(\lambda_0, \alpha_0, 0) \end{pmatrix}.$$

Or

$$\begin{aligned}\nabla_{\lambda^2}^2 \mathcal{L}(\lambda_0, \alpha_0, 0) &= \nabla_\lambda^2 J(\lambda_0, 0) = H, \\ \nabla_{\lambda\alpha}^2 \mathcal{L}(\lambda_0, \alpha_0, 0) &= A^t, \\ \nabla_{\alpha\lambda}^2 \mathcal{L}(\lambda_0, \alpha_0, 0) &= A, \\ \nabla_{\alpha^2}^2 \mathcal{L}(\lambda_0, \alpha_0, 0) &= 0.\end{aligned}\tag{2.8}$$

Notons W la matrice Jacobienne partielle au point (λ_0, α_0) . W est de la forme:

$$W = \begin{pmatrix} H & A^t \\ A & 0 \end{pmatrix}.$$

Démontrons que W est inversible.

Si H est inversible alors il en est de même pour la matrice W . En effet, on démontre que:

$$\begin{pmatrix} H & A^t \\ A & 0 \end{pmatrix}^{-1} = \begin{pmatrix} H^{-1} - H^{-1}A^t(AH^{-1}A^t)^{-1}AH^{-1} & H^{-1}A^t(AH^{-1}A^t)^{-1} \\ (AH^{-1}A^t)^{-1}AH^{-1} & -AH^{-1}A^t \end{pmatrix}.$$

Ce résultat s'obtient en prenant une matrice $\begin{pmatrix} C & D^t \\ D & E \end{pmatrix}$ et en résolvant le système:

$$\begin{aligned} CH + D^t A &= I_p, \\ CA^t &= 0, \\ DH + EA &= 0, \\ DA^t &= I_n. \end{aligned}$$

La condition, H est inversible, est une condition suffisante pour que W le soit mais elle n'est pas vérifiée en général. Cependant, la restriction de H à $N(A)$ est inversible. En effet, si $J(\lambda_0, 0) \neq 0$ alors:

$$\nabla_\lambda^2 J^{1/n}(\lambda_0, 0) = \frac{1}{n} J^{\frac{1}{n}-1}(\lambda_0, 0) \nabla_\lambda^2 J(\lambda_0, 0) + \frac{1}{n} \left(\frac{1}{n} - 1 \right) J^{\frac{1}{n}-2}(\lambda_0, 0) \nabla_\lambda J(\lambda_0, 0) \nabla_\lambda J^t(\lambda_0, 0),$$

et $\nabla_\lambda^2 J^{1/n}(\lambda_0, 0)$ est définie négative car $J^{1/n}$ est strictement concave.

Soit $z \in N(A)$; alors:

$$\begin{aligned} z^t \nabla_\lambda^2 J^{1/n}(\lambda_0, 0) z &= \frac{1}{n} J^{\frac{1}{n}-1}(\lambda_0, 0) z^t \nabla_\lambda^2 J(\lambda_0, 0) z + \\ &\quad \frac{1}{n} \left(\frac{1}{n} - 1 \right) J^{\frac{1}{n}-2}(\lambda_0, 0) z^t \nabla_\lambda J(\lambda_0, 0) \nabla_\lambda J^t(\lambda_0, 0) z. \end{aligned}$$

A l'optimum, $\nabla_\lambda J(\lambda_0, 0)$ est orthogonal à $N(A)$, donc

$$\nabla_\lambda J^t(\lambda_0, 0) z = 0 \quad \forall z \in N(A),$$

d'où

$$z^t \nabla_\lambda^2 J^{1/n}(\lambda_0, 0) z = z^t \nabla_\lambda^2 J(\lambda_0, 0) z, \quad (2.9)$$

ce qui exprime que la restriction de $\nabla_\lambda^2 J(\lambda_0, 0)$ à $N(A)$ est aussi définie négative et donc inversible.

Considérons la matrice P dont les lignes forment une base de $N(A)$. Pour tout z de $N(A)$, il existe $s \in \mathbb{R}^{p-n}$ tel que $z = P^t s$. La restriction de H à $N(A)$ s'exprime comme suit:

$$PHP^t,$$

et d'après la relation (2.9), la matrice PHP^t est inversible.

On montre maintenant que si la restriction de H à $N(A)$ est inversible alors la matrice W l'est aussi.

Le rang de A est plein (les vecteurs A_i^t constituant les colonnes de A^t sont linéairement indépendants; ils génèrent donc un espace de dimension n dans $\mathbb{R}^p$). Soit V une matrice telle que V^t soit constituée des vecteurs tels que (A^t, V^t) soit une base de $\mathbb{R}^p$.

La matrice $[A^t, V^t]$ est non singulière. Il existe une matrice $\begin{pmatrix} Y \\ Z \end{pmatrix}$ telle que l'on ait les relations:

$$A^t Y + V^t Z = I,$$

et

$$\begin{pmatrix} Y A^t & Y V^t \\ Z A^t & Z V^t \end{pmatrix} = \begin{pmatrix} I_n & 0 \\ 0 & I_{p-n} \end{pmatrix},$$

c'est à dire qu'on ait :

$$Y A^t = I_n,$$

$$Z A^t = 0,$$

$$Y V^t = 0,$$

$$Z V^t = I_{p-n}.$$

On a $A Z^t = 0$, donc $Z_i^t \in N(A)$, $i = 1, \dots, p-n$. Les colonnes de Z^t constituent une base de $N(A)$. Donc $Z H Z^t$ est inversible (restriction de H à $N(A)$).

Notons $\begin{pmatrix} C & D^t \\ D & E \end{pmatrix}$ la matrice inverse de $\begin{pmatrix} H & A^t \\ A & 0 \end{pmatrix}$. Alors:

$$\begin{cases} C = Z^t (Z H Z^t)^{-1} Z, \\ D = Y - Y H Z^t (Z H Z^t)^{-1} Z, \\ E = Y H Y^t - Y H Z^t (Z H Z^t)^{-1} Z H Y^t. \end{cases} \quad (2.10)$$

En effet:

$$(i) \quad C H + D^t A = I_p,$$

$$(ii) \quad C A^t = 0,$$

$$(iii) \quad D H + E A = 0,$$

$$(iv) \quad D A^t = I_n.$$

Dans la première équation, en multipliant à droite, par la matrice (Y^t, Z^t) , on obtient:

$$\begin{cases} C H Y^t + D^t = Y^t, \\ C H Z^t = Z^t, \end{cases}$$

qui peut s'écrire:

$$\begin{cases} C (A^t \ V^t) \begin{pmatrix} Y \\ Z \end{pmatrix} H Y^t + D^t = Y^t, \\ C (A^t \ V^t) \begin{pmatrix} Y \\ Z \end{pmatrix} H Z^t = Z^t. \end{cases}$$

En utilisant la relation (ii), on obtient:

$$\begin{cases} CV^t ZHY^t + D^t = Y^t, \\ CV^t ZHZ^t = Z^t, \end{cases}$$

ou encore, puisque ZHZ^t est inversible:

$$\begin{cases} CV^t ZHY^t + D^t = Y^t, \\ CV^t = Z^t(ZHZ^t)^{-1}, \end{cases}$$

ce qui permet de calculer D^t :

$$\begin{cases} D^t = Y^t - Z^t(ZHZ^t)^{-1}ZHY^t, \\ CV^t = Z^t(ZHZ^t)^{-1}. \end{cases}$$

Des équations

$$\begin{cases} CA^t = 0, \\ CV^t = Z^t(ZHZ^t)^{-1}, \end{cases}$$

on peut tirer

$$C(A^t \quad V^t) \begin{pmatrix} Y \\ Z \end{pmatrix} = (0 \quad Z^t(ZHZ^t)^{-1}) \begin{pmatrix} Y \\ Z \end{pmatrix},$$

et donc

$$C = Z^t(ZHZ^t)^{-1}Z.$$

En outre, considérons les équations (iii) et (iv)

$$\begin{cases} DH + EA = 0, \\ DA^t = I, \end{cases}$$

et multiplions par (Y^t, Z^t) :

$$\begin{cases} E = -DHY^t, \\ 0 = -DHZ^t. \end{cases}$$

On obtient donc, en remplaçant D par sa valeur:

$$E = YHY^t - YHZ^t(ZHZ^t)^{-1}ZHY^t.$$

La Jacobienne est donc inversible. L'inverse est donné par (2.10).

Par application du théorème des fonctions implicites, on démontre que:

- i- il existe un voisinage V_0 de 0 dans $\mathbb{R}^{np}$,
 - ii- il existe une fonction de classe C^∞ , de $V_0 \longrightarrow \mathbb{R}^{n+p}$, qui à ε associe $(\lambda_\varepsilon, \alpha_\varepsilon)$
- tels que:

$$\begin{cases} \nabla_{\lambda} \mathcal{L}(\lambda_{\varepsilon}, \alpha_{\varepsilon}, \varepsilon) = 0, \\ \nabla_{\alpha} \mathcal{L}(\lambda_{\varepsilon}, \alpha_{\varepsilon}, \varepsilon) = 0, \end{cases}$$

pour tout ε dans V_0 . ■

Nous notons cette fonction $\begin{pmatrix} \lambda(\varepsilon) \\ \alpha(\varepsilon) \end{pmatrix}$.

• **Jacobienne de la fonction** $\begin{pmatrix} \lambda \\ \alpha \end{pmatrix}$:

On a démontré que pour J régulière, la solution du problème (PE_{ε}) est aussi régulière par rapport aux perturbations sur les données. Pour déterminer la sensibilité, on calcule la Jacobienne par rapport à ε de la solution.

Pour ε tel que $\|\varepsilon\|$ suffisamment petit, on a:

$$\begin{cases} \nabla_{\lambda} \mathcal{L}(\lambda(\varepsilon), \alpha(\varepsilon), \varepsilon) = 0, \\ \nabla_{\alpha} \mathcal{L}(\lambda(\varepsilon), \alpha(\varepsilon), \varepsilon) = 0. \end{cases}$$

En dérivant par rapport à ε , on obtient la relation:

$$\begin{cases} \nabla_{\lambda^2}^2 \mathcal{L}(\lambda(\varepsilon), \alpha(\varepsilon), \varepsilon) \cdot \nabla_{\varepsilon} \lambda(\varepsilon) + \nabla_{\lambda \alpha}^2 \mathcal{L}(\lambda(\varepsilon), \alpha(\varepsilon), \varepsilon) \cdot \nabla_{\varepsilon} \alpha(\varepsilon) + \nabla_{\lambda \varepsilon}^2 \mathcal{L}(\lambda(\varepsilon), \alpha(\varepsilon), \varepsilon) = 0, \\ \nabla_{\alpha \lambda}^2 \mathcal{L}(\lambda(\varepsilon), \alpha(\varepsilon), \varepsilon) \cdot \nabla_{\varepsilon} \lambda(\varepsilon) + \nabla_{\alpha^2}^2 \mathcal{L}(\lambda(\varepsilon), \alpha(\varepsilon), \varepsilon) \cdot \nabla_{\varepsilon} \alpha(\varepsilon) + \nabla_{\alpha \varepsilon}^2 \mathcal{L}(\lambda(\varepsilon), \alpha(\varepsilon), \varepsilon) = 0. \end{cases}$$

En utilisant les relations données par (2.8), on obtient:

$$\begin{pmatrix} H & A^t \\ A & 0 \end{pmatrix} \begin{pmatrix} \nabla_{\varepsilon} \lambda(0) \\ \nabla_{\varepsilon} \alpha(0) \end{pmatrix} = - \begin{pmatrix} \nabla_{\lambda \varepsilon}^2 \mathcal{L}(\lambda_0, \alpha_0, 0) \\ \Lambda_0 \end{pmatrix},$$

où Λ_0 est la matrice $(n \times np)$, qui à ε associe le vecteur $\varepsilon \lambda_0$.

On en tire que:

$$\begin{pmatrix} \nabla_{\varepsilon} \lambda(0) \\ \nabla_{\varepsilon} \alpha(0) \end{pmatrix} = - \begin{pmatrix} H & A^t \\ A & 0 \end{pmatrix}^{-1} \begin{pmatrix} \nabla_{\lambda \varepsilon}^2 \mathcal{L}(\lambda_0, \alpha_0, 0) \\ \Lambda_0 \end{pmatrix}, \quad (2.11)$$

et

$$\nabla_{\varepsilon} \lambda = -PW^{-1} \begin{pmatrix} \nabla_{\lambda \varepsilon}^2 \mathcal{L}(\lambda_0, \alpha_0, 0) \\ \Lambda_0 \end{pmatrix},$$

où P est la matrice de projection telle que $P(\lambda, \alpha) = \lambda$.

Le calcul des valeurs singulières du Jacobien nous permet de déterminer la sensibilité de la solution λ_{ε} par rapport à ε ; si la plus grande des valeurs singulières en module est très "grande", alors le modèle est très sensible par rapport aux erreurs.

3.1.2. Cas des contraintes de type égalité et inégalité

On a le problème d'optimisation avec contraintes de type égalité et d'inégalité suivant:

$$\begin{cases} \max & J(\lambda, \varepsilon) \\ (A + \varepsilon)\lambda = M, \\ C\lambda \geq b. \end{cases} \quad (PI_\varepsilon)$$

On veut étudier la sensibilité de λ par rapport à ε .
Le Lagrangien associé dans ce cas est de la forme:

$$\mathcal{L}(\lambda, \alpha, \mu, \varepsilon) = J(\lambda, \varepsilon) + \langle \alpha, (A + \varepsilon)\lambda - M \rangle + \langle \mu, C\lambda - b \rangle,$$

où α est un vecteur de $\mathbb{R}^n$ et μ un vecteur à composantes positives.

Si λ_ε est la solution de (PI_ε) , alors il existe α_ε et μ_ε tels que $\mu_{\varepsilon_i} \geq 0$, $\forall i$, et vérifiant les relations:

$$\begin{cases} \nabla_\lambda J(\lambda_\varepsilon, \varepsilon) + (A + \varepsilon)^t \alpha_\varepsilon + C^t \mu_\varepsilon = 0, \\ (A + \varepsilon)\lambda_\varepsilon - M = 0, \\ \mu_{\varepsilon_i}(C\lambda_\varepsilon - b)_i = 0, \quad \forall i. \end{cases} \quad (2.12)$$

Notons C_F (resp. C_L) la sous matrice constituée des lignes de C pour lesquelles on a l'égalité (resp. l'inégalité stricte) c'est à dire les indices qui saturent les contraintes. Posons:

$$C\lambda = \begin{pmatrix} C_F \\ C_L \end{pmatrix} \lambda = \begin{pmatrix} b_F \\ b_L \end{pmatrix},$$

et

$$\mu = \begin{pmatrix} \mu_F \\ \mu_L \end{pmatrix}.$$

Définition 2. 12

- Une contrainte est dite *fortement active* si elle est active et le multiplicateur de Kuhn et Tucker associé est strictement positif.
- Une contrainte est dite *faiblement active* si elle est active et le multiplicateur de Kuhn et Tucker associé est nul.

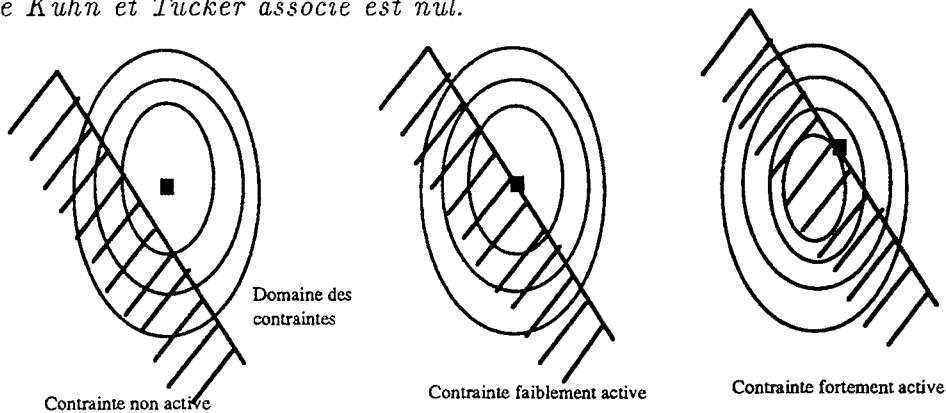


FIGURE .4. Contraintes.

On a le résultat suivant:

Théorème 2.13 Soit (λ, ε) solution du problème (PI_ε) .

Si la fonction J est de classe C^∞ par rapport à (λ, ε) quand ε varie dans un voisinage U_0 de 0, et si à l'optimum toutes les contraintes sont fortement actives alors il existe un voisinage de 0 tel que λ est une fonction de classe C^∞ par rapport à ε .

Démonstration :

Comme pour le cas des contraintes d'égalités, on se placera dans les conditions d'application du théorème des fonctions implicites.

Considérons la fonction:

$$(\lambda, \alpha, \mu_F, \mu_L, \varepsilon) \longrightarrow \begin{pmatrix} \nabla_\lambda J(\lambda, \varepsilon) + (A + \varepsilon)^t \alpha + C_F^t \mu_F + C_L^t \mu_L \\ (A + \varepsilon)\lambda - M \\ \mu_{F_i}(C_F \lambda - b)_i \\ \mu_{L_i}(C_L \lambda - b)_i \end{pmatrix}. \quad (2.13)$$

C'est une fonction de classe C^∞ de $(\lambda, \alpha, \mu, \varepsilon)$.

Pour $\varepsilon = 0$, la relation (2.12) s'écrit:

$$\begin{cases} \nabla_\lambda J(\lambda_0, 0) + A^t \alpha_0 + C^t \mu_0 = 0, \\ A\lambda - M = 0, \\ \mu_{0_i}(C_F \lambda_0 - b_F)_i = 0, & i \in I(\lambda_0), \\ \mu_{0_i}(C_L \lambda_0 - b_L)_i = 0, & i \notin I(\lambda_0). \end{cases}$$

Donc la fonction définie en (2.13) s'annule au point $(\lambda_0, \alpha_0, \mu_0, 0)$. La Jacobienne partielle $(\nabla_\lambda, \nabla_\alpha, \nabla_\mu)$ de cette fonction au point $(\lambda_0, \alpha_0, \mu_0, 0)$ s'exprime de la façon suivante:

$$\Omega = \begin{pmatrix} H & A^t & C_F^t & C_L^t \\ A & 0 & 0 & 0 \\ T_F & 0 & 0 & 0 \\ 0 & 0 & 0 & T_L \end{pmatrix},$$

où T_F est la matrice constituée par les lignes de C_F telle que:

$$T_{F_i} = \mu_{F_i} C_{F_i}, \quad \forall i \in I(\lambda).$$

T_L est une matrice diagonale de taille $p - n - |I(\lambda)|$ telle que:

$$T_{L_{ii}} = (C_L \lambda - b_L)_i > 0.$$

On a

$$\det(\Omega) = \det(T_L) \det \begin{pmatrix} H & A^t & C_F^t \\ A & 0 & 0 \\ T_F & 0 & 0 \end{pmatrix}.$$

Or

$$\det \begin{pmatrix} H & A^t & C_F^t \\ A & 0 & 0 \\ T_F & 0 & 0 \end{pmatrix} = \left(\prod_{i \in I} \mu_i \right) \det \begin{pmatrix} H & A^t & C_F^t \\ A & 0 & 0 \\ C_F & 0 & 0 \end{pmatrix},$$

et

$$\det(T_L) = \prod_{i \notin I} (C_F \lambda - b_F)_i,$$

d'où

$$\det(\Omega) = \left(\prod_{i \in I} \mu_i \right) \left(\prod_{i \notin I} (C_F \lambda - b_F)_i \right) \det \begin{pmatrix} H & \tilde{A}^t \\ \tilde{A} & 0 \end{pmatrix},$$

où $\tilde{A} = \begin{pmatrix} A \\ C_F \end{pmatrix}$.

Le Jacobien est nul si l'un des termes est nul.

En appliquant le même raisonnement que dans le cas d'un problème d'optimisation à contraintes linéaires de type égalité, on démontre que:

$$\det \begin{pmatrix} H & \tilde{A}^t \\ \tilde{A} & 0 \end{pmatrix} \neq 0.$$

Le terme $(\prod_{i \in I} \mu_i)(\prod_{i \notin I} (C_F \lambda - b_F)_i)$ est nul s'il existe $i_0 \in I(\lambda)$ tel que $\mu_{i_0} = 0$. Dans ce cas, la solution est faiblement active. En d'autres termes, il existe une face contenant l'optimum telle que la contrainte relative à cette face ne soit pas utile. Si toutes les contraintes sont fortement actives, le Jacobien est non nul. On peut donc appliquer le théorème des fonctions implicites.

- il existe un voisinage V_0 de 0,
- il existe une fonction C^∞ :

$$\varepsilon \longrightarrow (\lambda_\varepsilon, \alpha_\varepsilon, \mu_\varepsilon)$$

telle que pour tout $\varepsilon \in V_0$, $(\lambda_\varepsilon, \alpha_\varepsilon, \mu_\varepsilon, \varepsilon)$ vérifie la relation (2.12). ■

On notera cette fonction implicite $\begin{pmatrix} \lambda \\ \alpha \\ \mu \end{pmatrix}$.

Remarque 6:

Il est clair que le voisinage V_0 dans le théorème précédent nous assure que l'ensemble des contraintes actives à l'optimum ne change pas lorsque ε décrit V_0 .

- Jacobien par rapport à ε de la fonction $\begin{pmatrix} \lambda \\ \alpha \\ \mu \end{pmatrix}$

En dérivant les équations à l'optimum par rapport à ε au point $(\lambda(0), \alpha(0), \mu(0))$, on obtient la relation:

$$\begin{pmatrix} H & A^t & C_F^t & C_L^t \\ A & 0 & 0 & 0 \\ T_F & 0 & 0 & 0 \\ 0 & 0 & 0 & T_L \end{pmatrix} \nabla_\varepsilon \begin{pmatrix} \lambda(0) \\ \alpha(0) \\ \mu_F(0) \\ \mu_L(0) \end{pmatrix} = - \begin{pmatrix} \nabla_{\lambda_\varepsilon}^2 J(\lambda(0), 0) + \tilde{\alpha} \\ \Lambda_0 \\ 0 \\ 0 \end{pmatrix}.$$

où $\tilde{\alpha}$ est la matrice qui ε associe le vecteur $\varepsilon^t \alpha$.

On a donc $\nabla_\varepsilon \mu_L = 0$, d'où:

$$\nabla_\varepsilon \begin{pmatrix} \lambda(0) \\ \alpha(0) \\ \mu_F(0) \end{pmatrix} = - \begin{pmatrix} H & A^t & C_F^t \\ A & 0 & 0 \\ T_F & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} \nabla_{\lambda_\varepsilon}^2 J(\lambda(0), 0) + \tilde{\alpha} \\ \Lambda_0 \\ 0 \end{pmatrix},$$

ou encore:

$$\nabla_\varepsilon \begin{pmatrix} \lambda(0) \\ \alpha(0) \\ \mu_F(0) \end{pmatrix} = - \begin{pmatrix} H & A^t & C_F^t \\ A & 0 & 0 \\ C_F & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} \nabla_{\lambda_\varepsilon}^2 J(\lambda(0), 0) + \tilde{\alpha} \\ \Lambda_0 \\ 0 \end{pmatrix}. \quad (2.14)$$

De la même manière que précédemment, on tire que:

$$\nabla_\varepsilon \lambda(0) = -P \begin{pmatrix} H & A^t & C_F^t \\ A & 0 & 0 \\ C_F & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} \nabla_{\lambda_\varepsilon}^2 J(\lambda(0), 0) + \tilde{\alpha} \\ \Lambda_0 \\ 0 \end{pmatrix}.$$

Les valeurs singulières du Jacobien donnent la sensibilité par rapport à ε .

3. 2. Région de confiance

La régularité étant démontrée dans le cas des contraintes linéaires de type égalité et inégalité, on va caractériser dans cette partie le domaine de variation de λ_ε quand ε décrit un certain ellipsoïde. On supposera que les composantes de ε suivent une loi normale. Σ désignera la matrice de variance covariance.

Cas des contraintes de type égalité

Hypothèse 1: On suppose que les vecteurs N_i et ε sont tels que les déterminants ne changent pas de signe.

En faisant un développement du premier ordre au voisinage de 0 de la solution, on obtient:

$$\begin{pmatrix} \lambda(\varepsilon) \\ \alpha(\varepsilon) \end{pmatrix} - \begin{pmatrix} \lambda(0) \\ \alpha(0) \end{pmatrix} = \begin{pmatrix} \nabla \lambda(0) \\ \nabla \alpha(0) \end{pmatrix} \varepsilon + O(\|\varepsilon\|^2).$$

Notons:

$$\Delta\lambda = \lambda(\varepsilon) - \lambda(0),$$

$$\Delta\alpha = \alpha(\varepsilon) - \alpha(0).$$

En remplaçant dans l'équation (2.11):

$$\begin{pmatrix} H & A^t \\ A & 0 \end{pmatrix} \begin{pmatrix} \Delta\lambda(\varepsilon) \\ \Delta\alpha \end{pmatrix} = - \begin{pmatrix} L \\ \Lambda_0 \end{pmatrix} \varepsilon + O(\|\varepsilon\|^2). \quad (2.15)$$

où L désigne la matrice $\nabla_{\lambda\varepsilon}\mathcal{L}(\lambda(0), \alpha(0), 0)$.

Quand ε varie dans un ellipsoïde, on veut caractériser le domaine décrit par $\Delta\lambda$. Soit B^ρ l'ellipsoïde tel que ε lui appartient avec une certaine probabilité. Alors B^ρ est définie par:

$$B^\rho = \{y \mid \langle y, \Sigma^{-1}y \rangle \leq \rho^2\}.$$

Soit T la matrice définie par:

$$T = - \begin{pmatrix} H & A^t \\ A & 0 \end{pmatrix}^{-1} \begin{pmatrix} L \\ \Lambda_0 \end{pmatrix},$$

alors $\begin{pmatrix} \Delta\lambda \\ \Delta\alpha \end{pmatrix}$ varie dans l'image de B^ρ par T .

Caractérisation de l'image linéaire d'un ellipsoïde

Soit B l'ellipsoïde défini par

$$B = \{x \mid \langle x, Qx \rangle \leq \rho^2\},$$

où Q est une matrice symétrique définie positive. Notons $\|\cdot\|_Q$, la norme définie par le produit scalaire $\langle \cdot, Q \cdot \rangle$.

Soit $T: \mathbb{R}^{np} \longrightarrow \mathbb{R}^{n+p}$ une application linéaire surjective, non injective.

Soit $y \in \mathbb{R}^{n+p}$. Considérons l'ensemble:

$$T^{-1}(\{y\}) = \{x \in X \mid Tx = y\},$$

qui est une variété affine non vide parallèle à $N(T)$ dans $\mathbb{R}^{np}$.

On aura $y \in T(B)$ si et seulement si l'élément x^* de $T^{-1}(\{y\})$, de norme $\|\cdot\|_Q$ minimum, appartient à B .

Cet élément x^* n'est autre que $T_Q^+(y)$, où T_Q^+ est le pseudo-inverse de T (si $\mathbb{R}^{np}$ est muni de la norme $\|\cdot\|_Q$).

Pour caractériser $T(B)$, on a:

$$\begin{aligned} y \in T(B) &\iff \|T_Q^+(y)\|_Q^2 \leq \rho^2 \\ &\iff \langle T_Q^+(y), QT_Q^+(y) \rangle \leq \rho^2 \\ &\iff \langle y, T_Q^{+t}QT_Q^+y \rangle \leq \rho^2. \end{aligned}$$

Donc:

$$T(B) = \{ y \mid \langle y, T_Q^{+t} Q T_Q^+ y \rangle \leq \rho^2 \}.$$

• **Caractérisation de T_Q^+**

Soit x^* l'élément de $T^{-1}(\{y\})$ de norme $\|\cdot\|_Q$ minimum.
 x^* est orthogonal à $N(T)$ au sens de la norme définie par la matrice Q .
 L'orthogonal de $N(T)$ au sens $\|\cdot\|_Q$ est $Im(Q^{-1}T^t)$, l'image de $\mathbb{R}^{np}$ par $(Q^{-1}T^t)$.
 En effet, pour tout élément z de $N(T)$, on a:

$$Tz = 0,$$

donc

$$TQ^{-1}Qz = 0.$$

Les lignes de TQ^{-1} sont orthogonales à $N(T)$. Elles définissent donc une base de l'espace orthogonal à $N(T)$ (si $\mathbb{R}^{np}$ est muni de la norme $\|\cdot\|_Q$). Donc

$$x^* \in Im(Q^{-1}T^t).$$

Il existe v tel que

$$x^* = Q^{-1}T^t v,$$

donc

$$y = TQ^{-1}T^t v.$$

d'où

$$v = (TQ^{-1}T^t)^{-1}Ty$$

et finalement:

$$x^* = Q^{-1}T^t(TQ^{-1}T^t)^{-1}y.$$

Le pseudo-inverse au sens de la norme $\|\cdot\|_Q$ s'écrit:

$$T_Q^+ = Q^{-1}T^t(TQ^{-1}T^t)^{-1}. \quad (2.16)$$

En remplaçant T_Q^+ par sa valeur, on obtient:

$$T_Q^+ Q T_Q^+ = (TQ^{-1}T^t)^{-1},$$

donc $T(B)$ est l'ellipsoïde défini par:

$$T(B) = \{ y \mid \langle y, (TQ^{-1}T^t)^{-1} y \rangle \leq \rho^2 \}. \quad (2.17)$$

Dans notre cas $Q = \Sigma^{-1}$. On en conclut alors que l'image de l'ellipsoïde B par T est l'ellipsoïde défini par:

$$\{ y \mid y^t (T \Sigma T^t)^{-1} y \leq \rho^2 \},$$

qui est l'ellipsoïde de variation du vecteur $\begin{pmatrix} \Delta\lambda \\ \Delta\alpha \end{pmatrix}$.

Pour chercher le domaine de variation relatif à $\Delta\lambda$, il suffit de faire un raisonnement analogue au précédent avec P , la projection définie par :

$$P(\lambda, \alpha) = \lambda,$$

$\Delta\lambda$ décrit l'image de B par l'application PT . D'après (2.17):

$$PT(B) = \{ y \mid y^t (PT\Sigma T^t P^t)^{-1} y \leq \rho^2 \}. \quad (2.18)$$

Cas des contraintes de type inégalité:

En plus de l'hypothèse 1 faite dans le cas des contraintes du type égalité, on fait l'hypothèse suivante:

Hypothèse 2: On suppose que les erreurs ε ne font pas changer l'ensemble des contraintes actives à l'optimum.

Le développement au premier ordre de $\begin{pmatrix} \lambda(\varepsilon) \\ \alpha(\varepsilon) \\ \mu_F(\varepsilon) \end{pmatrix}$ au voisinage de 0 donne:

$$\begin{pmatrix} \lambda(\varepsilon) \\ \alpha(\varepsilon) \\ \mu_F(\varepsilon) \end{pmatrix} - \begin{pmatrix} \lambda(0) \\ \alpha(0) \\ \mu_F(0) \end{pmatrix} = \begin{pmatrix} \nabla\lambda(0) \\ \nabla\alpha(0) \\ \nabla\mu_F(0) \end{pmatrix} \varepsilon + O(\|\varepsilon\|^2) \quad (2.19)$$

En utilisant la relation (2.14), on a la relation au premier ordre suivante:

$$\begin{pmatrix} \Delta\lambda \\ \Delta\alpha \\ \Delta\mu_F \end{pmatrix} = - \begin{pmatrix} H & A^t & C_F^t \\ A & 0 & 0 \\ C_F & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} L \\ \Lambda_0 \\ 0 \end{pmatrix} \varepsilon.$$

où L est la matrice $\nabla_{\lambda_\varepsilon}^2 J(\lambda(0), 0) + \tilde{\alpha}$

On est ramené à un cas similaire au cas des contraintes de type égalité, avec

$$T = - \begin{pmatrix} H & A^t & C_F^t \\ A & 0 & 0 \\ C_F & 0 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} L \\ \Lambda_0 \\ 0 \end{pmatrix}.$$

Quand ε décrit B^ρ , $\Delta\lambda$ décrit son image par PT , où P est la projection définie par, $P(\lambda, \alpha, \mu) = \lambda$.

La région de confiance pour $\Delta\lambda$ au niveau $(1 - \delta)$ est définie par (cf [Ant92]) :

$$\Delta\lambda^t (PT\Sigma T^t P^t)^{-1} \Delta\lambda \leq (\chi_p^2)^{-1} (1 - \delta). \quad (2.20)$$

où χ_p^2 désigne la fonction de répartition de la loi du χ^2 à p degrés de liberté.

Pour $\delta = 0.05$, la relation (2.20) détermine l'ellipsoïde où le vecteur $\Delta\lambda$ varie avec une probabilité de 0.95.

De même, la région de confiance associée à chaque couple de composantes $(\Delta\lambda_i, \Delta\lambda_j)$, au niveau $(1 - \delta)$ est donnée par:

$$\Delta\lambda^t (P_{ij} T \Sigma T^t P_{ij}^t)^{-1} \Delta\lambda \leq (\chi_2^2)^{-1} (1 - \delta), \quad (2.21)$$

où

$$P_{ij}(\lambda, \alpha, \mu) = (\lambda_i, \lambda_j), \quad i < j, \quad i, j = 1, \dots, p.$$

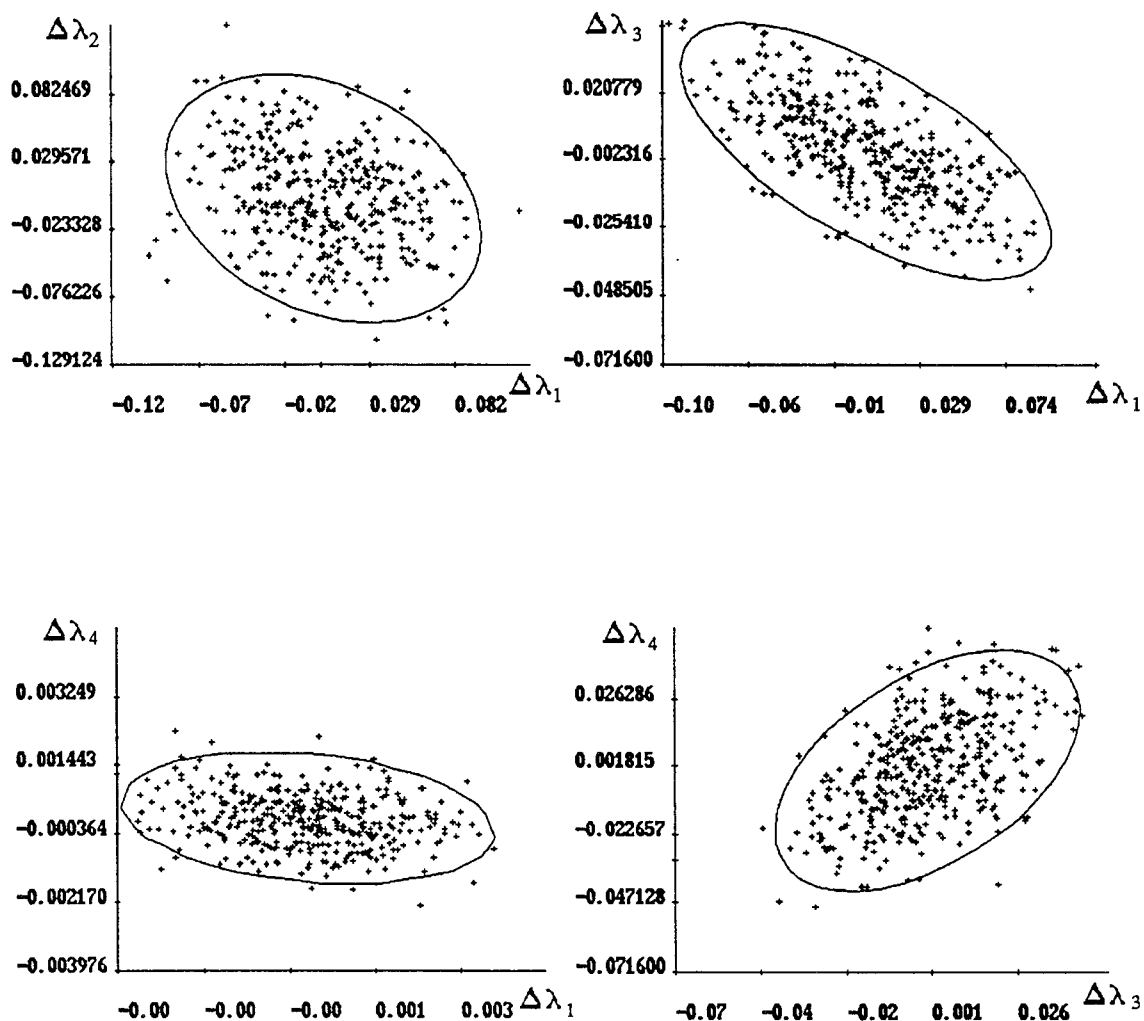
Exemples

Dans les exemples traités, on suppose que les erreurs ε_{ij} sont gaussiennes et de variances proportionnelles aux composantes. Elles sont choisies de la forme:

$$\varepsilon_{ij} = 0.05 \times A_{ij} \times X,$$

où X est une variable aléatoire gaussienne $\mathcal{N}(0, 1)$.

Dans le graphique, chaque point est solution du problème d'optimisation perturbé. Le nombre de perturbations, dans les exemples choisis, est 400. L'ellipse sur le plan défini par deux composantes $\Delta\lambda_i$ et $\Delta\lambda_j$, est donnée par la relation (2.21), avec $\delta = 0.05$.

Exemple 1FIGURE .5. Variations de $\Delta\lambda$.

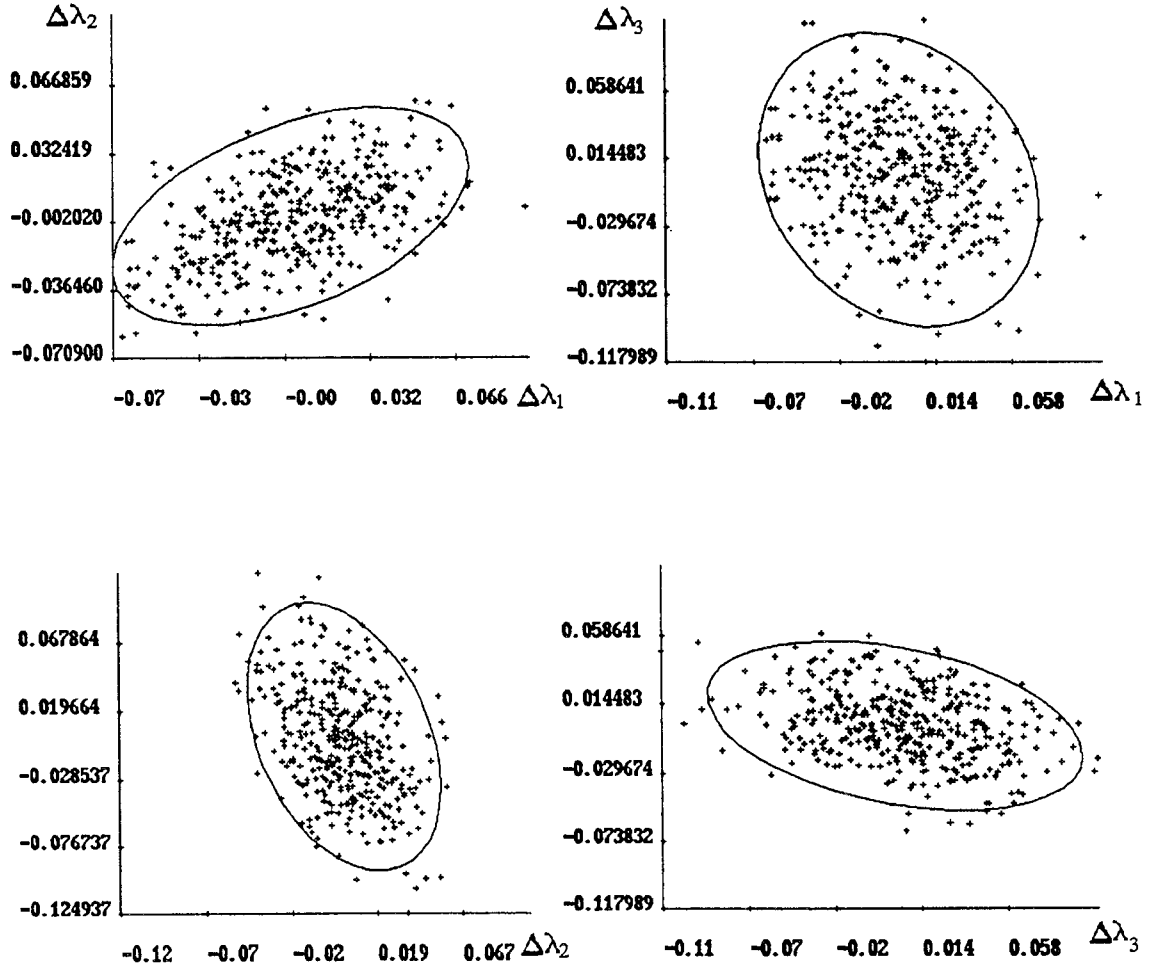
$$N_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad N_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad N_3 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \quad N_4 = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} \quad \text{et} \quad M = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

La solution λ du problème d'optimisation non perturbé est:

$$\lambda = \begin{pmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{pmatrix}$$

Remarquons qu'aucune contrainte du type inégalité n'est active.

Exemple 2

FIGURE .6. Variations de $\Delta\lambda$.

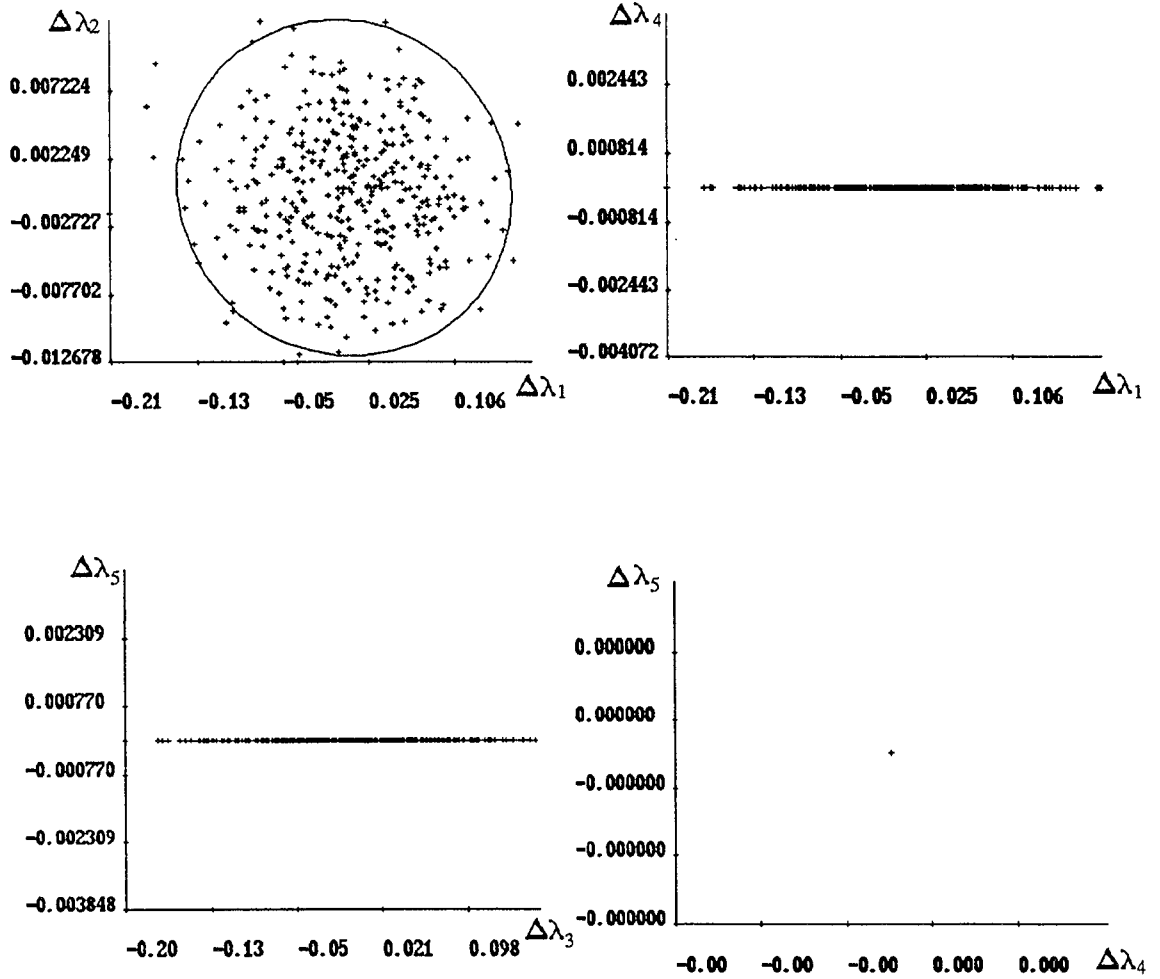
$$N_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, N_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, N_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, N_4 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, N_5 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \text{ et } M = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}.$$

La solution λ du problème d'optimisation non perturbé est:

$$\lambda = \begin{pmatrix} 0.6 \\ 0.6 \\ 0.2 \\ 0.2 \\ 0.6 \end{pmatrix}$$

Même remarque: aucune contrainte du type inégalité n'est active.

Exemple 3

FIGURE .7. Variations de $\Delta\lambda$.

$$N_1 = \begin{pmatrix} 20 \\ 10 \\ 0 \end{pmatrix}, N_2 = \begin{pmatrix} 20 \\ 20 \\ 10 \end{pmatrix}, N_3 = \begin{pmatrix} 10 \\ 20 \\ 0 \end{pmatrix}, N_4 = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, N_5 = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}, \text{ et } M = \begin{pmatrix} 25 \\ 25 \\ 1 \end{pmatrix}.$$

La solution λ du problème d'optimisation non perturbé est:

$$\lambda = \begin{pmatrix} \frac{23}{30} \\ 0.1 \\ \frac{23}{30} \\ 0 \\ 0 \end{pmatrix}.$$

Les contraintes $\lambda_4 \geq 0$ et $\lambda_5 \geq 0$ sont actives. L'hypothèse 2 impose à ces deux contraintes de rester actives; on aura donc:

$$\Delta\lambda_4 = 0 \quad \text{et} \quad \Delta\lambda_5 = 0.$$

quand ε décrit son domaine. La projection de l'ellipsoïde sur le plan relatif à la quatrième et la cinquième composantes, est réduit à un point.

Variations du mélange à réaliser

En commettant des erreurs ε pour définir les caractéristiques des produits de base, on obtient la solution λ_ε comme vecteur de fractions optimales pour fabriquer M . Dans le processus de fabrication, ceci se traduit par prendre les fractions λ_ε des vecteurs N_i ; le produit réalisé est donc:

$$M_\varepsilon = \sum_{i=1}^p \lambda_{\varepsilon_i} N_i.$$

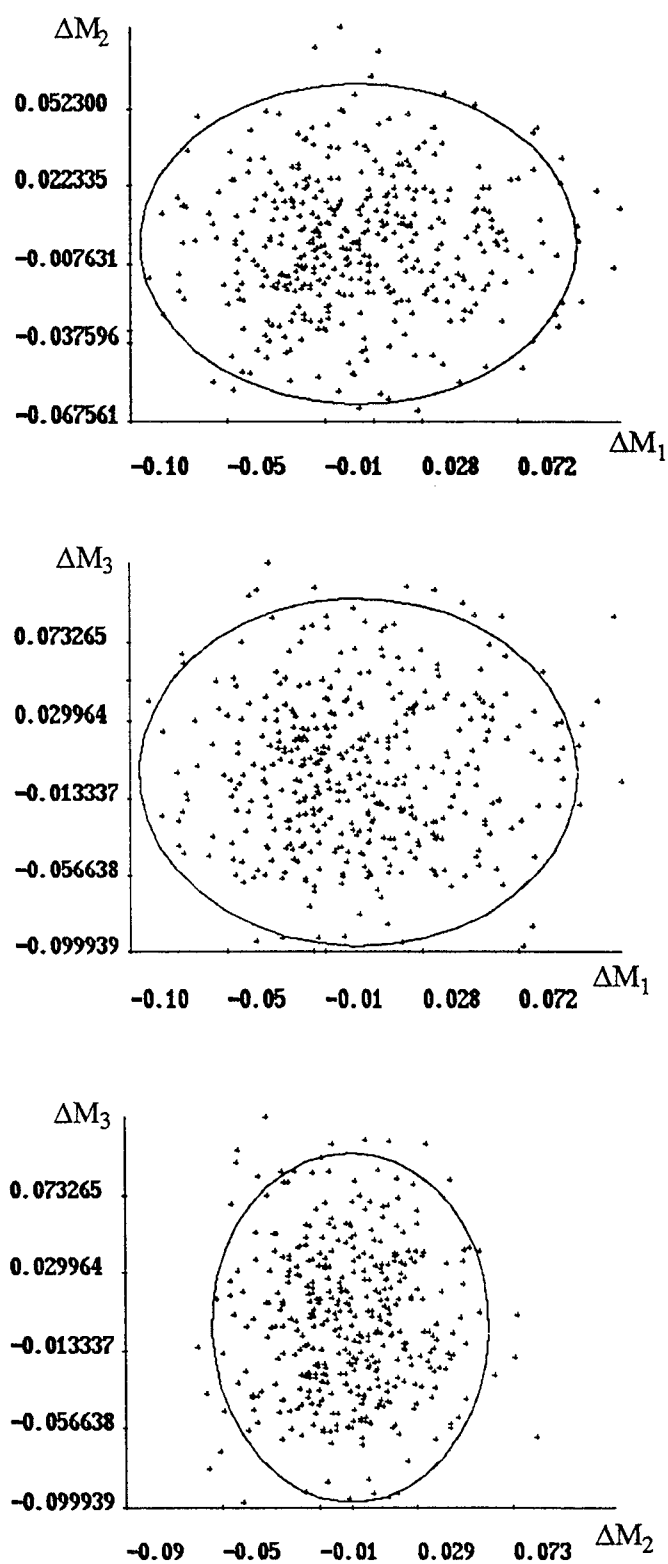
La variation $\Delta M = M(\varepsilon) - M$ est donnée par:

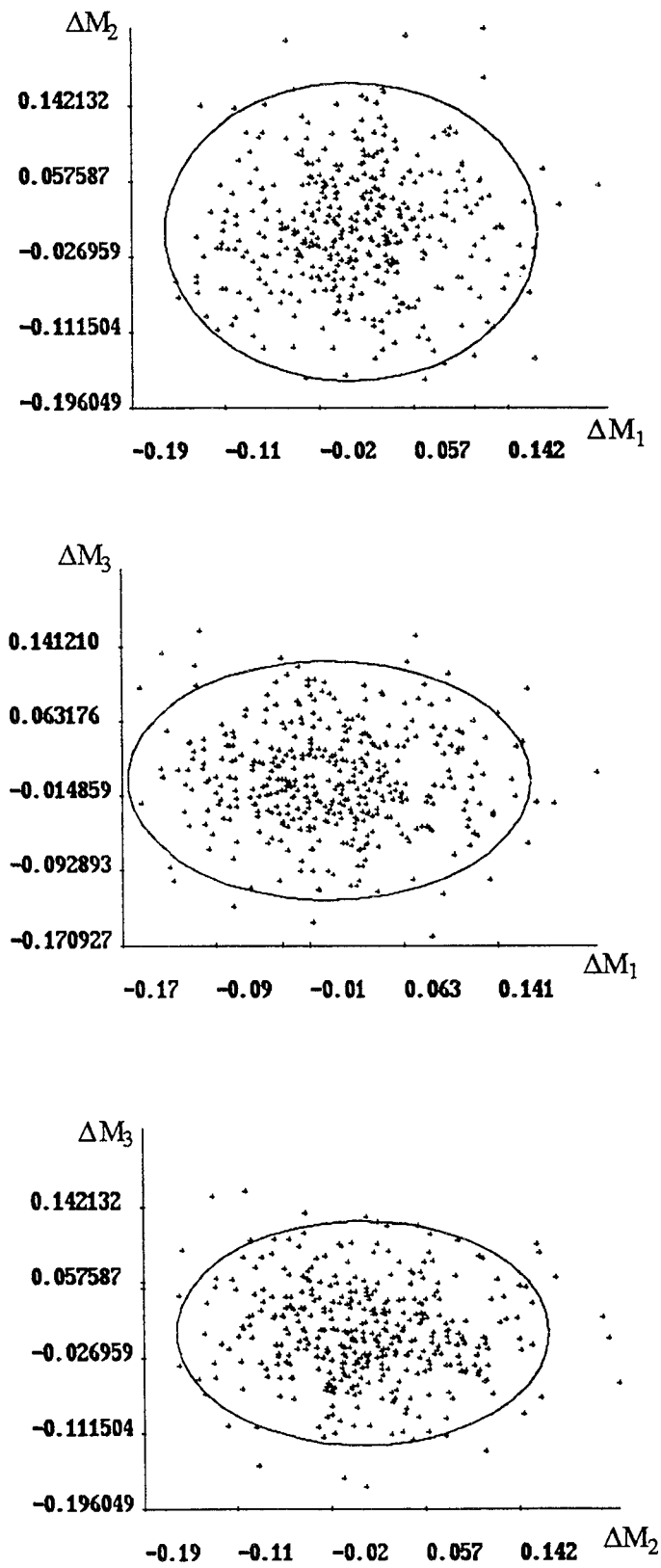
$$\Delta M = A \Delta \lambda.$$

Quand ε varie dans la boule B^ρ , $\Delta \lambda$ varie dans la boule donnée par (2.17), et ΔM décrit l'image de cette boule par la matrice A .

La région de confiance pour ΔM au niveau $(1 - \delta)$ est définie par

$$\Delta M^t (A P T \Sigma T^t P^t A^t)^{-1} \Delta M \leq (\chi_n^2)^{-1} (1 - \delta). \quad (2.22)$$

FIGURE .8. Variation de M (exemple 1).

FIGURE .9. Variation de M (exemple 2).

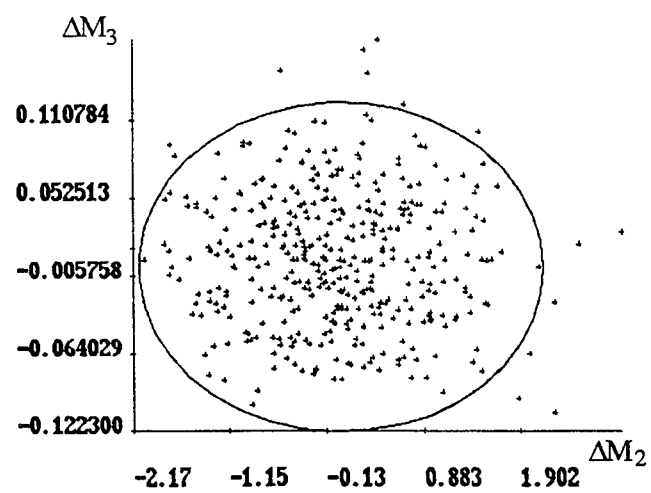
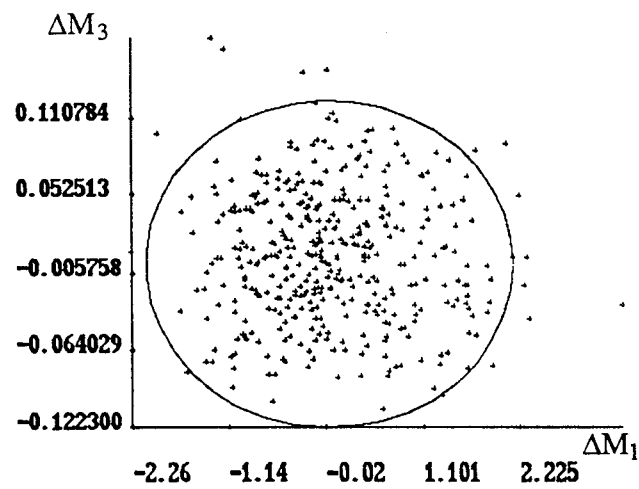
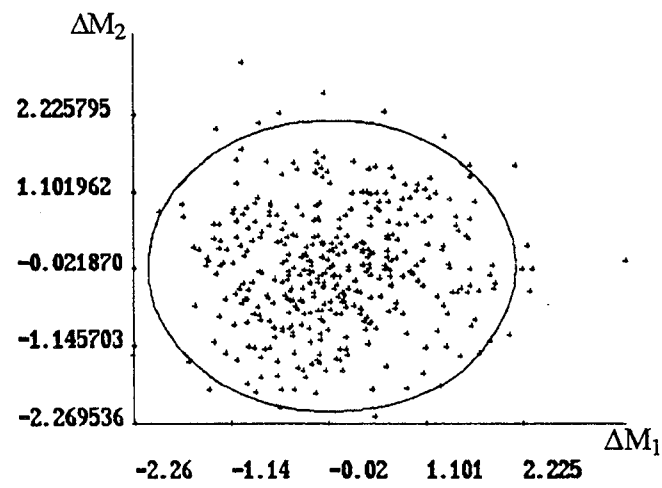


FIGURE .10. Variation de M (exemple 3).

Chapitre 3

Zonoïdes Associés à une Courbe

Introduction

Lors d'un processus discret de fabrication de mélanges à partir d'un certain nombre p fixé de produits de base, la modélisation géométrique a amené à utiliser les zonotopes. Cependant, en chimie, il y a des processus qui sont continus ou diffus. Dans ce cas, P. Valentin a introduit la notion de zonoïde associé à une famille différentielle ou ce que nous appellerons dans cette étude zonoïde associé à une fonction paramétrique régulière.

Un exemple de processus diffus est le principe de séparation par distillation. Ce principe consiste à mettre un produit dans le bouilleur d'une colonne à distiller, et récupérer en haut de la colonne, dans un laps de temps dt donné, le produit de distillation. Le produit récupéré a un certain nombre n de spécifications qui varient au cours du temps. A ces spécifications, on associe une fonction additive F (voir définition de l'additivité). Le mélange produit entre les instants t et $t + dt$ est représenté par $F(t)dt$.

Nous commencerons dans la première partie par un bref rappel sur les zonoïdes. Dans la seconde partie nous introduisons les zonoïdes définis par des fonctions paramétriques régulières, et nous donnerons quelques propriétés et résultats. Dans la troisième partie, nous étudierons la régularité de la surface constituant le bord du zonoïde connaissant celle de la fonction F . Pour cela, nous définissons une paramétrisation de cette surface et sous certaines hypothèses nous donnerons un résultat sur sa régularité. Dans la partie 4, nous étudierons le cas particulier où F vérifie la propriété de la sélectivité totalement régulière (STR): nous donnerons une paramétrisation plus simple de la surface et un résultat de régularité. Dans la dernière partie, dans le cas où F est STR, nous traiterons le problème de l'intersection d'un plan avec le zonoïde associé à F .

Soit $E = \mathbb{R}^n$ l'espace euclidien de dimension n muni du produit scalaire usuel $\langle ., . \rangle$. Soient $\mathcal{K}$ l'ensemble des convexes compacts de E , et $\mathcal{K}_S$ l'ensemble des convexes compacts à symétrie centrale. On désigne par B la boule unité dans E , et S la sphère unité, i.e., la frontière de B .

1. Zonoïde, rappels et Définitions

1. 1. Définitions et propriétés

Dans cette partie, nous rappelons quelques définitions et résultats sur les zonoïdes.

Définition 3. 1 (Distance de Hausdorff) Soient K et L deux éléments de $\mathcal{K}$. La métrique de Hausdorff est définie par:

$$\text{haus}(K, L) = \min \{ \varepsilon > 0 \mid K \subset L + \varepsilon B \quad \text{et} \quad L \subset K + \varepsilon B \}$$

On considère $\mathcal{K}$ muni de la métrique de Hausdorff.

Définition 3. 2 Un zonoïde est défini comme étant la limite, au sens de la métrique de Hausdorff, d'une séquence de zonotopes.

Remarque 1: De ce fait l'ensemble des zonoïdes constitue la fermeture, au sens de la métrique de Hausdorff dans $\mathcal{K}$, de l'ensemble des zonotopes.

Proposition 3. 3 (Bolker 69) Un zonoïde est un ensemble convexe compact à symétrie centrale.

Les zonoïdes apparaissent dans d'autres contextes. Bolker, dans [Bol69], fait un lien entre les zonoïdes et la théorie de la mesure et démontre le théorème suivant:

Théorème 3.4 Un zonoïde est l'image d'une mesure vectorielle non atomique.

Propriété 3. 5 L'image par une application affine d'un zonoïde est un zonoïde.

D'autres auteurs ont donné des caractérisations des zonoïdes; citons par exemple Weil 76 et Goodey 77 qui, à partir des distributions ont défini les zonoïdes par des propriétés sur les volumes mixtes.

1. 2. Exemples

1- Pour $n \geq 3$,

$$\left\{ x = (x_1, x_2, \dots, x_n); \quad \|x\|_1 = \left(\sum_{i=1}^n |x_i| \right) \leq 1 \right\}$$

n'est pas un zonoïde.

Plus généralement: l'ensemble

$$\left\{ x = (x_1, x_2, \dots, x_n); \quad \|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \leq 1 \right\}$$

est un zonoïde si et seulement si $p \geq 2$.

2- Un polytope est un zonoïde si et seulement si c'est un zonotope.

3- Dans $\mathbb{R}^2$, tout ensemble borné convexe à symétrie centrale est un zonoïde.

4- Le polaire d'un zonoïde n'est pas, en général, un zonoïde.

1. 3. Fonction support

Définition 3. 6 Soit $K \in \mathcal{K}$; la fonction support H_K de K dans la direction $\xi \in \mathbb{R}^n$ est définie par:

$$H_K(\xi) = \sup \{ \langle x, \xi \rangle \mid x \in K \}.$$

Dans le cas d'un zonoïde la fonction support s'exprime sous la forme suivante (cf [Lin75], [Wei76]):

Proposition 3. 7 Soit Z un zonoïde ayant 0 comme centre de symétrie; alors il existe une mesure de Borel positive ρ sur S telle que la fonction support de Z s'écrit:

$$H_Z(\xi) = \int_S |\langle x, \xi \rangle| \rho(dx), \quad \forall \xi \in E.$$

2. Zonoïdes associés à une courbe paramétrique régulière

2. 1. Définitions et propriétés

On désigne par Λ l'ensemble des fonctions de $L^\infty(I)$ à valeurs dans $[0, 1]$.

Définition 3. 8 Soit F une application définie de $I = [a, b]$ à valeurs dans $\mathbb{R}^n$ satisfaisant les conditions suivantes que l'on appellera $(\mathcal{H}_1)$:

- F est injective,
- F est C^p ,
- $F'(t) \neq 0$ pour tout $t \in I$.

F est une représentation paramétrique régulière de classe C^p d'une courbe Γ .

Définition 3. 9 L'ensemble défini par:

$$Z(F) = \left\{ M \in E \quad : \quad M = \int_I \lambda(t)F(t)dt, \quad \lambda \in \Lambda \right\},$$

est appelé le zonoïde associé à F .

Remarque 2: $Z(F)$ peut être noté aussi $Z(\Gamma)$ car il ne dépend pas de la paramétrisation régulière choisie.

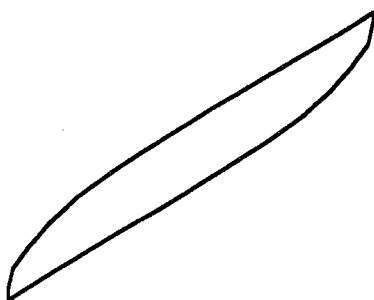


FIGURE .1. Exemple de zonoïde dans $\mathbb{R}^2$.

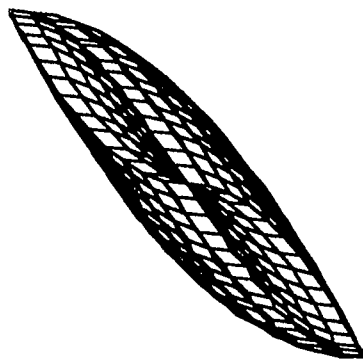


FIGURE .2. Exemple de zonoïde dans $\mathbb{R}^3$.

Propriété 3. 10 Le zonoïde $Z(F)$ admet comme symétrie centrale le point

$$C = \frac{1}{2} \int_I F(t)dt.$$

Démonstration : Soit M un point de $Z(F)$, il existe une fonction $\lambda \in \Lambda$, telle que $M = \int_I \lambda(t)F(t)dt$.

Or la fonction $1 - \lambda$ appartient à Λ , donc le point M^* défini par:

$$\begin{aligned} M^* &= \int_I (1 - \lambda(t)) F(t) dt \\ &= \int_I F(t) dt - M \\ &= 2C - M. \end{aligned}$$

est aussi un point du zonoïde. ■

Définition 3. 11 On appelle point distal, le point de $Z(F)$, noté S et défini par:

$$S = \int_I F(t) dt.$$

Propriété 3. 12 Si T est une application linéaire de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ alors $T(Z(F))$ est un zonoïde dans $\mathbb{R}^p$ et on a:

$$T(Z(F)) = Z(T(F)).$$

Démonstration : immédiate.

• **Fonction support :**

Proposition 3. 13 La fonction support du zonoïde $Z(F)$ dans la direction u est donnée par:

$$H_{Z(F)}(u) = \int_{\omega_u} \langle F(t), u \rangle dt,$$

où $\omega_u = \{ t \in I, \langle F(t), u \rangle \geq 0 \}$.

Démonstration : Soit $u \in \mathbb{R}^n$; on a:

$$\begin{aligned} H_{Z(F)}(u) &= \max_{x \in Z(F)} \langle x, u \rangle \\ &= \max_{\lambda \in \Lambda} \left\langle \int_I \lambda(t) F(t) dt, u \right\rangle \\ &= \max_{\lambda \in \Lambda} \int_I \lambda(t) \langle F(t), u \rangle dt \\ &= \int_I \langle F(t), u \rangle_+ dt \\ &= \int_{\omega_u} \langle F(t), u \rangle dt. \quad \blacksquare \end{aligned}$$

Remarque 3: Lorsque l'ensemble $\{ t \in I \mid \langle F(t), u \rangle = 0 \}$ est de mesure non nulle, alors la fonction $x \rightarrow \langle x, u \rangle$ peut atteindre la valeur $H_{Z(F)}(u)$ en plusieurs points x de $Z(F)$. Dans ce cas, l'ensemble des points d'appui du plan de normale u sur la surface du zonoïde constitue une face.

2. 2. Zonoïdes associés à une courbe et mesures vectorielles

Le théorème suivant, donnera un lien entre les zonoïdes définis dans la partie 1, et ceux associés à une fonction. Le résultat est particulièrement dans le sens du théorème 3. 4 .

Soit $\mathcal{B}(I)$, la tribu borélienne sur I . On a le résultat suivant:

Théorème 3.14 *Le zonoïde $Z(F)$ est l'image de $\mathcal{B}(I)$ par la mesure vectorielle ayant F comme densité de mesure par rapport à la mesure de Borel Lebesgue.*

Démonstration :

On désigne par μ la mesure de Borel Lebesgue; soit ν la mesure vectorielle définie par $\nu = F.\mu$.

Soit Z_ν le zonoïde défini comme étant l'image de $\mathcal{B}(I)$ par la mesure ν . Nous allons montrer que $Z_\nu = Z(F)$.

Pour tout $M \in Z_\nu$, il existe $\omega \in \mathcal{B}(I)$ tel que

$$\begin{aligned} M &= \int_{\omega} F(t) dt \\ &= \int_I \chi_{\omega}(t) F(t) dt \end{aligned}$$

(χ_{ω} désigne ici la fonction indicatrice de ω). Or $\chi_{\omega} \in \Lambda$, donc $M \in Z(F)$.
D'où:

$$Z_\nu \subset Z(F).$$

Réciproquement, montrons que $Z(F) \subset Z_\nu$. Pour cela, nous utiliserons le lemme suivant; pour une démonstration, nous renvoyons à Rockafellar dans [Rock70].

Lemme 3. 15 *Soient K_1 et K_2 sont deux ensembles convexes fermés non vides. Alors $K_1 \subset K_2$ si et seulement si pour tout $\xi \in S$, $H_{K_1}(\xi) \leq H_{K_2}(\xi)$.*

Soit $\xi \in S$, d'après la proposition proposition 3.13:

$$H_{Z(F)}(\xi) = \left\langle \int_{\omega_\xi} F(t) dt, \xi \right\rangle,$$

et par définition de ν , on a:

$$\begin{aligned} H_{Z(F)}(\xi) &= \langle \nu(\omega_\xi), \xi \rangle \\ &\leq \sup_{\omega \in \mathcal{B}(I)} \langle \nu(\omega), \xi \rangle. \end{aligned}$$

Donc pour tout $\xi \in \mathcal{S}$, on a:

$$H_{Z(F)}(\xi) \leq H_{Z_\nu}(\xi).$$

D'après le lemme 3.15:

$$Z(F) \subset Z_\nu,$$

d'où

$$Z(F) = Z_\nu. \blacksquare$$

Dans la suite, nous allons considérer des zonoïdes dans $\mathbb{R}^3$. Pour tout $\eta \in \mathcal{S}$, nous désignons par Π_η , le plan de normale η qui passe par 0.

$$\Pi_\eta = \{x \in \mathbb{R}^3 \mid \langle x, \eta \rangle = 0\}.$$

2. 3. Diagramme projectif

On généralise le premier diagramme projectif, introduit par Coxeter dans [Cox62] dans le cas des zonoèdres, à un diagramme projectif continu. On suppose qu'il existe $\mu \in \mathbb{R}^3$ tel que $\langle F(t), \mu \rangle > 0$, pour tout $t \in I$.

Définition 3. 16 On appelle diagramme projectif associé au zonoïde $Z(F)$ la courbe Γ_p définie comme étant la projection conique de la courbe Γ sur le plan

$$P_\mu = \{ x \in \mathbb{R}^3 \quad : \quad \langle x, \mu \rangle = 1 \quad \}.$$

Elle est définie par:

$$F_p(t) = \frac{F(t)}{\langle F(t), \mu \rangle}, \quad \forall t \in I.$$

Dans la suite, on suppose que $\mu = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Ainsi,

$$\Gamma_p = \{ (x(t), y(t)); \quad t \in I \quad \}$$

où

$$x(t) = \frac{F_1(t)}{F_3(t)},$$

$$y(t) = \frac{F_2(t)}{F_3(t)}.$$

A chaque vecteur η , $\eta \neq \mu$, on associe la droite Δ_η , intersection du plan Π_η et du plan P_μ . Soit Π_η^+ (resp. Π_η^-) le demi espace défini par les éléments $x \in \mathbb{R}^3$ tels que $\langle x, \eta \rangle > 0$ (resp. $\langle x, \eta \rangle < 0$). De même, soit Δ_η^+ (resp. Δ_η^-) le demi plan défini par les éléments u de P_μ tels que $\langle u, \eta \rangle > 0$ (resp. $\langle u, \eta \rangle < 0$).

Soit $u \in P_\mu$; on a les équivalences suivantes:

$$\begin{aligned} u \in \Pi_\eta \quad (i.e., \quad u \in \Delta_\eta) &\iff \eta \in \Pi_u \\ u \in \Pi_\eta^+ \quad (i.e., \quad u \in \Delta_\eta^+) &\iff \eta \in \Pi_u^+ \\ u \in \Pi_\eta^- \quad (i.e., \quad u \in \Delta_\eta^-) &\iff \eta \in \Pi_u^-. \end{aligned}$$

On oriente la droite Δ_η de telle sorte que Δ_η^+ (resp. Δ_η^-) soit à droite de Δ_η (resp. à gauche de Δ_η) voir figure (3).

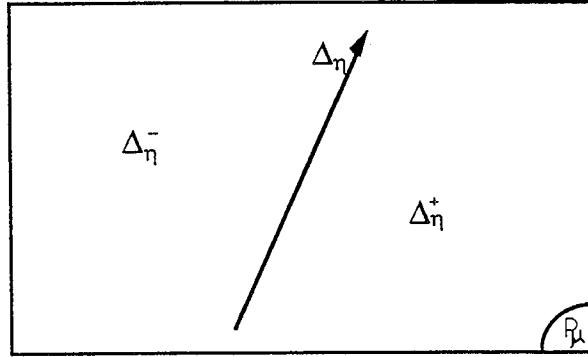


FIGURE .3. Orientation de Δ_η dans le P.D.P..

A chaque plan Π_η , on associe la droite orientée Δ_η , dans le diagramme projectif, intersection du plan Π_η et du plan P_μ . Si la courbe Γ et le plan Π_η se coupent en un point de paramètre t alors la courbe Γ_p et la droite Δ_η se coupent en un point ayant le même paramètre t .

Le point d'appui du plan de normale η sur $Z(F)$ est donné en intégrant F sur l'ensemble des points de I tel que le morceau de courbe de Γ_p correspondant se trouve du côté droit de la droite Δ_η (i.e., $\{t \in I \mid F_p(t) \in \Delta_\eta^+\}$).

Remarque 4: Le plan Π_μ (resp. $\Pi_{-\mu}$) n'est pas représenté par une droite dans le P.D.P.. Ce plan s'appuie sur le point distal S (resp. O) car on a supposé que $\langle F(t), \mu \rangle > 0, \forall t \in I$.

Exemple

Soit F la fonction définie de $[-1,1]$ vers $\mathbb{R}^3$ par:

$$F(t) = \begin{pmatrix} t \\ t^2 \\ 1 + t^2 \end{pmatrix}.$$

Le diagramme projectif associé est donné par:

$$\Gamma_p = \left\{ \begin{pmatrix} \frac{t}{1+t^2} \\ \frac{t^2}{1+t^2} \end{pmatrix}, \quad t \in [-1, 1] \right\}.$$

Dans la figure 4, nous donnons le diagramme projectif correspondant.

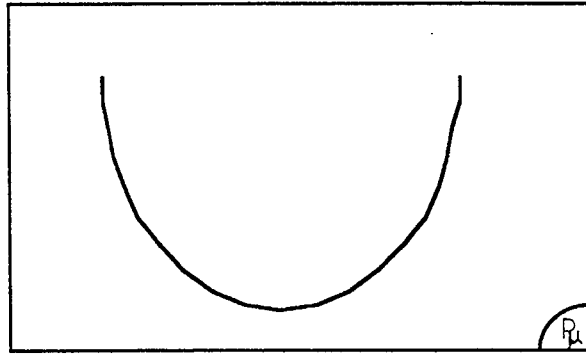


FIGURE .4. Diagramme projectif associé à F .

3. Surface Zonoïdale

3. 1. Définition d'une surface zonoïdale

Hypothèse ($\mathcal{H}_2$):

Le nombre de points d'intersection de la courbe $\Gamma = F(I)$ avec tout plan passant par O est plus petit qu'un nombre fixé m .

Pour définir la notion de surface zonoïdale, nous avons besoin du lemme suivant:

Lemme 3. 17 *Si l'hypothèse ($\mathcal{H}_2$) est vérifiée, alors $Z(F)$ est d'intérieur non vide*

Démonstration : $Z(F)$ est convexe; pour démontrer qu'il est d'intérieur non vide, il suffit de démontrer qu'il n'est pas contenu dans un plan.

Supposons qu'il existe un plan contenant $Z(F)$. Puisque $Z(F)$ contient O , alors ce plan passe par O . Soit η une normale à ce plan; alors on a:

$$H_{Z(F)}(\eta) = H_{Z(F)}(-\eta) = 0.$$

Donc

$$\int_I \langle F(t), \eta \rangle_+ dt = \int_I \langle F(t), -\eta \rangle_+ dt = 0.$$

d'où on peut déduire que:

$$\langle F(t), \eta \rangle = 0, \quad \forall t \in I,$$

ce qui entraîne que Γ est plane, ce qui contredit l'hypothèse $(\mathcal{H}_2)$. ■

Dans toute la suite, nous supposons que les hypothèses $\mathcal{H}_1$ et $\mathcal{H}_2$ sont vérifiées.

Définition 3. 18 On définit la surface zonoïdale $S(F)$ comme étant la frontière du zonoïde $Z(F)$.

Propriété 3. 19 L'unique point d'appui du plan de normale u sur le zonoïde $Z(F)$ est donné par:

$$\int_{\omega_u} F(t) dt.$$

Démonstration : De la proposition 3.13, la fonction support atteint sa valeur au point $\int_{\omega_u} F(t) dt$. Or d'après l'hypothèse précédente, le nombre de points d'intersection du plan de normale u avec la courbe Γ est inférieur à un nombre fixé m . Donc l'ensemble $\{t \mid \langle F(t), uu \rangle = 0\}$ est de mesure nulle, d'où l'unicité. ■

Remarques :

- Le point d'appui est obtenu en intégrant la fonction F sur les parties de I telles que $F(t)$ est du côté positif du plan Π_u .
- En général, ω_u n'est pas réduit à un intervalle sauf dans le cas de la sélectivité régulière (voir plus loin).

3. 2. Paramétrisation

On utilise les coordonnées sphériques (θ, φ) pour définir un point de $\mathcal{S}$. Soit $\eta(\theta, \varphi)$ le vecteur unitaire

$$\eta(\theta, \varphi) = (\cos\theta\cos\varphi, \sin\theta\cos\varphi, \sin\varphi),$$

et soit M l'application définie par:

$$\begin{aligned} M : \mathbb{R}/2\pi\mathbb{Z} \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] &\longrightarrow S(F), \\ (\theta, \varphi) &\longrightarrow \int_{\omega_{\eta(\theta, \varphi)}} F(t) dt. \end{aligned}$$

Proposition 3. 20 M est une application surjective de $\mathbb{R}/2\pi\mathbb{Z} \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ sur $S(F)$.

Démonstration : $Z(F)$ est un ensemble convexe d'intérieur non vide, donc par tout point P de $S(F)$ passe un hyperplan support à $Z(F)$. Soit $u \in \mathcal{S}$ une normale à ce plan; alors si (θ_0, φ_0) sont les coordonnées sphériques de u alors on a:

$$P = M(\theta_0, \varphi_0). \quad \blacksquare$$

Définition 3. 21 Soit t un paramètre tel que Π_η et Γ se coupent en $F(t)$. On dit qu'en t , $F(t)$ est une intersection de multiplicité $m < p$ (F est de classe C^p) si et seulement si:

$$\langle F^{(i)}(t), \eta \rangle = 0, \quad \forall i = 0, \dots, m-1,$$

et

$$\langle F^{(m)}(t), \eta \rangle \neq 0.$$

Une intersection est simple si $m = 1$.

Proposition 3. 22 Si le plan $\Pi_{\eta(\theta, \varphi)}$ a q intersections simples avec la courbe $\Gamma = F([a, b])$, correspondant aux paramètres $(a \leq t_1 < t_2 < \dots < t_q \leq b)$; alors:

$$M(\theta, \varphi) = \varepsilon_0 \int_a^{t_1} F(t) dt + \sum_{i=1}^{q-1} \varepsilon_i \int_{t_i}^{t_{i+1}} F(t) dt + \varepsilon_q \int_{t_q}^b F(t) dt,$$

où

$$\varepsilon_i = \begin{cases} 1 & \text{si } \langle F'(t_i), \eta(\theta, \varphi) \rangle > 0, \\ 0 & \text{sinon,} \end{cases} \quad i = 1, 2, \dots, q,$$

et

$$\varepsilon_0 = 1 - \varepsilon_1.$$

Démonstration : découle directement de la propriété 3.19.

Remarques :

- Les ε_i prennent alternativement les valeurs 0 et 1.
- Le point d'appui du plan de normale μ (resp. $-\mu$) sur $Z(F)$ est $\int_I F(t) dt$ (resp. 0); ces deux points sont appelés pôles du zonoïde.

3. 3. Courbe cumulée

Définition 3. 23 On appelle courbe cumulée, l'ensemble des points de $S(F)$ qui sont de la forme:

$$\int_a^u F(t) dt \quad \text{ou} \quad \int_u^b F(t) dt,$$

quand u décrit $[a, b]$.

Dans le diagramme projectif, lorsque la courbe Γ_p et la droite Δ_η n'ont qu'une seule intersection simple correspondant au paramètre u , le point d'appui du plan Π_η sur le zonoïde s'écrit:

$$\int_a^u F(t)dt \quad \text{ou} \quad \int_u^b F(t)dt.$$

Le plan de normale u s'appuie sur un point de la courbe cumulée.

Remarque 5: Par tout point de $S(F)$ appartenant à la courbe cumulée, passent plusieurs plans d'appui au zonoïde $Z(F)$.

Dans le diagramme projectif, ceci se traduit par le fait qu'il existe plusieurs droites que la courbe Γ_p traverse une seule fois, et en un même point.

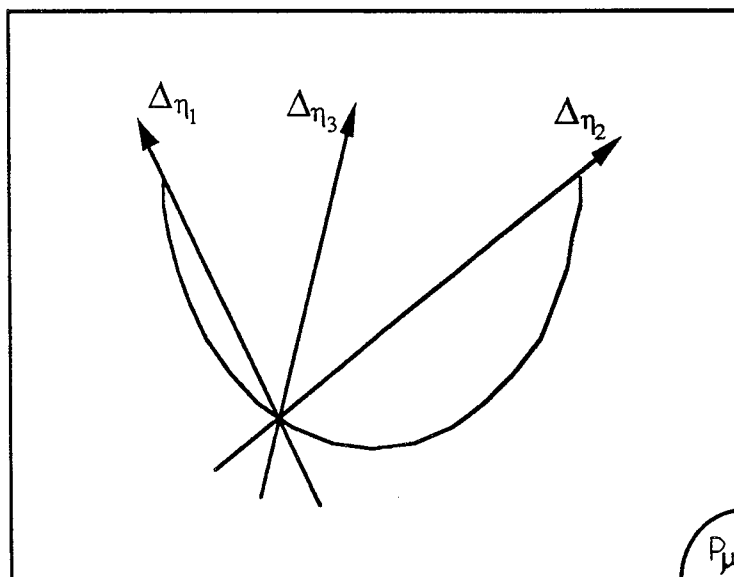


FIGURE 5. Les trois droites correspondent à un même point sur $S(F)$.

Dans la figure 5, les plans correspondants aux droites orientées Δ_{η_1} , Δ_{η_2} et $\underline{\Delta}_{\eta_3}$ admettent le même point d'appui sur $S(F)$.

3. 4. Régularité de la paramétrisation

Théorème 3.24 Soit $(\theta_0, \varphi_0) \in \Omega = \mathbb{R}/2\pi\mathbb{Z} \times (-\frac{\pi}{2}, \frac{\pi}{2})$. On suppose qu'il existe un voisinage ouvert $V_{(\theta_0, \varphi_0)}$ de (θ_0, φ_0) tel que pour tout $(\theta, \varphi) \in V_{(\theta_0, \varphi_0)}$, $\Pi_{\eta(\theta, \varphi)}$ et Γ ont exactement q intersections simples ($q \geq 2$).

Alors

- i) si F est C^p $p \geq 1$, alors M est aussi C^p dans un voisinage ouvert de (θ_0, φ_0) .
- ii) $\partial_\theta M(\theta, \varphi)$ et $\partial_\varphi M(\theta, \varphi)$ sont linéairement indépendants si et seulement si les points d'intersection de la courbe avec le plan de normale $\eta(\theta, \varphi)$ ne sont pas tous sur une même droite passant par 0.

Pour démontrer le théorème, on a besoin du lemme suivant:

Lemme 3. 25 *Soit t_0 une solution de l'équation:*

$$\langle \eta(\theta_0, \varphi_0), F(t) \rangle = 0, \quad (3.1)$$

telle que

$$\langle \eta(\theta_0, \varphi_0), F'(t_0) \rangle \neq 0. \quad (3.2)$$

Alors si F est de classe C^p , il existe un voisinage ouvert V_0 de (θ_0, φ_0) , un voisinage ouvert T_0 de t_0 et une fonction $g : V_0 \longrightarrow T_0$ de classe C^p telle que:

$$t_0 = g(\theta_0, \varphi_0) \text{ et } \langle \eta(\theta, \varphi), F(g(\theta, \varphi)) \rangle = 0 \quad \forall (\theta, \varphi) \in V_0.$$

En plus, $g(\theta, \varphi)$ est l'unique solution de l'équation (3.1) pour tout $t \in T_0$.

Démonstration :

Soit $G(\theta, \varphi, t)$ la fonction définie par:

$$G(\theta, \varphi, t) = \langle \eta(\theta, \varphi), F(t) \rangle,$$

G est une fonction C^p . D'après (3.1),

$$G(\theta_0, \varphi_0, t_0) = 0,$$

et d'après (3.2)

$$\partial_t G(\theta_0, \varphi_0, t_0) \neq 0.$$

Donc d'après le théorème de la fonction implicite:

- il existe V_0 voisinage ouvert de (θ_0, φ_0) , il existe T_0 voisinage ouvert de t_0 , et il existe $g : V_0 \longrightarrow T_0$ une fonction C^p tels que

$$t_0 = g(\theta_0, \varphi_0),$$

et

$$G(\theta, \varphi, g(\theta, \varphi)) = 0 \quad \forall (\theta, \varphi) \in V_0. \quad \blacksquare$$

Démonstration du théorème : Soient $a < t_1 < t_2 < \dots < t_q < b$ les paramètres des q intersections de $\Pi_{\eta(\theta_0, \varphi_0)}$ et de Γ . Puisque ces intersections sont simples, alors elles vérifient $G(\theta_0, \varphi_0, t_i) = 0$, et $\partial_t G(\theta_0, \varphi_0, t_i) \neq 0$ pour $i = 1, \dots, q$. D'après le lemme précédent, pour chaque t_i , $i=1, \dots, q$,

- il existe U_i voisinage ouvert de (θ_0, φ_0) , il existe T_i , voisinage ouvert de t_i , et il existe une fonction de classe C^p $g_i : U_i \longrightarrow T_i$ tels que

$$t_i = g_i(\theta_0, \varphi_0) \text{ et } G(\theta, \varphi, g_i(\theta, \varphi)) = 0 \quad \forall (\theta, \varphi) \in U_i.$$

Soit $U = \cap_{i=1}^q U_i$, voisinage ouvert de (θ_0, φ_0) . Puisque les t_i sont tous distincts on peut restreindre U de tel sorte que les $g_i(U)$ soient disjoints. On a, d'après la proposition 3.22, en remplaçant les t_i par $g_i(\theta, \varphi)$:

$$M(\theta, \varphi) = \varepsilon_0 \int_a^{g_1(\theta, \varphi)} F(t) dt + \sum_{i=1}^{q-1} \varepsilon_i \int_{g_i(\theta, \varphi)}^{g_{i+1}(\theta, \varphi)} F(t) dt + \varepsilon_q \int_{g_q(\theta, \varphi)}^b F(t) dt. \quad (3.3)$$

Donc M est une fonction de classe C^p de $U \cap V_{(\theta_0, \varphi_0)}$ sur $S(F)$.

Démontrons (ii). Rappelons que les dérivées partielles de g_i , $i = 1, \dots, q$, par rapport à θ et φ sont données par:

$$\begin{aligned} \partial_\theta g_i(\theta, \varphi) &= -\frac{\langle \partial_\theta \eta(\theta, \varphi), F(g_i(\theta, \varphi)) \rangle}{\langle \eta(\theta, \varphi), F'(g_i(\theta, \varphi)) \rangle}, \\ \partial_\varphi g_i(\theta, \varphi) &= -\frac{\langle \partial_\varphi \eta(\theta, \varphi), F(g_i(\theta, \varphi)) \rangle}{\langle \eta(\theta, \varphi), F'(g_i(\theta, \varphi)) \rangle}. \end{aligned} \quad (3.4)$$

En dérivant (3.3) par rapport à θ et φ et en utilisant les relations (3.4), on a: en dérivant M par rapport à θ ,

$$\partial_\theta M(\theta, \varphi) = -\sum_{i=1}^q (\varepsilon_{i-1} - \varepsilon_i) F(t_i) \frac{\langle \partial_\theta \eta(\theta, \varphi), F(t_i) \rangle}{\langle \eta(\theta, \varphi), F'(t_i) \rangle},$$

en dérivant M par rapport à φ ,

$$\partial_\varphi M(\theta, \varphi) = -\sum_{i=1}^q (\varepsilon_{i-1} - \varepsilon_i) F(t_i) \frac{\langle \partial_\varphi \eta(\theta, \varphi), F(t_i) \rangle}{\langle \eta(\theta, \varphi), F'(t_i) \rangle}.$$

Or

$$\frac{-(\varepsilon_{i-1} - \varepsilon_i)}{\langle \eta(\theta, \varphi), F'(t_i) \rangle} = \frac{1}{|\langle \eta(\theta, \varphi), F'(t_i) \rangle|},$$

donc:

$$\partial_\theta M(\theta, \varphi) = \sum_{i=1}^q F(g_i(\theta, \varphi)) \frac{\langle \partial_\theta \eta(\theta, \varphi), F(g_i(\theta, \varphi)) \rangle}{|\langle \eta(\theta, \varphi), F'(g_i(\theta, \varphi)) \rangle|}, \quad (3.5)$$

$$\partial_\varphi M(\theta, \varphi) = \sum_{i=1}^q F(g_i(\theta, \varphi)) \frac{\langle \partial_\varphi \eta(\theta, \varphi), F(g_i(\theta, \varphi)) \rangle}{|\langle \eta(\theta, \varphi), F'(g_i(\theta, \varphi)) \rangle|}. \quad (3.6)$$

Supposons que $F(g_i(\theta, \varphi))$, $i = 1, \dots, q$ soient colinéaires; alors pour tout i , il existe α_i tel que

$$F(g_i(\theta, \varphi)) = \alpha_i F(g_1(\theta, \varphi)). \quad (3.7)$$

Si on remplace ceci dans les équations (3.5) et (3.6), on aura $\partial_\theta M$ et $\partial_\varphi M$ linéairement dépendants.

Réciproquement, supposons que $\partial_\theta M(\theta, \varphi)$ et $\partial_\varphi M(\theta, \varphi)$, sont linéairement dépendants; il existe $(\alpha, \beta) \neq (0, 0)$ tels que

$$\begin{aligned} 0 &= \alpha \partial_\theta M(\theta, \varphi) + \beta \partial_\varphi M(\theta, \varphi) \\ &= \sum_{i=1}^q \langle v(\theta, \varphi), F(g_i(\theta, \varphi)) \rangle c_i F(g_i(\theta, \varphi)), \end{aligned}$$

où

$$c_i = \frac{1}{|\langle \eta(\theta, \varphi), F'(t_i) \rangle|}$$

et

$$v = \alpha \partial_\theta \eta + \beta \partial_\varphi \eta.$$

On a donc

$$0 = \sum c_i \langle v, F(t_i) \rangle F(t_i). \quad (3.8)$$

En faisant le produit scalaire des deux membres de (3.8) avec le vecteur v , nous obtenons:

$$0 = \langle 0, v \rangle = \sum c_i \langle v, F(t_i) \rangle^2.$$

Les constantes c_i sont toutes positives, donc le vecteur v est orthogonal à tous les $F(t_i)$, $i = 1, \dots, q$. Or $F(t_i)$ et v appartiennent au même plan $\Pi_{\eta(\theta, \varphi)}$, ce qui prouve que les $F(t_i)$ sont tous colinéaires. ■

Remarque 6: Dans le premier diagramme projectif le point (ii) du théorème s'exprime par:

(ii') - $\partial_\theta M(\theta, \varphi)$ et $\partial_\varphi M(\theta, \varphi)$ sont linéairements indépendants si et seulement si le nombre d'intersections distinctes entre la courbe Γ_p et la droite Δ_η est supérieur ou égal à 2.

3. 5. Etude d'un exemple

$$F(t) = \begin{pmatrix} t \\ t^3 - t \\ 1 \end{pmatrix} \quad \forall t \in [-1, 1].$$

Dans le premier diagramme projectif, la courbe Γ_p peut être paramétrée par:

$$(t, t^3 - t) \quad t \in [-1, 1].$$

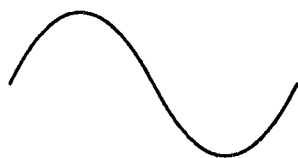


FIGURE .6. Courbe Γ_p .

On rappelle qu'à tout vecteur η , correspond une droite orientée Δ_η dans le diagramme projectif. Le point d'appui du plan Π_η sur la surface du zonoïde, est défini en intégrant la fonction F sur l'ensemble des paramètres t tels que $F_p(t)$ est du côté droit de Δ_η .

La régularité locale donnée par le théorème s'applique dans le cas d'intersections du type illustré dans la figure 7.

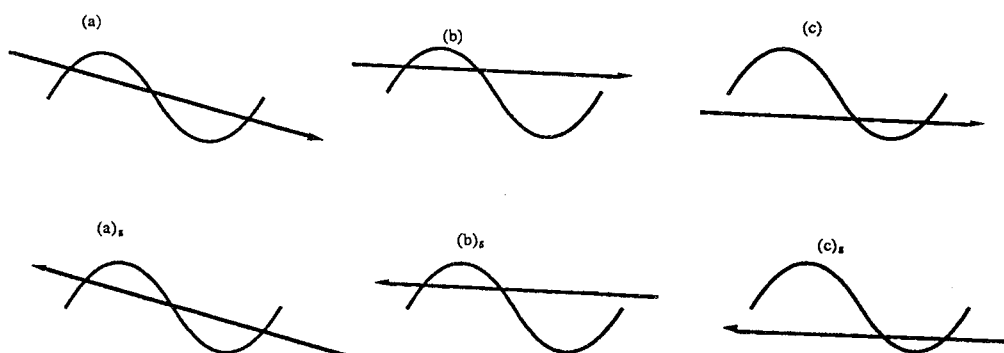


FIGURE .7. Cas réguliers.

Dans chacun des cas, la régularité locale peut être étendue sur l'ensemble des paramètres qui définissent le même type d'intersections. L'image de cet ensemble de paramètres est un morceau (ou patch) de la surface zonoïdale.

On dira qu'un morceau de la surface zonoïdale est un q -patch, s'il est l'image d'un ensemble de paramètres (θ, φ) qui définissent exactement q -intersections entre la droite orientée $\Delta_{\eta(\theta, \varphi)}$ et la courbe Γ_p .

Dans cet exemple, on a 3 types de patch qu'on a noté (a) , (b) et (c) et leurs symétriques (a_s) , (b_s) et (c_s) . Les patches correspondant à (b) et (c) sont tous deux définis par un même nombre d'intersections: 2, mais on ne peut pas trouver un déplacement continu de (b) à (c) sans passer par un point ne vérifiant pas les hypothèses de la régularité locale.

La surface du zonoïde est constituée de 6 patches et à chacun de ces patches correspond une expression différente de la surface:

- un point sur le patch du type (a) défini par trois intersections s'écrit:

$$M(\theta, \varphi) = \int_a^{g_1(\theta, \varphi)} F(t) dt + \int_{g_2(\theta, \varphi)}^{g_3(\theta, \varphi)} F(t) dt,$$

et sur le patch symétrique un point s'écrit:

$$M(\theta, \varphi) = \int_{g_1(\theta, \varphi)}^{g_2(\theta, \varphi)} F(t) dt + \int_{g_3(\theta, \varphi)}^b F(t) dt.$$

- Un point sur le patch du type (b) défini par deux intersections s'écrit:

$$M(\theta, \varphi) = \int_a^{g_1(\theta, \varphi)} F(t) dt + \int_{g_2(\theta, \varphi)}^b F(t) dt,$$

et sur le patch symétrique, il s'écrit:

$$M(\theta, \varphi) = \int_{g_1(\theta, \varphi)}^{g_2(\theta, \varphi)} F(t) dt.$$

- Un point sur le patch du type (c) défini par deux intersections s'écrit:

$$M(\theta, \varphi) = \int_{g_2(\theta, \varphi)}^{g_3(\theta, \varphi)} F(t) dt,$$

et sur le patch symétrique, il s'écrit:

$$M(\theta, \varphi) = \int_a^{g_2(\theta, \varphi)} F(t) dt + \int_{g_3(\theta, \varphi)}^b F(t) dt.$$

Les patches définis ci-dessus, sont séparés par des courbes formées de points qui correspondent à des paramètres ne vérifiant pas les hypothèses du théorème, ce qui se traduit dans le diagramme projectif par les cas suivants:

- (i)- la droite Δ_η passe par une des extrémités de la courbe Γ_p ,
- (ii)- Γ_p traverse la droite Δ_η en un seul point,
- (iii)- Δ_η est tangente à la courbe Γ_p .

Dans le cas (i), Δ_η passe par une extrémité. La courbe séparatrice est obtenue en faisant pivoter Δ_η autour de cette même extrémité. Le cas (ii), correspond à la courbe cumulée (voir courbe cumulée). Dans le cas (iii), Δ_η est tangente à Γ_p et puisque cette dernière est paramétrée par un polynôme de degré 3, la courbe ne peut traverser la droite qu'au maximum une seule fois; on se ramène donc au cas (ii). Dans la figure 8, on a représenté les situations dans le premier diagramme projectif qui correspondent aux courbes séparatrices.

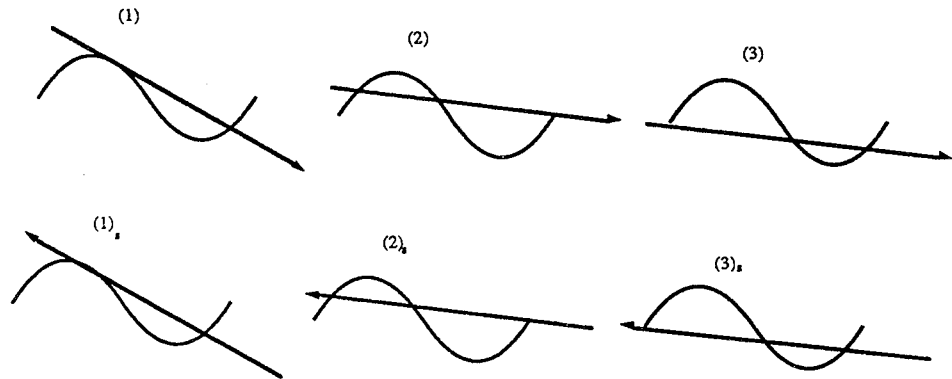


FIGURE .8. Cas non réguliers.

Le contact entre deux patchs le long la courbe cumulée est seulement C^0 ; en effet, comme on a vu dans la remarque 7, sur tout point de la courbe cumulée peuvent s'appuyer plusieurs plans d'appui au zonoïde.

Les points particuliers de la surface du zonoïde $Z(F)$ sont donnés par les configurations dans la figure 9.

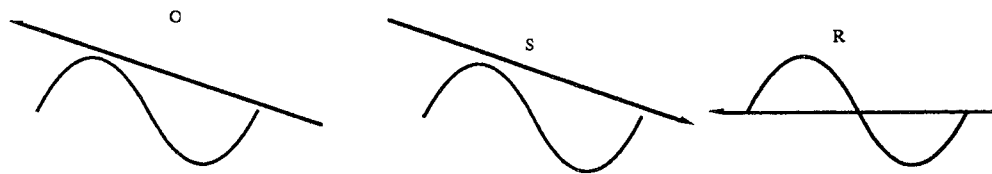


FIGURE .9. Points particuliers.

Les points 0 et S sont appelés pôles du zonoïde. Toute droite orientée Δ_η laissant à gauche (resp. à droite) la courbe dans le diagramme projectif correspond à un

plan qui s'appuie sur 0 (resp. S). S est appelé aussi point distal. Le point R dans ce cas correspond à un point qui appartient à la frontière des 3 patchs (a) , (b) et (c) . R est un point de la courbe cumulée.

Une lecture sur le diagramme projectif, nous permet d'établir un graphe planaire qui met en évidence l'adjacence des différents patchs sur la surface du zonoïde (figure 10). Le zonoïde est donné dans la figure 11.

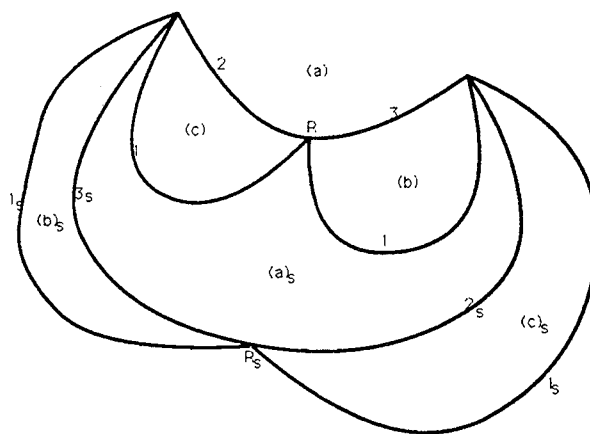


FIGURE .10. Graphe planaire.

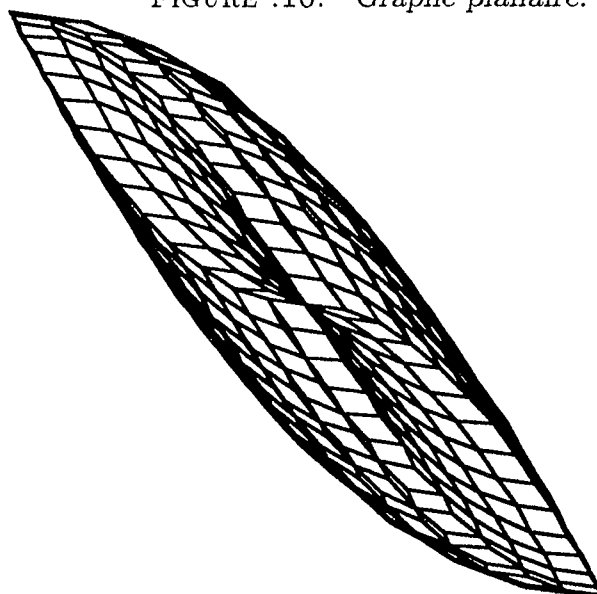


FIGURE .11. zonoïde associé.

4. Sélectivité totalement régulière (STR)

4. 1. Définition et propriétés

Définition 3. 26 On dira que la fonction F définie sur I à valeurs dans $\mathbb{R}^3$ a la propriété de la sélectivité totalement régulière (STR) si et seulement si pour tout $t_1, t_2, t_3 \in I$ tels que $t_1 < t_2 < t_3$, on a :

$$\det(F(t_1), F(t_2), F(t_3)) \neq 0.$$

En abrégé, on dira que F est STR.

Définition 3. 27 On appelle courbe convexe dans $\mathbb{R}^2$, une courbe contenue dans la frontière de sa propre enveloppe convexe.

Propriété 3. 28 Si F est une fonction STR alors le nombre d'intersections de la courbe Γ avec tout plan Π passant par 0 est inférieur ou égal à 2.

Cela permet d'énoncer la propriété suivante:

Propriété 3. 29 Si F est STR, alors son diagramme projectif est constitué par une courbe convexe.

Exemple :

1- Soit F la fonction définie par:

$$F(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ 1 \end{pmatrix} \quad \forall t \in [-\pi, \pi[.$$

F est une fonction STR.

2- Soit F la fonction définie par:

$$F(t) = \begin{pmatrix} t \\ t^3 - t \\ 1 \end{pmatrix} \quad \forall t \in [-1, 1].$$

F n'est pas STR. Le nombre d'intersections du plan passant par 0 de normale $(-1, 1, 0)$ est 3.

4. 2. Paramétrisation de la surface zonoïdale dans le cas STR

Proposition 3. 30 [Sza91] *Si F est une fonction STR alors $S(F)$ peut être exprimée comme réunion de deux surfaces paramétriques définies par:*

$$C^+ = \left\{ \int_u^v F(t)dt; \quad a \leq u \leq b, \quad u \leq v \leq b \right\},$$

et son symétrique

$$C^- = \left\{ \int_a^u F(t)dt + \int_v^b F(t)dt; \quad a \leq u \leq b, \quad u \leq v \leq b \right\}.$$

Remarquons ici que les deux surfaces C^+ et C^- sont définies sur le même domaine des paramètres qui est le triangle $\{(u, v) \mid a \leq u \leq b, \quad u \leq v \leq b\}$. L'intersection entre C^+ et C^- est la courbe cumulée.

Le résultat suivant, nous permet d'écrire $S(F)$ explicitement comme une seule surface paramétrique:

Théorème 3.31 *On suppose que F est STR sur $[a, b]$. Soit $\tilde{F}$ la fonction périodique de période $T = b - a$ définie de $\mathbb{R}$ à valeurs dans $\mathbb{R}^3$ telle que $\tilde{F}(t) = F(t) \forall t \in [a, b]$.*

Alors $S(F)$ est donnée par la surface paramétrique suivante:

$$\left\{ \int_u^v \tilde{F}(t)dt \quad a \leq u \leq b, \quad u \leq v \leq u + T \right\}. \quad (3.9)$$

Démonstration :

Notons C la surface paramétrique définie par (3.9). Alors on peut écrire:

$$C = \left\{ \int_u^v F(t)dt, \quad a \leq u \leq b, \quad u \leq v \leq b \right\} \cup \left\{ \int_u^v \tilde{F}(t)dt, \quad a \leq u \leq b, \quad b \leq v \leq u + T \right\},$$

donc:

$$C = C^+ \cup \left\{ \int_u^v \tilde{F}(t)dt, \quad a \leq u \leq b, \quad b \leq v \leq u + T \right\}.$$

Il reste à montrer que:

$$C^- = \left\{ \int_u^v \tilde{F}(t)dt, \quad a \leq u \leq b, \quad b \leq v \leq u + T \right\}.$$

On a:

$$\begin{aligned}
C^- &= \left\{ \int_a^u F(t)dt + \int_v^b F(t)dt, \quad a \leq u \leq b, \quad u \leq v \leq b \right\} \\
&= \left\{ \int_a^u F(t)dt + \int_v^b F(t)dt, \quad a \leq v \leq b, \quad a \leq u \leq v \right\} \\
&= \left\{ \int_b^{u+T} \tilde{F}(t)dt + \int_v^b F(t)dt, \quad a \leq v \leq b, \quad a \leq u \leq v \right\} \\
&= \left\{ \int_v^b F(t)dt + \int_b^{u'} \tilde{F}(t)dt, \quad a \leq v \leq b, \quad b \leq u' \leq v+T \right\} \\
&= \left\{ \int_v^{u'} \tilde{F}(t)dt \quad a \leq v \leq b, \quad b \leq u' \leq u+T \right\} \\
&= \left\{ \int_u^v \tilde{F}(t)dt \quad a \leq u \leq b, \quad b \leq v \leq u+T \right\}.
\end{aligned}$$

d'où

$$C = C^+ \cup C^-. \quad \blacksquare$$

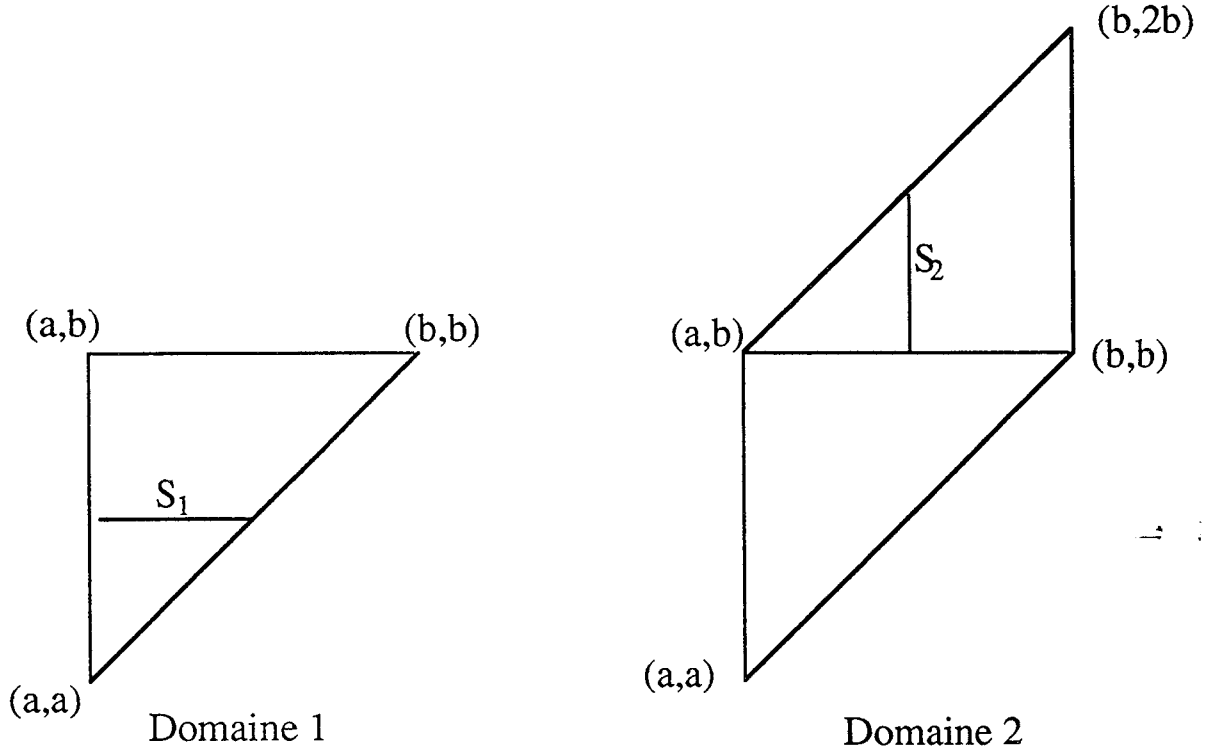


FIGURE .12. Domaines de paramètres

Dans la figure 12, pour la proposition 3.30, les surfaces C^+ et C^- sont définies sur le même domaine 1; pour le théorème 3. 31, C^+ (resp. C^-) est définie sur le triangle (a, a) , (a, b) et (b, b) , (resp. (a, b) , (b, b) et $(b, 2b)$). La surface $S(F)$ est définie sur le domaine 2. Les segments S_1 et S_2 ont la même courbe image sur C^- .

4. 3. Régularité de la paramétrisation

La régularité de cette paramétrisation est donnée par le théorème suivant:

Théorème 3.32 *Soit F une fonction STR de $[a, b]$ à valeurs dans $\mathbb{R}^3$. Soit $\tilde{F}$ la fonction périodique de période $T = b - a$ telle que $\tilde{F}(t) = F(t) \forall t \in [a, b]$. Considérons la paramétrisation définie par:*

$$\tilde{M}(u, v) \longrightarrow \int_u^v \tilde{F}(t) dt, \quad \begin{array}{l} a \leq u \leq b, \\ u \leq v \leq u + T. \end{array}$$

Alors si $\tilde{F}$ est de classe C^p au voisinage de u et au voisinage de v , $\tilde{M}$ est de classe C^{p+1} au voisinage de (u, v) .

De plus:

$\partial_u \tilde{M}(u, v)$ et $\partial_v \tilde{M}(u, v)$ sont linéairement indépendants si et seulement si $v \neq u$ et $v \neq u + T$.

Démonstration : Pour tout $u \in [a, b[$ et $v \in [u, u + T]$, on a:

$$\tilde{M}(u, v) = \int_u^v \tilde{F}(t) dt,$$

donc:

$$\partial_u \tilde{M}(u, v) = -\tilde{F}(u), \quad (3.10)$$

$$\partial_v \tilde{M}(u, v) = \tilde{F}(v). \quad (3.11)$$

Donc si $\tilde{F}$ est de classe C^p alors $\tilde{M}$ est de classe C^{p+1} .

F étant STR, $\forall u, v \in [a, b[, u \neq v$, $F(u)$ et $F(v)$ sont linéairement indépendants, ce qui entraîne que $\tilde{F}(u)$ et $\tilde{F}(v)$ le sont aussi $\forall u \in [a, b[, \forall v \in]u, u + T[$. ■

4. 4. Exemples

• Exemple 2

$$F(t) = \begin{pmatrix} t^2 \\ t \\ 1 \end{pmatrix} \quad \forall t \in [0, 1].$$

La fonction F est STR sur $[0, 1]$.

Le diagramme projectif est donné par la courbe de la figure 13.



FIGURE 13. Courbe Γ_p .

Les patchs réguliers correspondent aux types d'intersection donnés dans la figure 14.

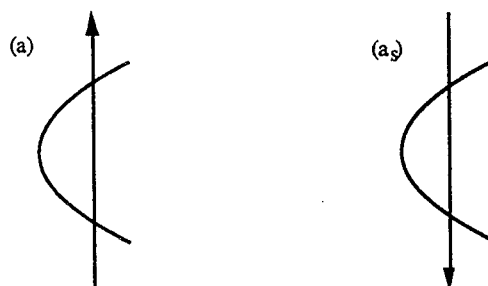


FIGURE .14. Cas réguliers.

Les points irréguliers de $S(F)$ correspondent aux cas où, dans le premier diagramme projectif, la courbe ne traverse la droite qu'une seule fois (figure 15). Ils sont donc sur la courbe cumulée.

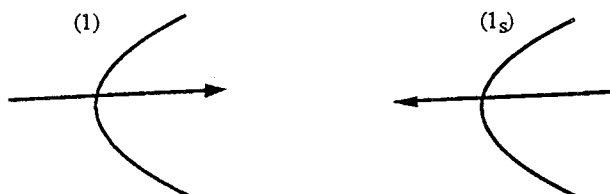


FIGURE .15. Cas non réguliers.

En considérant la paramétrisation définie par $\tilde{M}(u, v)$, la surface du zonoïde est donc constituée de deux patchs réguliers C^∞ qui se raccordent C^0 le long de la courbe cumulée. Le raccordement C^0 est dû au fait que la fonction $\tilde{F}$, définie dans la théorie 3. 31, est discontinue au point 1.

Le graphe planaire est donné dans ce cas par la figure 16.

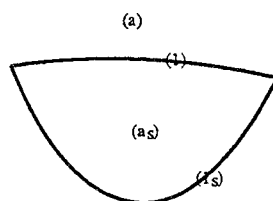


FIGURE .16. Graphe planaire.

La fonction F étant STR, d'après la proposition proposition 3.30, elle peut s'écrire comme réunion de deux surfaces paramétriques C_1 et C_2 qui sont définies par:

$$\left\{ \begin{pmatrix} \frac{1}{3}(v^3 - u^3) \\ \frac{1}{2}(v^2 - u^2) \\ v - u \end{pmatrix} ; \quad \forall u \in [0, 1] \quad u \leq v \leq b \right\},$$

et

$$\left\{ S - \begin{pmatrix} \frac{1}{3}(v^3 - u^3) \\ \frac{1}{2}(v^2 - u^2) \\ v - u \end{pmatrix} ; \quad \forall u \in [0, 1] \quad u \leq v \leq b \right\}.$$

Le zonoïde associé est donné par la figure 17.

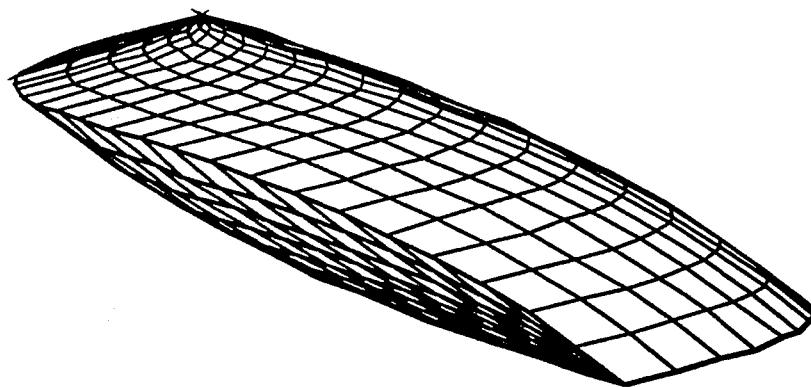


FIGURE .17. Zonoïde associé.

• Exemple 3 (Fonction STR périodique):

$$F(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ 1 \end{pmatrix} \quad \forall t \in [-\pi, \pi[.$$

F est STR sur $[-\pi, \pi[$.

Le diagramme projectif est donnée par la figure 18.

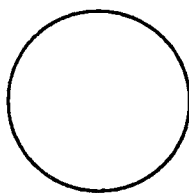


FIGURE .18. Courbe Γ_p .

La situation qui correspond aux points réguliers est donnée par toute droite intersectant Γ_p deux fois (figure 19).

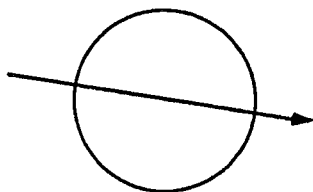


FIGURE .19. Cas réguliers.

Les points irréguliers de la surface de $Z(F)$, correspondent aux cas où la droite est tangente à la courbe (figure 20). Dans ce cas la courbe Γ_p est entièrement d'un côté de la droite, ce qui correspond au point O ou au point distal S . On remarque dans cet exemple que les points de la courbe cumulée constituent des points réguliers. Donc, la surface du zonoïde est constituée par un seul patch dont les points irréguliers sont O et S .

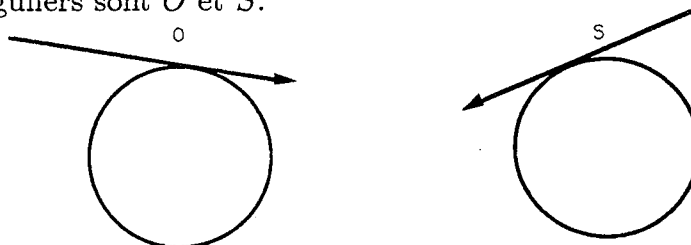


FIGURE .20. Cas non réguliers.

La fonction F est STR et périodique. Donc, d'après la théorème 3. 31, $S(F)$ peut s'écrire aussi sous la forme paramétrique suivante:

$$S(F) = \left\{ \int_u^v F(t)dt; \quad u \in [-\pi, \pi[, \quad v \in [u, u + 2\pi[\right\}.$$

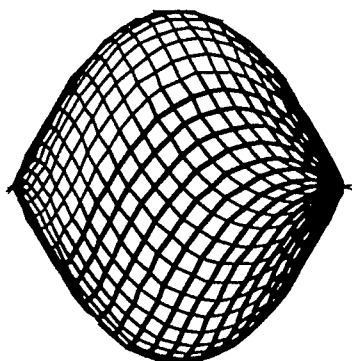


FIGURE .21. Zonoïde associé.

5. Coupe zonoïde-plan

Il est en général difficile, voire même impossible, de trouver une relation explicite de l'intersection d'un zonoïde avec un plan.

Dans cette partie, nous présentons une méthode numérique de suivi de coupe dans le cas d'un zonoïde associé à une fonction présentant la propriété de la sélectivité régulière.

5. 1. Problème d'intersection d'un plan avec un zonoïde

Soit H le plan défini par:

$$\{ x \in \mathbb{R}^3 \mid \langle x, \eta \rangle = \alpha \}.$$

Soit $F : I \longrightarrow \mathbb{R}^n$ une fonction STR. On cherche à déterminer l'intersection de H avec le zonoïde $Z(F)$.

Notons:

$$\alpha_{max} = H_Z(\eta) = \max_{x \in Z} \langle x, \eta \rangle,$$

$$\alpha_{min} = \min_{x \in Z} \langle x, \eta \rangle.$$

Le plan H coupe Z si et seulement si

$$\alpha_{min} \leq \alpha \leq \alpha_{max}.$$

On désigne par D le domaine triangulaire:

$$D = \{ (x, y) \mid x \in [a, b], \quad x \leq y \leq b \}$$

Notons C^+ la surface définie par la paramétrisation suivante:

$$\begin{aligned} D &\longrightarrow \mathbb{R}^3 \\ (x, y) &\longrightarrow \int_x^y F(t) dt \end{aligned}$$

et C^- celle définie par:

$$\begin{aligned} D &\longrightarrow \mathbb{R}^3 \\ (x, y) &\longrightarrow S - \int_x^y F(t) dt \end{aligned}$$

D'après la proposition 3.30, la surface $S(F)$ est alors donnée par:

$$S(F) = C^+ \cup C^-.$$

On cherchera l'intersection de H avec C^+ et C^- .

L'ensemble des points d'intersection de H avec C^+ s'écrit:

$$\left\{ \int_x^y F(t)dt \mid \int_x^y \langle F(t), \eta \rangle dt = \alpha \right\},$$

et l'ensemble des points d'intersection de H avec C^- s'écrit:

$$\left\{ S - \int_x^y F(t)dt \mid \langle S, \eta \rangle - \int_x^y \langle F(t), \eta \rangle dt = \alpha \right\}.$$

Un point de C^+ (resp. C^-) a pour paramètre (x, y) , s'il s'écrit sous la forme $\int_x^y F(t)dt$ (resp. $S - \int_x^y F(t)dt$).

Considérons la fonction

$$\Phi_\eta(x, y) = \int_x^y \langle F(t), \eta \rangle dt,$$

définie sur le domaine D .

H coupe C^+ aux points de paramètres (x, y) vérifiant:

$$\Phi_\eta(x, y) = \alpha.$$

De même H coupe C^- aux points de paramètres (x, y) vérifiant:

$$\Phi_\eta(x, y) = \langle S, \eta \rangle - \alpha.$$

Ainsi la recherche de la valeur des paramètres correspondant aux points de la coupe de $S(F)$ par H devient un problème de détermination des paramètres des lignes des niveaux α et $\langle S, \eta \rangle - \alpha$ de la fonction Φ_η .

Remarque 7: Les surfaces C^+ et C^- s'intersectent en une courbe fermée. Cette courbe est constituée par les deux courbes définies par:

$$\left\{ \int_a^x F(t)dt, \quad x \in [a, b] \right\},$$

et sa symétrique:

$$\left\{ \int_x^b F(t)dt, \quad x \in [a, b] \right\}.$$

C'est la courbe cumulée qu'on appelle aussi courbe séparatrice.

Lorsqu'un point M est sur la courbe séparatrice, il appartient à C^+ et à C^- et donc il peut s'écrire sous deux formes selon qu'il est considéré comme un point de C^+ ou de C^-

$$\begin{aligned} M &= \int_a^x F(t)dt && \in C^+ \\ &= S - \int_x^b F(t)dt && \in C^-, \end{aligned}$$

ou

$$\begin{aligned} M &= \int_x^b F(t)dt && \in C^+ \\ &= S - \int_a^x F(t) && \in C^-. \end{aligned}$$

Les paramètres de la forme (a, u) et (u, b) correspondent donc aux paramètres d'un même point selon qu'il est considéré dans C^+ ou C^- .

5. 2. Méthodes de résolution:

Le problème est de déterminer (x, y) tel que :

$$\Phi_\eta(x, y) = \alpha. \quad (3.12)$$

L'ensemble des solutions $\begin{pmatrix} x \\ y \end{pmatrix}$ de (3.12) est une courbe dans le plan qu'on peut représenter avec les coordonnées paramétriques:

$$x = x(s), \quad y = y(s)$$

En remplaçant dans (3.12) et en dérivant par rapport à s , on obtient:

$$\frac{\partial \Phi_\eta}{\partial x}(x(s), y(s))x'(s) + \frac{\partial \Phi_\eta}{\partial y}(x(s), y(s))y'(s) = 0. \quad (3.13)$$

Une solution particulière de (3.13) est donnée par:

$$\begin{aligned} x'(s) &= \frac{\partial \Phi_\eta}{\partial y}(x(s), y(s)), \\ y'(s) &= -\frac{\partial \Phi_\eta}{\partial x}(x(s), y(s)), \end{aligned} \quad (3.14)$$

qui s'écrit encore en remplaçant Φ_η par sa valeur:

$$\begin{aligned} x'(s) &= \langle F(y(s)), \eta \rangle, \\ y'(s) &= \langle F(x(s)), \eta \rangle. \end{aligned} \quad (3.15)$$

L'ensemble des points solutions de (3.12) est donc solution du système d'équations différentielles (3.15).

Connaissant une solution de départ $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$, on peut utiliser une méthode de résolution numérique de système d'équations différentielles ordinaires.

5.2.1. Algorithme de résolution numérique de (3.15)

Pour résoudre numériquement (3.15), supposons que les fonctions Ψ_1 et Ψ_2 représentent un schéma numérique de résolution; alors:

$$\begin{cases} x_{n+1} = x_n + h\Psi_1(x_n, y_n, h), \\ y_{n+1} = y_n + h\Psi_2(x_n, y_n, h), \end{cases} \quad (3.16)$$

où h est le pas.

Test d'arrêt:

Pour établir un test d'arrêt, nous devons prévoir deux cas:

- **premier cas:** la courbe solution est une courbe fermée dans le domaine D des paramètres. Dans ce cas l'itération doit s'arrêter pour ne pas tourner une deuxième fois sur la courbe.

- **deuxième cas:** A une étape de l'itération (3.16), le point courant $\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix}$ est à l'extérieur du domaine (i.e. $x_{k+1} < a$ ou $y_{k+1} > b$). Dans ce cas l'itération est arrêtée, et le dernier point de l'itération que nous allons prendre sera l'intersection du segment $\left[\begin{pmatrix} x_k \\ y_k \end{pmatrix}, \begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} \right]$ avec:

$$\begin{aligned} * & \left[\begin{pmatrix} a \\ a \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \right], \text{ si } x_{k+1} < a, \\ * & \left[\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} b \\ b \end{pmatrix} \right], \text{ si } y_{k+1} > b. \end{aligned}$$

Le dernier point de l'itération est un point sur la frontière du domaine, il correspond aux paramètres d'un point de la courbe séparatrice.

5.2.2. Algorithme de suivi de coupe

On suppose que $\alpha_{min} \leq \alpha \leq \alpha_{max}$, c'est à dire que H coupe C^+ .

- **premier cas**

On suppose que $\Phi_\eta(a, b) - \alpha > \alpha_{max}$ ou $\Phi_\eta(a, b) - \alpha < \alpha_{min}$. Ainsi H ne coupe pas C^- . Dans ce cas, l'intersection de H avec C^+ est une courbe fermée. $\curvearrowright$

Soit $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ un point initial vérifiant $\Phi_\eta(x_0, y_0) = \alpha$; en utilisant le schéma (3.16) pour résoudre (3.15), on obtient une suite de couples de paramètres:

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \dots, \begin{pmatrix} x_p \\ y_p \end{pmatrix}.$$

Les points correspondants:

$$\left(\int_{x_i}^{y_i} F(t) dt \right)_{i=0, \dots, p}$$

représentent une discrétisation ordonnée de la coupe de C^+ par H .

• deuxième cas

$$\alpha_{min} \leq \Phi_{\eta}(a, b) - \alpha \leq \alpha_{max},$$

H coupe les surfaces C^+ et C^- .

Pour simplifier, supposons que le nombre de points d'intersection de la courbe séparatrice et du plan H est égale à 2. Dans ce cas $H \cap C^+$ (de même $H \cap C^-$) est constituée par un seul morceau de courbe.

Soit $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ un point de la courbe séparatrice (i.e., soit $x_0 = a$, soit $y_0 = b$) vérifiant $\Phi_{\eta}(x_0, y_0) = \alpha$. En résolvant numériquement par (3.16) le système (3.15), on fait un suivi d'intersection de H avec C^+ en déterminant des paramètres $\begin{pmatrix} x_i \\ y_i \end{pmatrix}$ des points de la coupe. Le dernier couple $\begin{pmatrix} x_p \\ y_p \end{pmatrix}$ trouvé contient les paramètres d'un point sur la courbe séparatrice. D'après la remarque précédente, $\begin{pmatrix} x'_p \\ y'_p \end{pmatrix}$ définis par:

$$\begin{cases} \begin{cases} x'_p = a \\ y'_p = x_p \end{cases} & \text{si } y_p = b \\ \begin{cases} x'_p = y_p \\ y'_p = b \end{cases} & \text{si } x_p = a \end{cases} \quad (3.17)$$

représentent les paramètres d'un même point sur la surface du zonoïde que celui défini par $\begin{pmatrix} x_p \\ y_p \end{pmatrix}$. On passe alors des paramètres $\begin{pmatrix} x_p \\ y_p \end{pmatrix}$ de C^+ aux paramètres $\begin{pmatrix} x'_p \\ y'_p \end{pmatrix}$ de C^- , et puis on continue le suivi de coupe dans C^- . Le dernier couple de paramètres $\begin{pmatrix} x_l \\ y_l \end{pmatrix}$ trouvé contient les paramètres d'un point sur la courbe séparatrice.

Puisque le plan H ne coupe la courbe séparatrice que deux fois alors $\begin{pmatrix} x_l \\ y_l \end{pmatrix}$ et $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ représentent les paramètres d'un même point.

On obtient ainsi une suite de couples de paramètres $\begin{pmatrix} x_i \\ y_i \end{pmatrix}_{i=1, \dots, l}$ tels que les points

$$\int_{x_i}^{y_i} F(t) dt$$

représentent une discrétisation ordonnée de la coupe.

Remarque 8: Lorsque le nombre de points d'intersection de la courbe séparatrice et de H est supérieur à 2, on suit la même démarche que précédemment et le suivi

d'intersection se déroulera en passant plusieurs fois de C^+ à C^- et inversement jusqu'à ce qu'on revienne au point de départ.

5.2.3. Exemple

Pour

$$F(t) = \begin{pmatrix} e^{-0.2t} \\ e^{-t} \\ e^{-5t} \end{pmatrix}; \quad \forall t \in [0, 1].$$

On prend $\eta = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ et H le plan d'équation $\langle x, \eta \rangle = 0.05$.

$$\alpha_{min} = 0, \quad \alpha_{max} = 0.2.$$

Les paramètres des points de la coupe de C^+ (resp C^-) avec H correspondent aux paramètres du niveau $\alpha = 0.05$ (resp $\alpha = 0.15$).

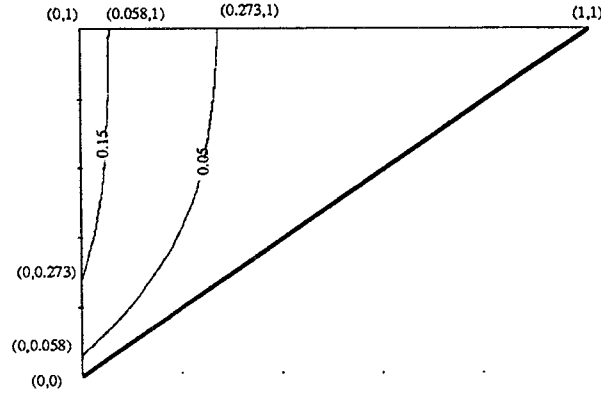


FIGURE .22. Courbes de paramètres des points d'intersection de H avec C^+ et C^- .

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0.058 \end{pmatrix}, \quad \begin{pmatrix} x_p \\ y_p \end{pmatrix} = \begin{pmatrix} 0.273 \\ 1 \end{pmatrix},$$

et

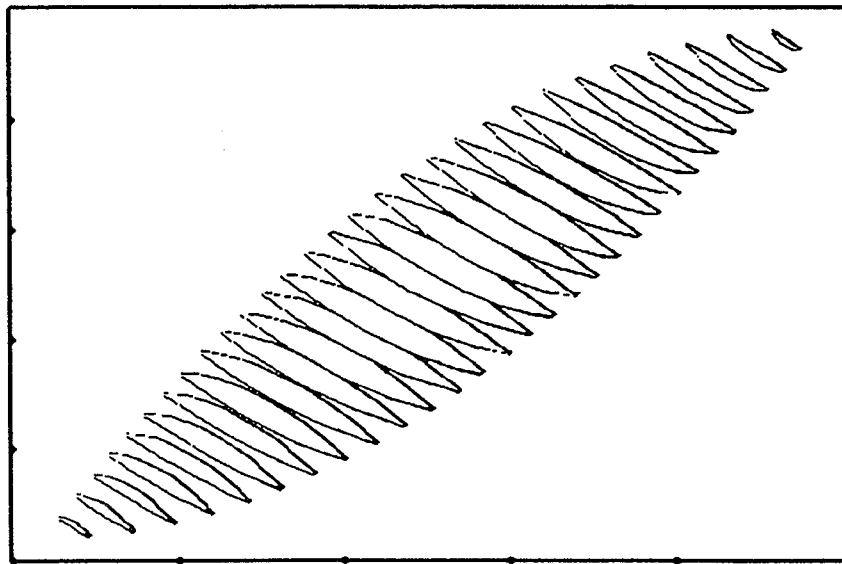
$$\begin{pmatrix} x_{p'} \\ y_{p'} \end{pmatrix} = \begin{pmatrix} 0 \\ 0.273 \end{pmatrix}, \quad \begin{pmatrix} x_l \\ y_l \end{pmatrix} = \begin{pmatrix} 0.058 \\ 1 \end{pmatrix}.$$

Remarque 9: Le fait que F admette la propriété de la sélectivité régulière est nécessaire pour appliquer cette méthode. Lorsque cette propriété n'est pas vérifiée, l'approximation du zonoïde par un zonotope (cf chapitre suivant), permet de ramener le problème à un problème d'intersection d'un zonotope avec un plan.

5. 3. Quelques exemples de visualisation par des coupes parallèles

$$F(t) = \begin{pmatrix} t^2 \\ t \\ 1 \end{pmatrix}; \forall t \in [0, 1].$$

$$\eta = (0, 0, 1)$$



$$\eta = (1, 1, 1)$$

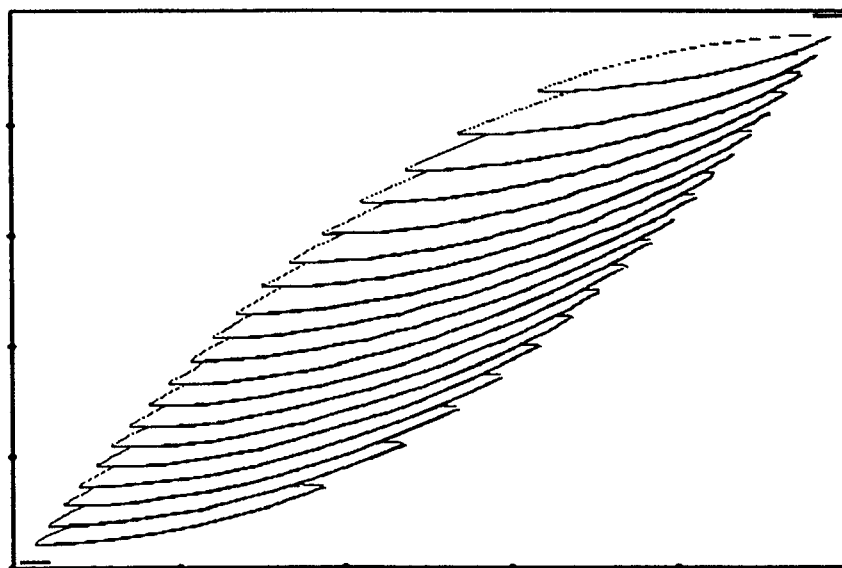
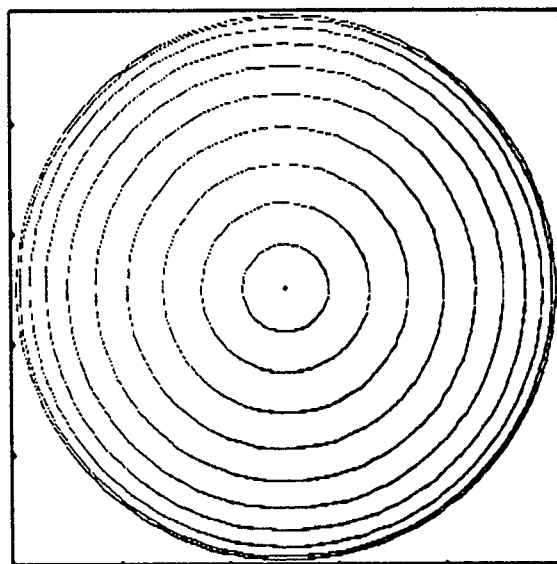


FIGURE .23. Visualisation par des coupes parallèles.

$$F(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ 1 \end{pmatrix}; \forall t \in [-\Pi, \Pi].$$

$$\eta=(0,0,1)$$



$$\eta=(1,1,1)$$

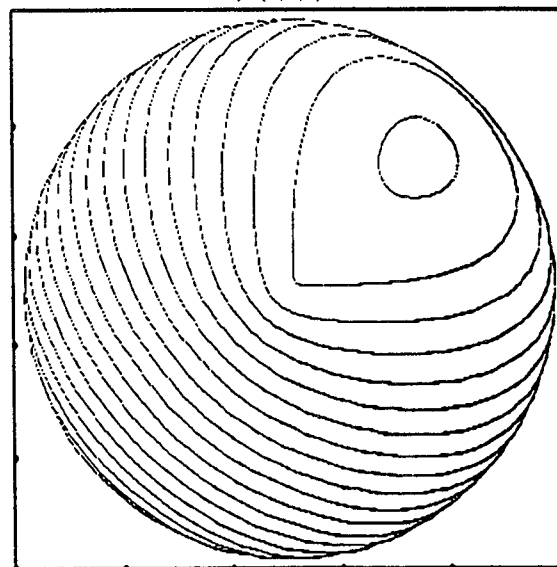


FIGURE .24. Visualisation par des coupes parallèles.

Chapitre 4

Approximation de zonoïdes par des zonotopes

Introduction

Dans ce chapitre, nous traitons le problème de l'approximation d'un zonoïde par un zonotope. D'après la définition, un zonoïde est la limite, au sens de la métrique de Hausdorff, d'une séquence de zonotopes. Donc, étant donné un zonoïde, cette définition assure l'existence d'un zonotope qui l'approche.

Dans ce cadre, deux problèmes se posent: pour un zonoïde Z donné

- définir une séquence de zonotopes qui converge vers le zonoïde Z ,
- étant donné qu'un zonotope est défini par la somme de Minkowski de segments, quel est le nombre de segments nécessaires pour approcher Z à ε près?

Dans la première partie nous donnerons les principaux résultats donnés par la théorie générale des zonoïdes.

Dans la seconde partie, nous étudierons le cas d'un zonoïde $Z(F)$ associé à une fonction F définie sur un intervalle de $\mathbb{R}$ à valeurs dans $\mathbb{R}^n$. Nous donnerons une méthode de construction de séquences de zonotopes qui convergent vers $Z(F)$, et nous démontrerons, sous certaines hypothèses, que le nombre de segments suffisants pour créer un zonotope qui approche $Z(F)$ à ε près est en $O(\frac{1}{\sqrt{\varepsilon}})$. Ce nombre est meilleur que celui donné dans la théorie générale celui ci est de l'ordre de $O(\frac{1}{\varepsilon^2})$.

La troisième partie sera consacrée à quelques exemples d'applications.

Dans la quatrième partie, nous étendrons l'étude aux cas où la fonction est définie sur un domaine de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}^n$. Nous construirons une séquence de zonotopes qui converge vers $Z(F)$. Le nombre de segments suffisants pour approcher $Z(F)$ est ici en $O(\frac{1}{\varepsilon})$.

Dans ce chapitre, on reprendra les mêmes notations et définitions que dans le chapitre précédent. $\mathcal{K}$ est l'ensemble des convexes compacts de $\mathbb{R}^n$.

1. Rappels de quelques résultats sur l'approximation

1. 1. Notions de convergence

Définition 4. 1 Soient $(K_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\mathcal{K}$ et K un élément de $\mathcal{K}$. On dit que la suite $(K_n)_{n \in \mathbb{N}}$ converge scalairement vers K si et seulement si

$$\forall u \in \mathbb{R}^n \quad H_{K_n}(u) \longrightarrow H_K(u) \quad \text{quand } n \rightarrow +\infty$$

Définition 4. 2 Soient $(K_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\mathcal{K}$ et K un élément de $\mathcal{K}$. On dit que la suite $(K_n)_{n \in \mathbb{N}}$ converge au sens de Hausdorff vers K si et seulement si

$$\text{haus}(K_n, K) \longrightarrow 0 \quad \text{quand } n \rightarrow +\infty$$

Dans le cas de la dimension finie, nous avons la proposition suivante:

Proposition 4. 3 Soient $(K_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\mathcal{K}$ et K un élément de $\mathcal{K}$, alors les propositions suivantes sont équivalentes:

- i- $\lim_{n \rightarrow +\infty} \text{haus}(K_n, K) = 0$,
- ii- K_n converge scalairement vers K quand $n \rightarrow +\infty$.

Nous avons aussi le résultat suivant qui lie la distance de Hausdorff entre deux convexes compacts et leurs fonctions support.

Lemme 4. 4 (Ghouila-houri) Soient A et B deux parties convexes compactes non vides dans $\mathbb{R}^n$ alors:

$$\text{haus}(A, B) = \sup_{u \in S} |H_A(u) - H_B(u)|.$$

(pour une démonstration, voir Hörmander [Hör54]).

1. 2. Approximation d'un ensemble convexe compact par un polytope

Tout ensemble convexe compact peut être approché par un polytope. Les principaux résultats sont donnés par les deux théorèmes suivants (voir Valentine [Vale 64]):

Théorème 4.5 *Si K est un ensemble convexe compact dans $\mathbb{R}^n$, alors pour tout $\varepsilon > 0$, il existe un polytope convexe P tel que*

$$P \subset K \subset P_\varepsilon,$$

où $P_\varepsilon = P + \varepsilon B$.

De ce théorème découle le corollaire suivant:

Corollaire 4.6 *Si K un élément de $\mathcal{K}$, alors il existe une séquence de polytopes P_i qui converge vers K , au sens de la métrique de Hausdorff.*

Théorème 4.7 *Soit K un élément de $\mathcal{K}$ tel que $0 \in \text{int}(K)$; alors pour tout $\delta > 1$, il existe un polytope convexe P tel que*

$$P \subset K \subset \delta P,$$

où $\delta P = \{\delta x; \quad x \in P\}$.

1. 3. Approximation de zonoïdes par la somme de Minkowski de segments

Dans le cas de l'approximation de la boule unité B , le problème du nombre $N(B, \varepsilon)$ de segments pour définir un zonotope qui approche B à ε près, a fait l'objet de diverses études. Le résultat de Gordon dans [Gor85] dans ce sens est donné par le théorème suivant:

Théorème 4.8 [Gor85] *Soit B la boule unité de $\mathbb{R}^n$, alors le nombre, $N = N(B, \varepsilon)$, minimal nécessaire pour générer un zonotope Z_N tel que l'on ait:*

$$\text{haus}(B, Z_N) < \varepsilon,$$

vérifie:

$$N(B, \varepsilon) \leq c\varepsilon^{-2}n,$$

où c est une constante.

Cependant, ce résultat n'est pas optimal pour $n = 2$: il suffit de prendre $N(B, \varepsilon) = \varepsilon^{-1/2}$.

Sur l'approximation de zonoïdes par des zonotopes, nous citons les deux résultats suivants:

Linquist dans [Lin75] donne le résultat suivant:

Théorème 4.9 [Lin75]: Soit Z un zonoïde défini par la mesure positive ρ sur $\mathcal{S}$ (i.e., $H_Z(u) = \int_{\mathcal{S}} |\langle x, u \rangle| \rho(dx)$, $\forall u \in E$). Soit $\varepsilon > 0$ donné et $\delta = \varepsilon/\rho(\mathcal{S})$. Soit $(\omega_1, \omega_2, \dots, \omega_r)$ une partition de $\mathcal{S}$ telle que $\max_{i=1}^r (\max_{v, w \in \omega_i} \|v - w\|) \leq \delta$.

Alors le zonotope $Z_r = \{\sum_{i=1}^r \lambda_i y_i \rho(\omega_i), \forall \lambda_i \in [-1, 1]\}$ avec $y_i \in \omega_i$ approche Z à ε près, au sens de la métrique de Hausdorff.

Le second résultat est donné par Bourgain, Lindenstrauss et Milman. Dans l'article [BLM89], ils donnent différents résultats sur l'approximation de zonoïdes par des zonotopes. Nous avons retenu le théorème suivant:

Théorème 4.10 [BLM89] Pour tout $\tau > 0$, il existe une constante $c(\tau)$ telle que pour tout zonoïde Z dans $\mathbb{R}^n$, $n \geq 2$

$$N(Z, \varepsilon) \leq c(\tau) \varepsilon^{-(2+\tau)} (\log(n))^3 n,$$

où $N(Z, \varepsilon)$ désigne le nombre minimal de segments nécessaire pour approcher Z à ε près.

Notons ici que les résultats cités dans cette partie sont basés sur des approches purement théoriques et qu'aucun d'eux ne fournit une méthode de construction du zonotope qui approche le zonoïde donné.

2. Approximation d'un zonoïde associé à une fonction

Rappelons qu'on désigne par $\mathcal{S}$ la sphère unité dans $\mathbb{R}^n$, $I = [a, b]$ et par Λ l'ensemble des fonctions mesurables définies sur I à valeurs dans $[0, 1]$. Soient $F : I \rightarrow \mathbb{R}^n$ une fonction intégrable et Γ la courbe dans $\mathbb{R}^n$ correspondant à F .

Considérons le zonoïde $Z(F)$ associé à F . Il s'écrit:

$$Z(F) = \left\{ \int_I \lambda(t) F(t) dt \mid \lambda \in \Lambda \right\}.$$

Pour définir une séquence de zonotopes qui approche $Z(F)$, et en s'inspirant du théorème 4. 9, Szafran dans [Sza91] a proposé la séquence de zonotopes donnée par la proposition suivante:

Proposition 4. 11 *Soit F une fonction intégrable de I dans $\mathbb{R}^n$, alors le zonoïde $Z(F)$ est la limite d'une suite de zonotopes $Z_p(F)$ définis comme somme des segments $[O, N_i]$ avec:*

$$N_i = \int_{s_i}^{s_{i+1}} F(t) dt \quad \text{pour} \quad a = s_0 < s_1 < \dots < s_p = b,$$

avec $\lim_{p \rightarrow \infty} \max_{i=1}^p (s_i - s_{i-1}) = 0$.

Si cette proposition nous donne une méthode pour construire une suite de zonotopes qui converge vers $Z(F)$, elle ne permet pas à p donné d'estimer l'erreur d'approximation du zonoïde $Z(F)$ par le zonotope $Z_p(F)$. Dans la partie suivante, nous donnerons, par une approche différente, d'autres méthodes pour définir des suites de zonotopes qui convergent vers $Z(F)$; et dont on connaît l'erreur d'approximation et l'ordre de convergence.

2. 1. Construction du zonotope approximant

Pour $u \in \mathcal{S}$, d'après la proposition 3.13, la fonction support de $Z(F)$ s'écrit:

$$H_{Z(F)}(u) = \int_a^b \langle F(t), u \rangle_+ dt. \quad (4.1)$$

En utilisant une formule de quadrature, on peut écrire pour $u \in \mathcal{S}$,

$$H_{Z(F)}(u) = I_p(u, F) + R_p(u, F), \quad (4.2)$$

où

$$I_p(u, F) = \sum_{i=0}^p c_i \langle F(t_i), u \rangle_+, \quad (4.3)$$

et $R_p(u, F)$ est le reste de la formule de quadrature $I_p(u, F)$, avec

$$a \leq t_0 < t_1 < \dots < t_p \leq b,$$

les noeuds d'intégration et c_i des constantes positives.

Considérons l'ensemble Z_p défini par:

$$Z_p = \left\{ \sum_{i=0}^p \lambda_i (c_i F(t_i)) \mid \lambda_i \in [0, 1] \right\}. \quad (4.4)$$

Z_p est le zonotope somme des $(p+1)$ segments $[0, c_i F(t_i)]$.

Pour $u \in \mathcal{S}$, notons $H_{Z_p}(u)$ la fonction support de Z_p . Nous avons donc :

$$H_{Z_p}(u) = \sum_{i=0}^p \langle c_i F(t_i), u \rangle_+. \quad (4.5)$$

Puisque les c_i sont positifs, les égalités (4.3) et (4.5) conduisent à la relation

$$H_{Z_p}(u) = I_p(u, F). \quad (4.6)$$

Nous avons la proposition suivante:

Proposition 4. 12 *Si pour tout $u \in \mathcal{S}$ la formule de quadrature, ayant c_i comme poids et t_i ($i = 0, \dots, p$) comme noeuds, est convergente pour la fonction $t \longrightarrow \langle F(t), u \rangle_+$, alors Z_p converge vers $Z(F)$ au sens de Hausdorff.*

Démonstration : Pour tout $u \in \mathcal{S}$, la formule de quadrature est convergente pour la fonction $t \longrightarrow \langle F(t), u \rangle_+$. Nous avons donc:

$$\sum_{i=0}^p \langle c_i F(t_i), u \rangle_+ \xrightarrow{p \rightarrow +\infty} \int_I \langle F(t), u \rangle_+ dt.$$

D'après les relations (4.1) et (4.6), on a : $(\forall u \in \mathcal{S})$

$$H_{Z_p}(u) \xrightarrow{p \rightarrow +\infty} H_{Z(F)}(u),$$

donc Z_p converge scalairement vers $Z(F)$. D'après la proposition 4.3, Z_p converge au sens de Hausdorff vers $Z(F)$. ■

Plus précisément, le théorème suivant donne l'erreur d'approximation de $Z(F)$ par Z_p .

Théorème 4.13 : Soient F une fonction intégrable sur $[a, b]$, $I_p(., F)$ une formule de quadrature ayant $c_i > 0$ comme poids et t_i ($i = 0, \dots, p$) comme noeuds et $R_p(., F)$ le reste correspondant. Considérons le zonotope défini par:

$$Z_p = \sum_{i=0}^p [0, c_i F(t_i)], \quad (4.7)$$

alors

$$\text{haus}(Z(F), Z_p) = \sup_{u \in \mathcal{S}} |R_p(u, F)|. \quad (4.8)$$

Démonstration : Les relations (4.1), (4.2) et (4.6) entraînent que pour $u \in \mathcal{S}$, on a:

$$\begin{aligned} |H_{Z(F)}(u) - H_{Z_p}(u)| &= \left| \int_a^b \langle F(t), u \rangle_+ dt - I_p(u, F) \right| \\ &= |R_p(u, F)|. \end{aligned}$$

Or $Z(F)$ et Z_p sont convexes compacts, donc en utilisant le lemme 4.4, on obtient le résultat. ■

2. 2. Majoration de l'erreur d'approximation $|R_p(u, F)|$ et ordre de convergence

Pour faire une majoration de l'erreur, nous faisons l'hypothèse suivante:

Hypothèses :

H_1) on suppose que pour tout $u \in \mathcal{S}$, le nombre d'intersections du plan de normale u passant par 0 avec la courbe de Γ est borné par une constante fixe m .

H_2) on suppose que la fonction F est de classe C^2 sur I .

Dans ce cas, pour tout u , la fonction $t \rightarrow \langle F(t), u \rangle_+$ est une fonction continue de classe C^0 qui est C^2 par morceaux (avec, au plus, $m + 1$ morceaux).

Nous utiliserons comme formule de quadrature la méthode des trapèzes composée sur p intervalles (la régularité de $\langle F(t), u \rangle_+$ ne justifie pas d'aller plus loin).

Dans ce qui suit, on prend les noeuds de la forme:

$$t_i = a + ih, \quad i = 0, \dots, p,$$

où

$$h = \frac{b-a}{p}.$$

Lemme 4.14 : Soit $g : [c, d] \rightarrow \mathbb{R}$ une fonction continue sur $[c, d]$; on suppose que g est dérivable sauf sur un ensemble fini $\mathcal{F}$ de points de $[c, d]$ et que g' est continue et bornée sur $[c, d] \setminus \mathcal{F}$: on posera $F_1 = \sup_{t \in [c, d] \setminus \mathcal{F}} |g'(t)|$. Notons:

$$T(g) = \frac{d-c}{2}(g(c) + g(d)),$$

$$Q(g) = \int_c^d g(t) dt,$$

$$E(g) = T(g) - Q(g).$$

Alors :

$$|E(g)| \leq \frac{(d-c)^2}{4} F_1.$$

Démonstration : D'après les hypothèses, on peut montrer que la fonction g est absolument continue. Donc, elle peut s'écrire sous la forme:

$$g(t) = g(c) + \int_c^t g'(u) du \quad (4.9)$$

$$= g(c) + \int_c^d \Psi(t, u) g'(u) du, \quad (4.10)$$

où

$$\Psi(t, u) = \begin{cases} 1 & \text{si } u \leq t, \\ 0 & \text{sinon.} \end{cases}$$

En utilisant l'expression de g donnée par (4.9), on obtient:

$$T(g) = g(c)(d-c) + \frac{(d-c)}{2} \int_c^d g'(u) du,$$

et en utilisant la relation (4.10),

$$Q(g) = g(c)(d-c) + \int_c^d \int_c^d \Psi(t, u) g'(u) du dt,$$

donc

$$\begin{aligned} E(g) &= T(g) - Q(g) \\ &= \int_c^d \left(\frac{(d-c)}{2} - \int_c^d \Psi(t, u) dt \right) g'(u) du \\ &= \int_c^d \left(\frac{(d-c)}{2} - (d-u) \right) g'(u) du, \end{aligned}$$

d'où nous tirons la majoration:

$$|E(g)| \leq \int_c^d \left| \frac{(d-c)}{2} - (d-u) \right| du \max_{u \in [c, d]} |g'(u)|,$$

ou encore:

$$|E(g)| \leq \frac{(d-c)^2}{4} F_1. \quad \blacksquare$$

Le lemme suivant donne le résultat classique de l'erreur d'intégration numérique par la méthode des trapèzes dans le cas où la fonction est de classe C^2 .

Lemme 4.15 *Soit $g : [c, d] \longrightarrow \mathbb{R}$ une fonction telle que g'' soit continue sur $[c, d]$ ($g \in C^2([c, d])$). Posons $F_2 = \max_{t \in [c, d]} |g''(t)|$.*

Alors $|E(g)| \leq \frac{(d-c)^3}{12} F_2$.

Si on applique l'un ou l'autre des deux lemmes précédents sur chaque intervalle $[t_i, t_{i+1}]$ pour intégrer une fonction g dont la dérivée seconde est continue sauf en un nombre fini de points, l'erreur est donnée par le théorème suivant:

Théorème 4.16 *Soit $g : [a, b] \longrightarrow \mathbb{R}$ une fonction continue sur $[a, b]$. On suppose que g est 2 fois continûment dérivable sauf en un ensemble fini $\mathcal{F}$ de points de $[a, b]$ et que g' et g'' sont continues bornées sur $[a, b] \setminus \mathcal{F}$:*

$$F_1 = \sup_{t \in [c, d] \setminus \mathcal{F}} |g'(t)|,$$

$$F_2 = \sup_{t \in [c, d] \setminus \mathcal{F}} |g''(t)|.$$

On note $I_p(g)$ la formule des trapèzes sur $[a, b]$ de pas $h = \frac{b-a}{p}$ et

$$Q(g) = \int_a^b g(t) dt,$$

$$e_p(g) = I_p(g) - Q(g).$$

Alors si $m = \text{card}(\mathcal{F})$, on a :

$$|e_p(g)| \leq kh^2,$$

$$\text{avec } k = \left(\frac{b-a}{12} F_2 + \frac{m}{4} F_1 \right).$$

Démonstration : On a $t_i = a + ih$, $i = 0, \dots, p$. Soit $\tilde{m}$ le nombre d'intervalles $[t_i, t_{i+1}]$ qui contiennent des points de $\mathcal{F}$. on a $\tilde{m} \leq m$.

$$\begin{aligned} |e_p(g)| &\leq (p - \tilde{m}) \frac{h^3}{12} F_2 + \tilde{m} \frac{h^2}{4} F_1 \\ &\leq p \frac{h^3}{12} F_2 + m \frac{h^2}{4} F_1 \\ &\leq \left(\frac{b-a}{12} F_2 + \frac{m}{4} F_1 \right) h^2. \quad \blacksquare \end{aligned}$$

On peut donc énoncer le résultat suivant:

Théorème 4.17 : Soit $F : I \longrightarrow \mathbb{R}^n$ une fonction de classe C^2 , et soit $Z(F)$ le zonoïde associé à F . Alors la séquence de zonotopes Z_p , $p \in \mathbb{N}$, où Z_p est défini par la somme des segments :

$$N_0 = \frac{h}{2}F(a), \quad N_i = hF(t_i), \quad i = 1, \dots, p-1, \quad N_p = \frac{h}{2}F(b),$$

converge vers $Z(F)$ au sens de la métrique de Hausdorff, et on a :

$$\text{haus}(Z(F), Z_p) \leq Kp^{-2}.$$

où K est une constante qui dépend de F .

Démonstration : Pour tout $u \in \mathcal{S}$, en appliquant le résultat du théorème précédent à la fonction $t \longrightarrow \langle F(t), u \rangle_+$, on obtient :

$$\left| \int_a^b \langle F(t), u \rangle_+ dt - I_p(u, F) \right| \leq \left(\frac{b-a}{12} F_2(u) + \frac{m}{4} F_1(u) \right) h^2,$$

avec :

$$F_1(u) = \max_{t \in I} |\langle F'(t), u \rangle|,$$

$$F_2(u) = \max_{t \in I} |\langle F''(t), u \rangle|.$$

Or d'après l'hypothèse H_2 , $F_1(u)$ et $F_2(u)$ sont bornées par rapport à $u \in \mathcal{S}$. Soient

$$F_1 = \sup_{u \in \mathcal{S}} F_1(u),$$

et

$$F_2 = \sup_{u \in \mathcal{S}} F_2(u).$$

Alors

$$\left| \int_a^b \langle F(t), u \rangle_+ dt - I_p(u, F) \right| \leq \left(\frac{b-a}{12} F_2 + \frac{m}{4} F_1 \right) h^2,$$

pour tout u dans $\mathcal{S}$.

On aboutit donc au résultat :

$$|H_{Z(F)}(u) - H_{Z_p}(u)| \leq \frac{K}{p^2} \quad \forall u \in \mathcal{S},$$

où

$$K = \left(\frac{b-a}{12} F_2 + \frac{m}{4} F_1 \right) (b-a)^2,$$

dépend de la fonction F et de ses dérivées première et seconde mais ne dépend pas de u .

D'où d'après le lemme 4.4:

$$\text{haus}(Z(F), Z_p) \leq \frac{K}{p^2}. \quad \blacksquare$$

Corollaire 4. 18 *Soient F une fonction définie de I à valeurs dans $\mathbb{R}^n$ et $Z(F)$ le zonoïde associé. Pour créer un zonotope qui approche $Z(F)$ à ε près, il suffit que le nombre de segments soit en $O(\frac{1}{\sqrt{\varepsilon}})$.*

Remarques :

- 1- La convergence d'ordre 2 par rapport au nombre de segments dans le théorème précédent est due à l'utilisation de la méthode des trapèzes. Plus généralement, on peut utiliser n'importe quelle méthode d'intégration numérique pourvu qu'elle soit d'ordre supérieure ou égale à 1 pour préserver la convergence quadratique.
- 2- Si l'on utilise une méthode d'ordre supérieur à 1, on définira une manière différente de construire une séquence de zonotopes qui converge vers $Z(F)$. Cependant la convergence en $O(\frac{1}{p^2})$ ne peut être améliorée; en effet pour $u \in \mathcal{S}$, la fonction $t \longrightarrow \langle F(t), u \rangle_+$ est seulement C^0 .
- 3- En appliquant la méthode des rectangles avec point milieu, on obtient les zonotopes Z_p comme somme des segments $[O, N_i]$ où

$$N_i = [0, hF(\frac{t_i + t_{i+1}}{2})], \quad i = 1, \dots, p. \quad (4.11)$$

De même en appliquant la méthode de Simpson, on obtient les zonotopes Z_p comme somme de Minkowski des segments $[O, N_i]$ où : (on suppose que p est pair)

$$\begin{aligned} N_0 &= [0, \frac{h}{6}F(a)], \\ N_i &= [0, \frac{2h}{3}F(t_i)], \quad \text{si } i \text{ est impair,} \\ N_i &= [0, \frac{h}{3}F(t_i)], \quad \text{si } i \text{ est pair,} \\ N_p &= [0, \frac{h}{6}F(b)]. \end{aligned} \quad (4.12)$$

- 4- Le nombre de segments $O(\frac{1}{\sqrt{\varepsilon}})$ dans le cas des zonoïdes associés à une fonction est meilleur que celui donné par la théorie générale et il est du même ordre que celui de l'approximation de la boule unité dans $\mathbb{R}^2$.

2. 3. Exemples numériques:

Pour estimer la distance de Hausdorff entre deux ensembles, on fait la discrétisation:

$$haus(A, B) = \max_{i=1, \dots, l} |H_A(u_i) - H_B(u_i)|,$$

où $u_i, i = 1, \dots, l$ sont l vecteurs uniformément répartis dans $\mathcal{S}$.

Dans les exemples qui suivent, sauf précision contraire, Z_p est défini par les vecteurs donnés par le théorème 4. 17.

• Exemple dans $\mathbb{R}^2$

Exemple 1: Soient $I = [0, \pi]$ et $F : I \longrightarrow \mathbb{R}^2$ telle que:

$$F(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}.$$

Pour $u = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathcal{S}$,

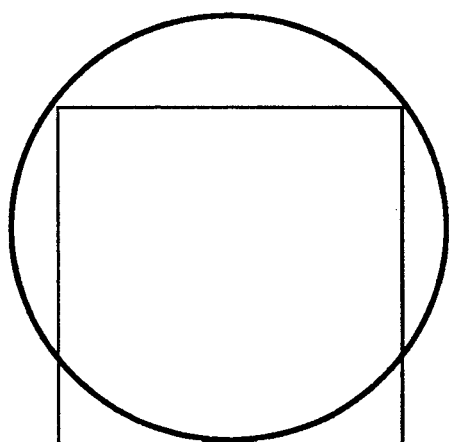
$$H_{Z(F)}(u) = \int_0^\pi (x \cos(t) + y \sin(t))_+ dt.$$

Dans ce cas, on démontre que $Z(F)$ est un cercle de centre (0,1) et de rayon 1.

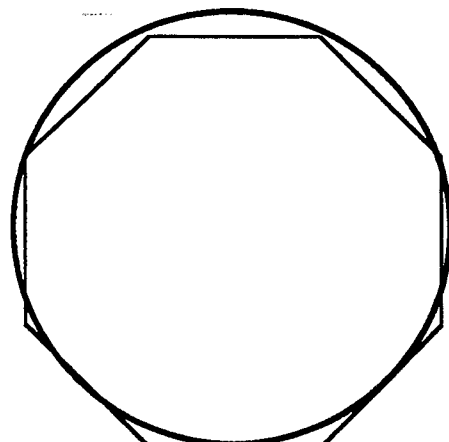
Dans le tableau suivant, on fait une illustration de la convergence quadratique:

p	$haus(Z_p, Z(F))$	ordre
2	0.4292034710	
4	0.1038809237	2.0473
8	0.0257682277	2.0113
16	0.0064294761	2.0028
2	0.4292034710	
20	0.0041138125	2.0159
200	0.0000409371	2.002
2000	0.0000003113	2.119

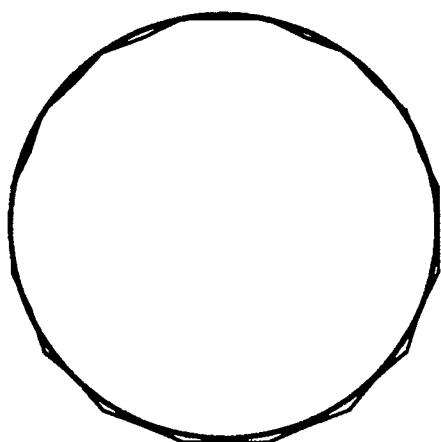
La valeur de l'ordre correspondant à $p = 2000$, est moins bonne. Ceci est dû au fait qu'à p élevé, on fait la somme d'un grand nombre de quantités très petites.



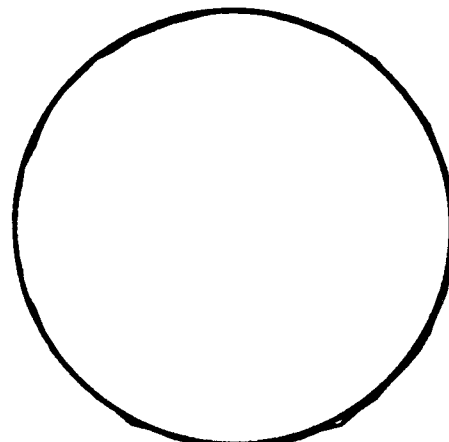
p = 2, (3 segments)



p = 4, (5 segments)



p = 8, (9 segments)



p = 16, (17 segments)

FIGURE .1. Exemple 1, approximation du cercle ($N_p = -N_0$)

Exemple 2:

$$F(t) = \begin{pmatrix} t \\ 1+t^2 \end{pmatrix}, \quad t \in [0, 2,]$$

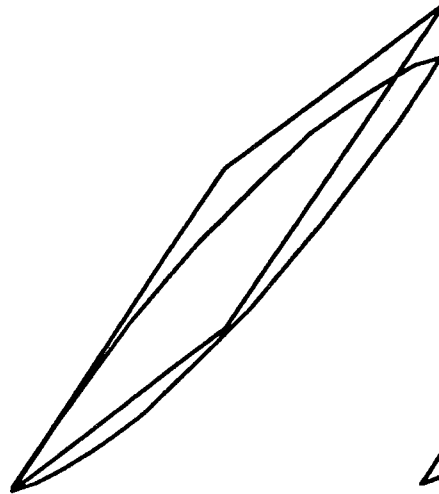
Pour $u = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathcal{S}$,

$$H_{Z(F)}(u) = \int_0^2 (x \cdot t + y \cdot (1+t^2))_+ dt.$$

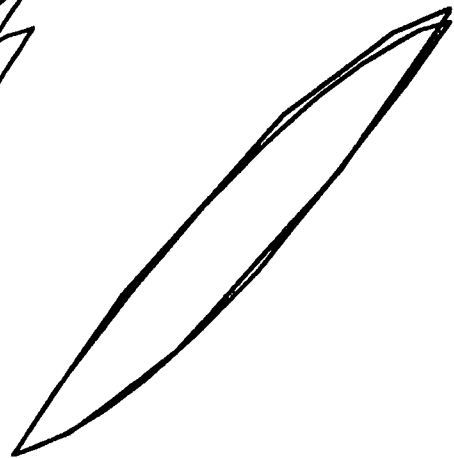
Dans les tableaux suivants, on fait une illustration de la convergence quadratique:

p	$haus(Z_p, Z(F))$	ordre
2	30.1516621485	
4	7.5365131105	2.0003
8	1.8827258510	2.0011
16	0.4692790361	2.0043
32	0.1159173324	2.017
64	0.0275769064	2.07
128	0.0054918000	2.33

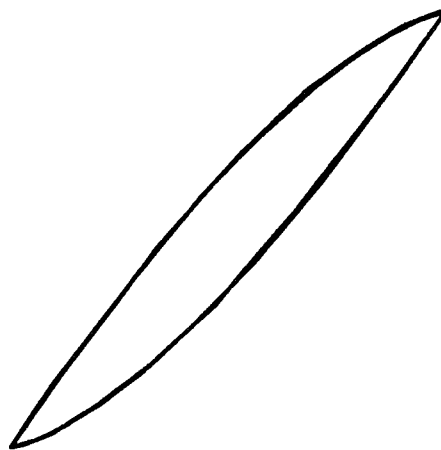
p	$haus(Z_p, Z(F))$	ordre
2	30.1535182343	
20	0.3015185489	2.000
200	0.0030012619	2.002



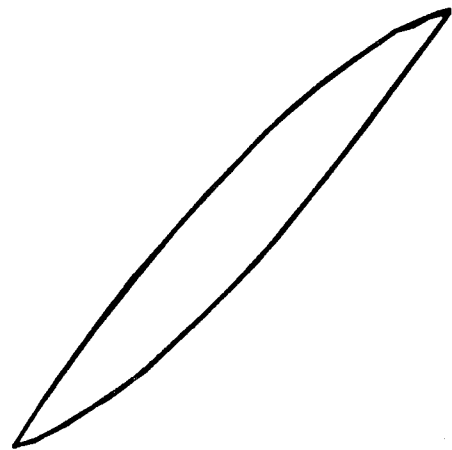
p = 2



p = 4

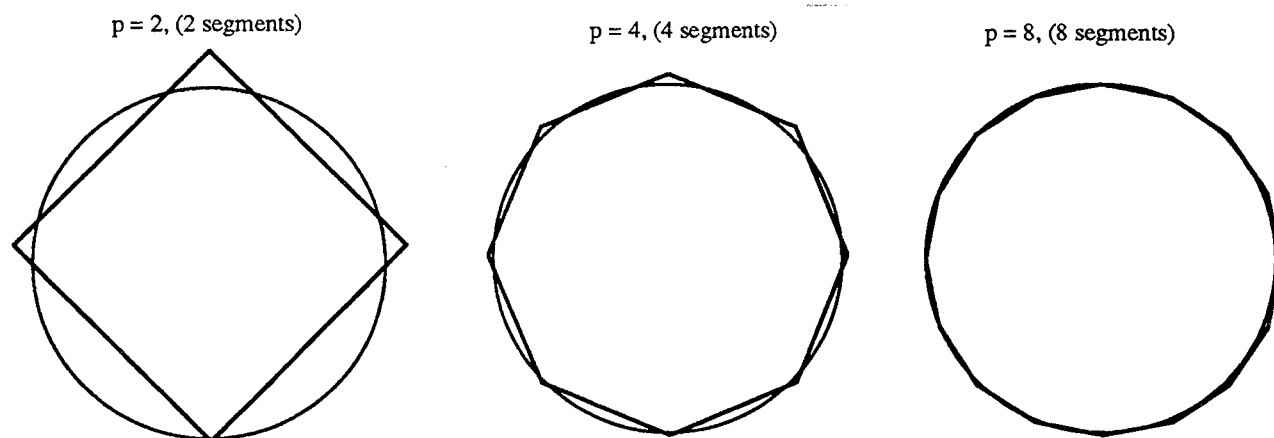


p = 8

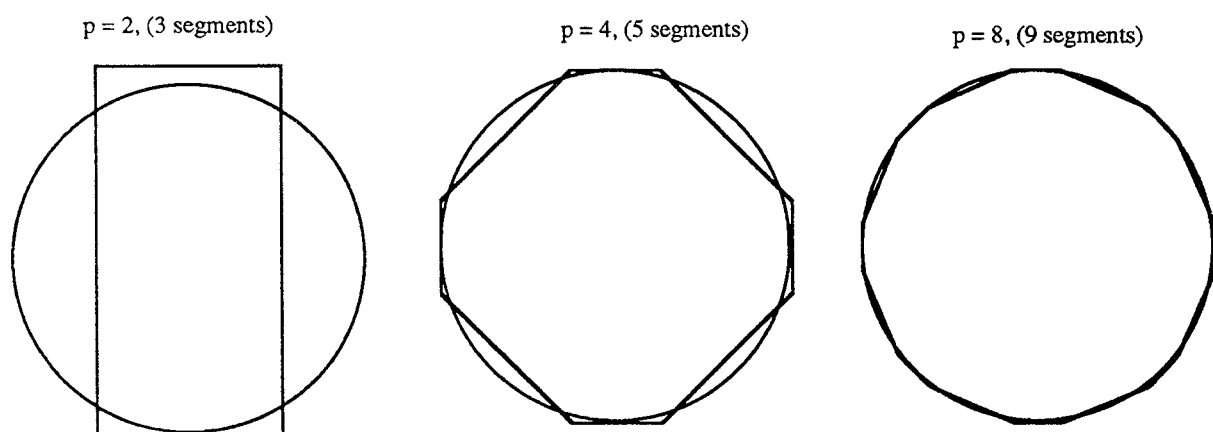


p = 16

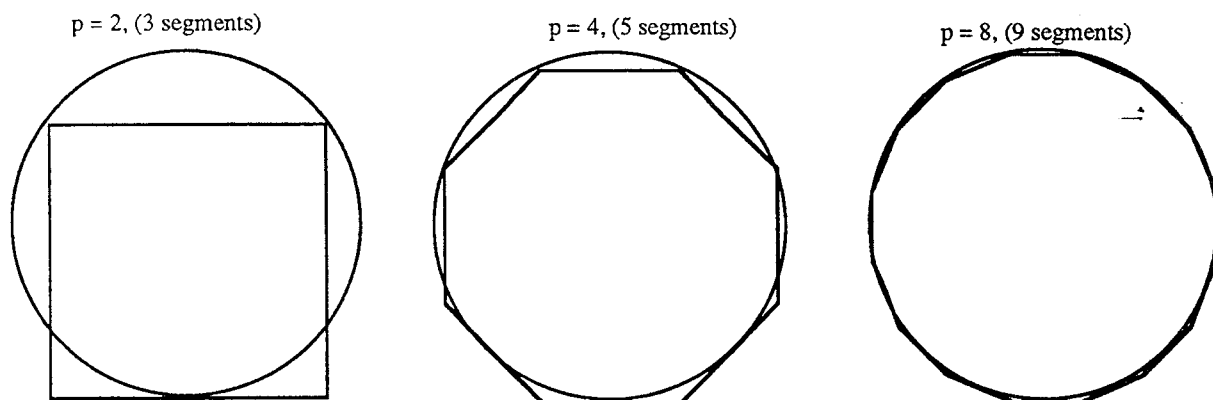
FIGURE .2. Exemple 2, approximation.



Premières itérations associées à la discrétisation donnée par (4.11)



Premières itérations associées à la discrétisation donnée par (4.12)



Premières itérations associées à la discrétisation donnée par le théorème 4.17

FIGURE .3. Comparaison entre différentes séquences.

• Exemples dans $\mathbb{R}^3$

Exemple 3

$$F(t) = \begin{pmatrix} t^2 \\ t \\ 1 \end{pmatrix}, \quad \forall t \in [0, 1].$$

p	$haus(Z_p, Z(F))$	ordre	p	$haus(Z_p, Z(F))$	ordre
4	0.012428		64	0.000049	2.000880
8	0.003340	1.895424	128	0.000012	1.993164
16	0.000703	2.248478	256	0.000003	2.013226
32	0.000196	1.843059	512	0.000001	1.962027

La figure 4 donne les zonotopes correspondants à quelques valeurs de p .

Exemple 4

$$F(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ 1 \end{pmatrix} \quad \forall t \in [-\pi, \pi]$$

p	$haus(Z_p, Z(F))$	ordre	p	$haus(Z_p, Z(F))$	ordre
4	0.429203		64	0.001607	2.000578
8	0.103881	2.046730	128	0.000402	1.999661
16	0.025768	2.011255	257	0.000101	1.997912
32	0.006430	2.002765	512	0.000025	1.991353

La figure 5 donne les zonotopes correspondants à quelques valeurs de p .

Exemple 5

$$F(t) = \begin{pmatrix} t^3 \\ t \\ 1 \end{pmatrix} \quad \forall t \in [-1, 1]$$

p	$haus(Z_p, Z(F))$	ordre	p	$haus(Z_p, Z(F))$	ordre
4	0.073064		64	0.000274	1.977107
8	0.018357	1.992826	128	0.000067	2.030645
16	0.003920	2.227554	256	0.000018	1.922169
32	0.001080	1.859781	512	0.000004	2.107282

La figure 6 donne les zonotopes correspondants à quelques valeurs de p .

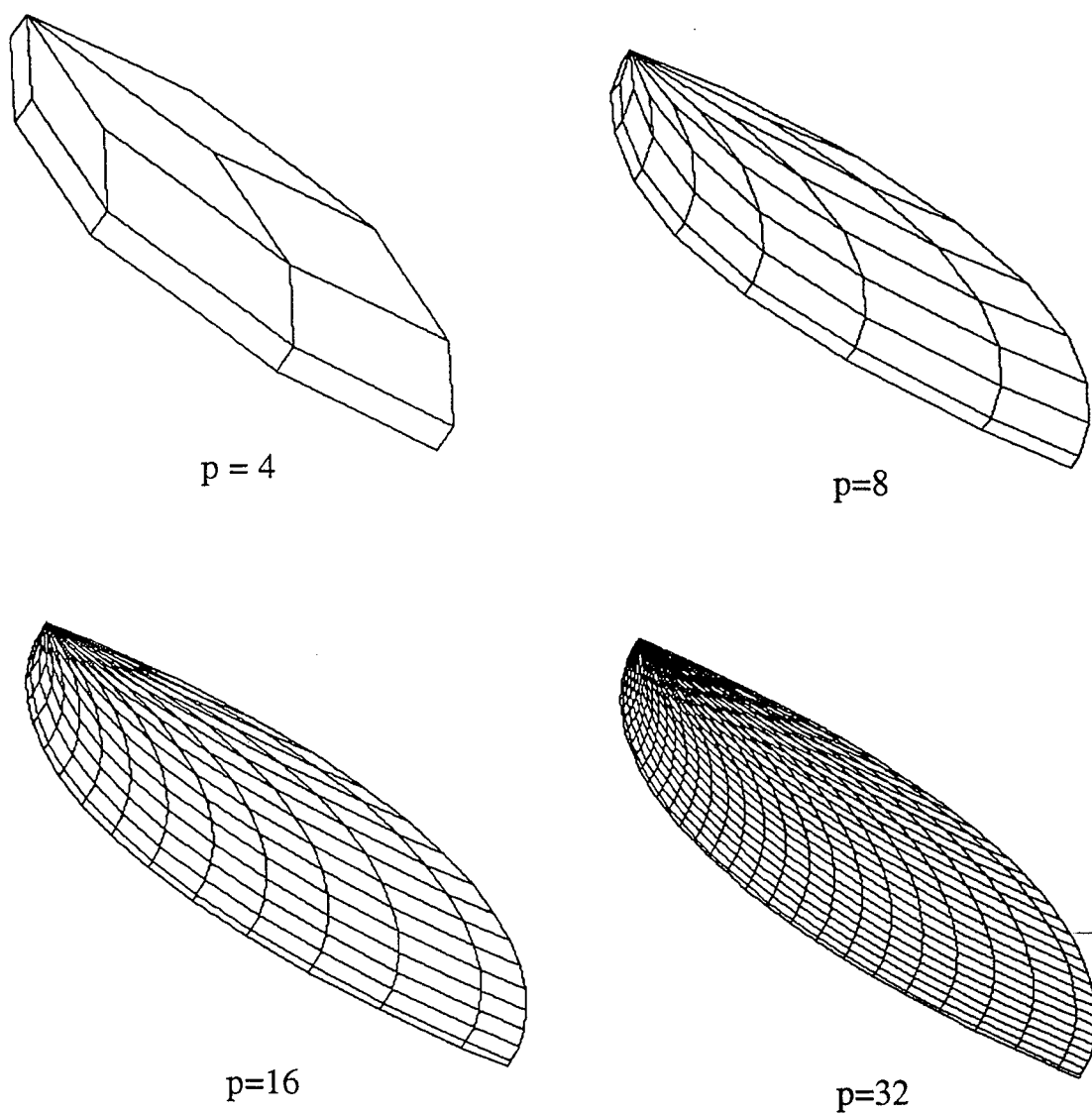


FIGURE .4. Exemple 3, convergence de Z_p dans $\mathbb{R}^3$.

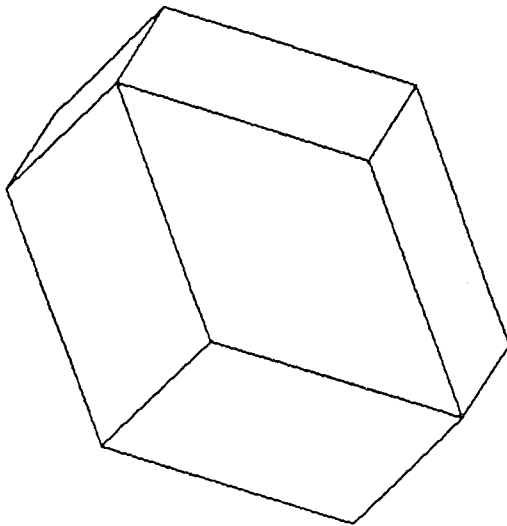
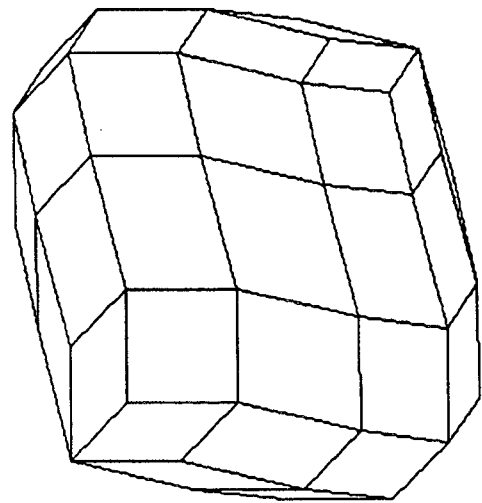
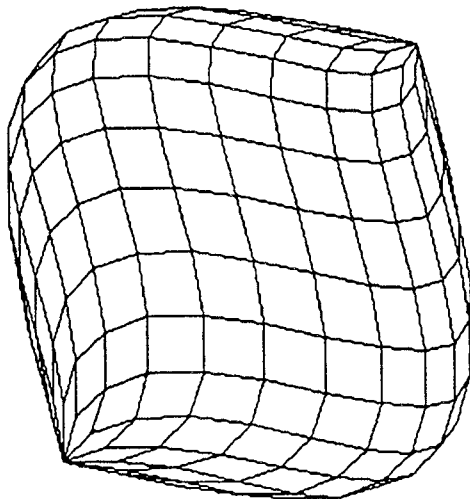
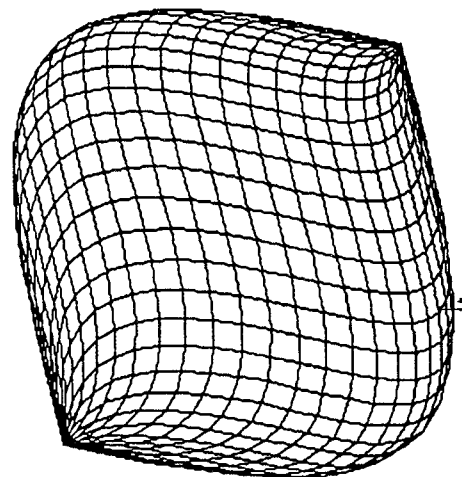
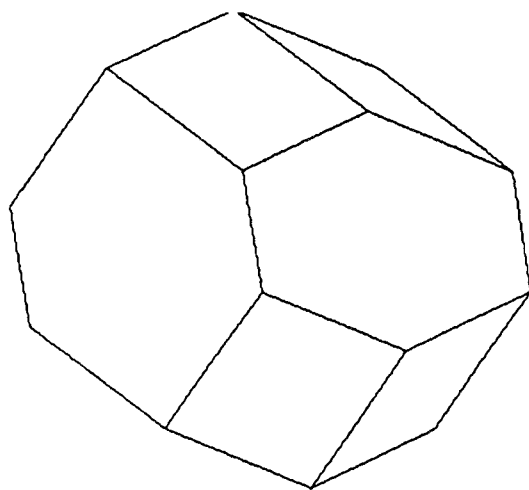
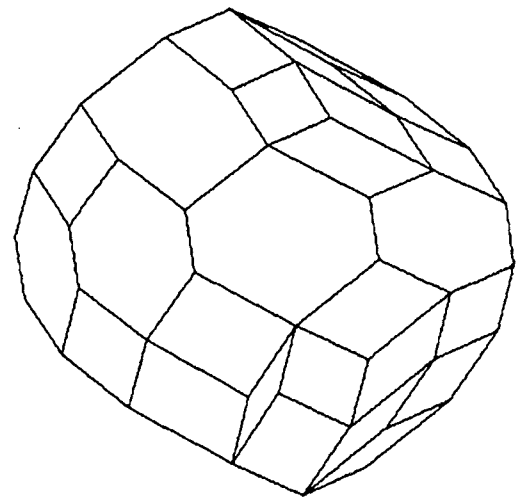
 $p = 4$  $p = 8$  $p = 16$  $p = 32$

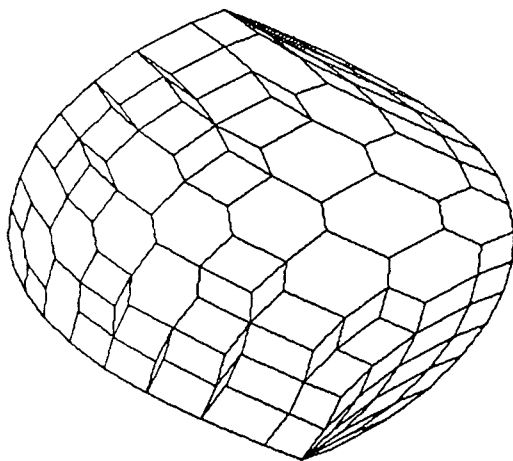
FIGURE .5. Exemple 4, convergence de Z_p dans $\mathbb{R}^3$.



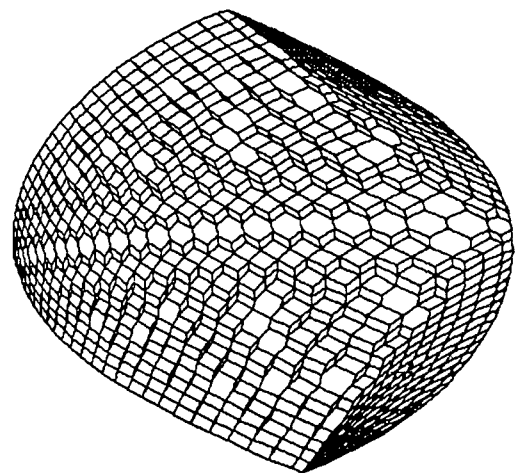
$p = 4$



$p=8$



$p=16$



$p=50$

FIGURE .6. Exemple 5, convergence de Z_p dans $\mathbb{R}^3$.

2. 4. Approximation dans le cas de la sélectivité régulière dans $\mathbb{R}^3$

Définition 4. 19 Un zonotope Z_p est régulier, si les vecteurs de base N_i , $i = 0, \dots, p$ sont tels que:

- il existe $\mu \in \mathbb{R}^3$ tel que $\langle N_i, \mu \rangle > 0$,
- toutes les demi-droites $\{\lambda N_i, \lambda \geq 0\}$ sont des rayons extrémaux du cône positif engendré par les vecteurs N_i , $i = 1, \dots, p$.

Remarque 1: Dans le diagramme projectif, cette définition se traduit par le fait que tous les points M_i sont des sommets de leur enveloppe convexe.

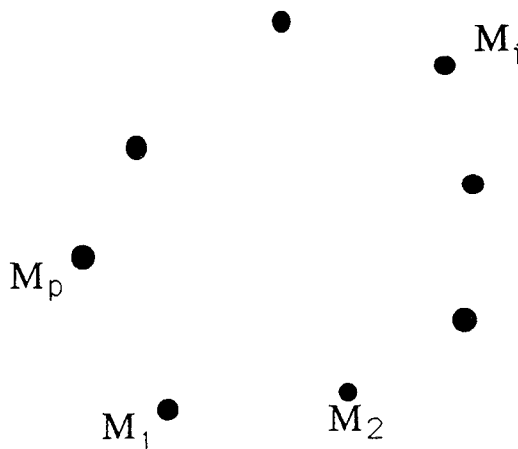


FIGURE .7. Premier diagramme projectif associé à un zonopèdre régulier

Proposition 4. 20 Si $F : [a, b] \longrightarrow \mathbb{R}^3$ admet la propriété de la sélectivité régulière alors tout zonopèdre Z_p défini par la relation (4.4) est régulier.

Démonstration : Si Z_p est un zonopèdre défini par la relation (4.4), alors les vecteurs de bases N_i sont de la forme:

$$N_i = c_i F(t_i), \quad i = 1, \dots, p,$$

et les c_i sont des constantes positives.

Pour tout i, j, k vérifiant $0 \leq i < j < k \leq p$, on a $t_i < t_j < t_k$. Puisque F est STR, et les poids c_i sont positifs alors pour tout i, j, k vérifiant $0 \leq i < j < k \leq p$, on a

$$\det(N_i, N_j, N_k) = c_i c_j c_k \det(F(t_i), F(t_j), F(t_k)),$$

garde un signe constant. ■

3. Quelques exemples d'applications

• Appartenance d'un point à un zonoïde

Un point M de E est un élément de $Z(F)$, s'il existe une fonction $\lambda_M \in \Lambda$ telle que $M = \int_I \lambda_M(t)F(t)dt$. Le problème de la recherche de λ_M n'est pas facile et même en général impossible à résoudre. Une solution approchée consiste à se ramener à la détermination de l'appartenance de M à un zonotope Z_p (p donné) qui approche $Z(F)$.

• Intersection d'un hyperplan avec un zonoïde

Dans le cas où F est STR, on peut appliquer la méthode décrite dans le chapitre 3. Plus généralement, en approchant F par un zonotope, on peut se ramener à un problème d'intersection d'un plan avec un zonotope.

• Volume d'un zonoïde

Nous avons vu dans le chapitre 1 que pour un zonotope Z défini comme somme de Minkowski de segments $[0, N_i]$ $i = 0, \dots, p$, le volume de Z s'écrit:

$$Vol(Z) = \frac{1}{n!} \sum_{i_j \in \{0, \dots, p\}} |\det(N_{i_1}, N_{i_2}, \dots, N_{i_n})|.$$

Quand on passe au cas continu, le volume du zonoïde $Z(F)$ est donné par,

$$Vol(Z(F)) = \frac{1}{n!} \int_I \int_I \dots \int_I |\det(F(t_1), F(t_2), \dots, F(t_n))| dt_1 dt_2 \dots dt_n. \quad (4.13)$$

Dans le cas où F est une fonction STR à valeurs dans $\mathbb{R}^3$, l'expression (4.13) se simplifie:

Proposition 4. 21 [Val91] *Soit F une fonction STR sur $[a, b]$ de primitive N vérifiant $N(a) = 0$ et $Z(F)$ le zonoïde associé à F . Alors le volume du zonoïde $Z(F)$ est donné par:*

$$Vol(Z(F)) = \langle N(b), \int_a^b F(t) \wedge N(t) dt \rangle. \quad (4.14)$$

La valeur du volume d'un zonoïde peut être très difficile à calculer d'une manière exacte. On peut l'approcher par la valeur du volume du zonotope Z_p (p donné) qui approche $Z(F)$.

Exemples

Pour chaque exemple, le volume du zonoïde est connu. On calcule le volume du zonotope approchant et l'erreur relative. L'erreur relative notée e_p est donnée par la formule suivante:

$$e_p = \left| \frac{\text{Vol}(Z(F)) - \text{Vol}(Z_p)}{\text{Vol}(Z(F))} \right|.$$

1-

$$F(t) = \begin{pmatrix} 5 \\ 6t \\ 6t^2 \end{pmatrix}, \quad \forall t \in [0, 1].$$

$$\text{Vol}(Z(F)) = 1.$$

2-

$$F(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ 1 \end{pmatrix}, \quad \forall t \in [-\pi, \pi].$$

$$\text{Vol}(Z(F)) = 4\pi^2.$$

3-

$$F(t) = \begin{pmatrix} e^{-t} \\ e^{-2t}/2 \\ e^{-3t}/3 \end{pmatrix}, \quad \forall t \in [0, +\infty[.$$

$$\text{Vol}(Z(F)) = \frac{\sqrt{2}}{1080}.$$

4-

$$F(t) = \begin{pmatrix} t^3 \\ t \\ 1 \end{pmatrix}, \quad \forall t \in [-1, 1].$$

$$\text{Vol}(Z(F)) = \frac{313}{1680}.$$

Dans le tableau, on donne les erreurs relatives en fonction de p pour chacun des exemples.

Dans la figure 8, on donne la variation du $\log(e_p)$ en fonction de $\log(p)$.

On remarque que pour chaque exemple les points sont alignés, avec une pente très proche de -2 (problème à l'exemple 3, surtout pour les premières valeurs de p ; comme le montre le graphique). Donc, la convergence vers la valeur 0 de la suite e_p est quadratique.

Erreur relative e_p

p	Exemple 1	Exemple 2	Exemple 3	Exemple 4
20	0.006898	0.009130	0.332901	0.018441
30	0.002968	0.003915	0.200290	0.007966
40	0.001642	0.002164	0.123822	0.004414
50	0.001041	0.001371	0.082453	0.002799
60	0.000718	0.000945	0.058377	0.001932
70	0.000525	0.000691	0.043325	0.001413
80	0.000400	0.000527	0.033348	0.001078
90	0.000316	0.000415	0.026418	0.000849
100	0.000255	0.000336	0.021418	0.000687

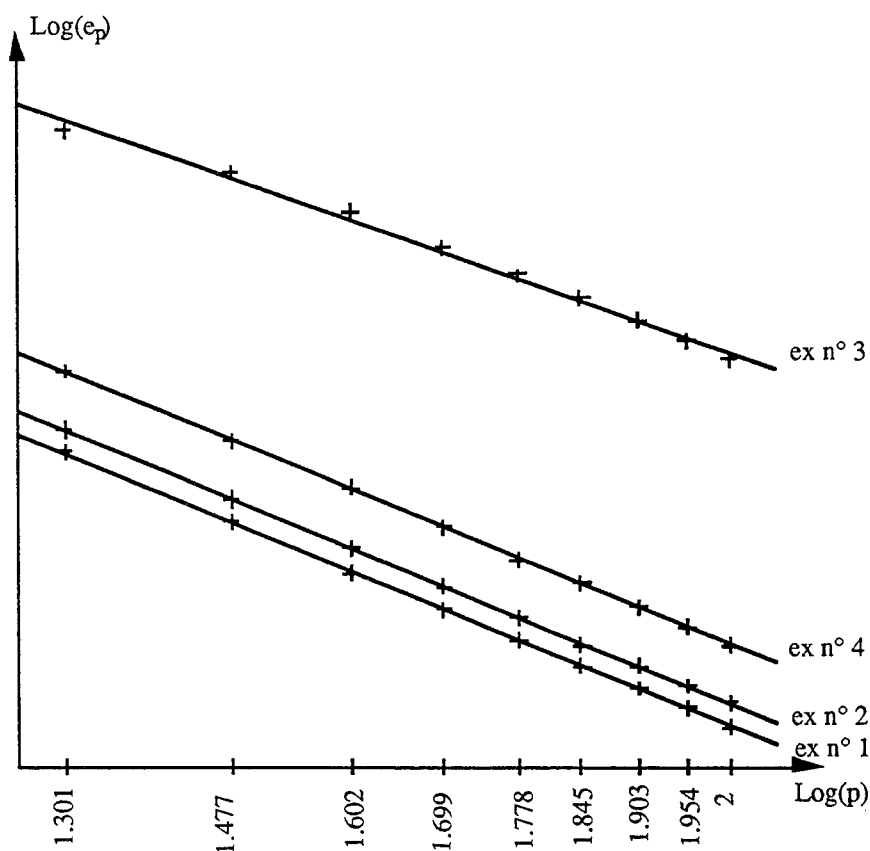


FIGURE .8. Ordre de convergences

L'équation de chaque droite (moindres carrés) est:

exemple 1 : $y = -2.047x + 1.1446$,

exemple 2 : $y = -2.050x + 1.4329$,

exemple 3 : $y = -1.745x + 4.2626$,

exemple 4 : $y = -2.043x + 2.1180$.

4. Zonoïde associé à une fonction de plusieurs variables

Soient $F : D = [a, b] \times [c, d] \longrightarrow \mathbb{R}^n$ une fonction de classe C^2 , Λ l'ensemble des fonctions mesurables définies sur D à valeurs dans $[0, 1]$.
Soit $Z(F)$ le zonoïde défini par :

$$Z(F) = \left\{ \int \int_D F(x, y) \lambda(x, y) dx dy; \quad \lambda \in \Lambda \right\}.$$

4. 1. Construction du zonotope approximant

De la même manière que dans la partie 2, on définira une suite de zonotopes Z_q qui converge, au sens de la métrique de Hausdorff, vers $Z(F)$ quand $q \rightarrow +\infty$. Comme précédemment, nous pouvons établir que la fonction support de $Z(F)$ s'écrit, pour $u \in \mathcal{S}$:

$$H_{Z(F)}(u) = \int \int_D \langle F(x, y), u \rangle_+ dx dy. \quad (4.15)$$

En utilisant une méthode d'intégration numérique à deux variables, on peut écrire:

$$H_{Z(F)}(u) = I_q(u, F) + R_q(u, F), \quad (4.16)$$

où

$$I_q(u, F) = \sum_{i=0}^q c_i \langle F(x_i, y_i), u \rangle_+. \quad (4.17)$$

Les (x_i, y_i) sont les noeuds, les c_i sont des poids positifs, et $R_q(u, F)$ est le reste de la formule d'intégration.

Soit Z_q le zonotope défini par:

$$Z_q = \sum_{i=0}^q [0, c_i F(x_i, y_i)]. \quad (4.18)$$

Puisque les poids c_i sont positifs, on démontre que:

$$H_{Z_q}(u) = I_q(u, F). \quad (4.19)$$

Dans ce cas, on a le même résultat donné par le théorème 4. 13, c'est à dire que:

$$haus(Z(F), Z_q) = \sup_{u \in \mathcal{S}} |R_q(u, F)|. \quad (4.20)$$

4. 2. Majoration de $|R_q(u, F)|$ et ordre de convergence

Pour majorer $|R_q(u, F)|$, nous avons besoin de résultats sur l'erreur d'intégration à plusieurs variables. Nous utiliserons l'extension de la méthode des trapèzes à deux variables. On a les deux lemmes suivants:

Lemme 4. 22 Soit $f : D = [a, b] \times [c, d] \longrightarrow \mathbb{R}$ une fonction de classe C^2 .

Posons

$$F_{ij} = \sup_{(x,y) \in D} \left| \frac{\partial^{i+j} f}{\partial x^i \partial y^j}(x, y) \right|, \quad i, j = 0, 1, 2 \quad i + j \leq 2$$

$$T(f) = \frac{(b-a)(d-c)}{4} (f(a, c) + f(a, d) + f(b, c) + f(b, d)),$$

$$Q(f) = \iint_D f(x, y) dx dy,$$

et

$$E(f) = T(f) - Q(f).$$

Alors:

$$|e(f)| \leq \frac{(b-a)^3(d-c)}{12} F_{20} + \frac{(b-a)^2(d-c)^2}{2} F_{11} + \frac{(b-a)(d-c)^3}{12} F_{02}.$$

Démonstration :

$$\begin{aligned} f(x, y) &= f(a, c) + \partial_x f(a, c)(x - a) + \partial_y f(a, c)(y - c) + \partial_{xy} f(a, c)(x - a)(y - c) \\ &\quad + \int_0^1 (1-t) \partial_{x^2} f(a + t(x-a), c)(x-a)^2 dt \\ &\quad + \int_0^1 (1-t) \partial_{xy} f(a + t(x-a), c)(x-a)(y-c) dt \\ &\quad + \int_0^1 (1-t) \partial_{y^2} f(a, c + t(y-c))(y-c)^2 dt. \end{aligned}$$

En substituant cette expression dans $T(f)$ et $Q(f)$, et en majorant, on obtient le résultat. ■

Lemme 4. 23 Soit f une fonction définie sur $D = [a, b] \times [c, d]$ à valeurs dans $\mathbb{R}$. On suppose que f est k -lipschitzienne. Alors :

$$|E(f)| \leq k(b-a)(d-c)\sqrt{(b-a)^2 + (d-c)^2}. \quad (4.21)$$

Démonstration : La fonction f est k -lipschitzienne, donc

$$|f(x, y) - f(a, c)| \leq k\sqrt{(b-a)^2 + (d-c)^2}, \quad \forall (x, y) \in D,$$

donc pour tout $(x, y) \in D$, on a:

$$f(a, c) - k\sqrt{(b-a)^2 + (d-c)^2} \leq f(x, y) \leq f(a, c) + k\sqrt{(b-a)^2 + (d-c)^2}.$$

En appliquant Q et T à chacun des membres des inégalités ci-dessus, on obtient le résultat. ■

On suppose que $D = [0, 1] \times [0, 1]$, et on considère la subdivision de D , en carrés élémentaires, dont la longueur des côtés est $h = 1/p$. On note:

$$x_i = ih \quad \text{et} \quad y_j = jh.$$

La formule des trapèzes à deux variables appliquée sur les p^2 carrés s'écrit:

$$\sum_{i,j=1}^p r_{ij} f(x_i, y_j),$$

où

$$r_{ij} = h^2 \begin{cases} 1/4 & \text{si } (x_i = 0 \text{ ou } 1) \text{ et } (y_j = 0 \text{ ou } 1), \\ 1/2 & \text{si } (x_i = 0 \text{ ou } 1) \text{ et } (y_j \neq 0 \text{ et } 1), \\ & \text{ou si } (y_j = 0 \text{ ou } 1) \text{ et } (x_i \neq 0 \text{ et } 1), \\ 1 & \text{si } x_i \neq 0, \quad x_i \neq 1, \quad y_j \neq 0, \quad y_j \neq 1. \end{cases} \quad (4.22)$$

Théorème 4.24 Soit f une fonction définie sur $D = [0, 1] \times [0, 1]$ à valeurs dans $\mathbb{R}$, k -lipschitzienne, et de classe C^2 sauf sur un ensemble $\mathcal{F}$ de points où elle est seulement C^0 . On suppose qu'il existe une constante m telle que le nombre de carrés élémentaires contenant des points de l'ensemble $\mathcal{F}$ est inférieur à mp , pour tout p suffisamment grand. Notons:

$I_{p^2}(f)$ la formule des trapèzes sur les p^2 carrés,

$$Q(f) = \int_0^1 \int_0^1 f(x, y) dx dy,$$

et

$$E_{p^2}(f) = I_{p^2}(f) - Q(f).$$

Alors il existe une constante K ne dépendant pas de p telle que

$$|E_{p^2}(f)| \leq Kh^2.$$

Démonstration : Notons pour i et j positifs ou nuls, $i + j \leq 2$

$$F_{ij} = \max_{(x,y) \in D \setminus \mathcal{F}} \frac{\partial^{i+j} f}{\partial x^i \partial y^j}(x, y),$$

et $M = \frac{F_{20}(u)}{12} + \frac{F_{11}(u)}{2} + \frac{F_{02}(u)}{12}$. Alors en appliquant les deux lemmes précédents, on obtient la majoration suivante:

$$|E_{p^2}(f)| \leq \sum_{i=0}^{mp} \frac{kh^3}{\sqrt{2}} + \sum_{i=0}^{p^2} Mh^4.$$

Et donc:

$$|E_{p^2}(f)| \leq \left(\frac{mk}{\sqrt{2}} + M\right)h^2. \quad \blacksquare$$

Théorème 4.25 Soit $F : D = [0, 1] \times [0, 1] \longrightarrow \mathbb{R}^n$, une fonction de classe C^2 sur D , et soit $Z(F)$ le zonoïde associé. Soit

$$\mathcal{F}_u = \{(x, y) \in D \mid \langle F(x, y), u \rangle = 0\}.$$

On suppose qu'il existe une constante m telle pour tout $u \in \mathcal{S}$ le nombre de carrés élémentaires qui contiennent des points de $\mathcal{F}_u$ est inférieur à mp . Considérons Z_{p^2} le zonotope défini par la somme de segments:

$$[0, r_{ij}F(x_i, y_j)] \quad i = 0, \dots, p, \quad j = 0, \dots, p$$

où les r_{ij} sont les constantes définies par la relation (4.22).

Alors la séquence de zonotopes Z_{p^2} , $p \in \mathbb{N}$ converge vers $Z(F)$, au sens de la métrique de Hausdorff et il existe une constante K ne dépendant pas de p telle que :

$$\text{haus}(Z(F), Z_{p^2}) \leq \frac{K}{p^2}.$$

Démonstration : F étant de classe C^2 sur D , il existe une constante k telle pour tout $u \in \mathcal{S}$, la fonction $(x, y) \longrightarrow \langle F(x, y), u \rangle_+$ est k -lipschitzienne.

Pour i et j positifs, $i + j \leq 2$, posons :

$$F_{ij}(u) = \max_{(x, y) \in D} \left\langle \frac{\partial^{i+j} F}{\partial_{x^i} \partial_{y^j}}(x, y), u \right\rangle,$$

$$\text{et } M(u) = \frac{F_{20}(u)}{12} + \frac{F_{11}(u)}{2} + \frac{F_{02}(u)}{12}.$$

Pour tout $u \in \mathcal{S}$, la fonction $(x, y) \longrightarrow \langle F(x, y), u \rangle_+$ satisfait les hypothèses du théorème 4. 24, donc pour tout $u \in \mathcal{S}$

$$\left| \int_0^1 \int_0^1 \langle F(x, y), u \rangle_+ dx dy - I_{p^2}(u, F) \right| \leq \frac{K(u)}{p^2},$$

$$\text{où } K(u) = \frac{mk}{\sqrt{2}} + M(u).$$

Donc en posant

$$M = \max_{u \in \mathcal{S}} M(u),$$

on a:

$$\left| \int_0^1 \int_0^1 \langle F(x, y), u \rangle_+ dx dy - I_{p^2}(u, F) \right| \leq \frac{K}{p^2}, \quad \forall u \in \mathcal{S}$$

où

$$K = \frac{mk}{\sqrt{2}} + M \quad \blacksquare$$

Corollaire 4. 26 Soient $F : D \longrightarrow \mathbb{R}^n$ et $Z(F)$ le zonoïde associé. Le nombre q de segments suffisant pour créer un zonotope Z_q qui approche $Z(F)$ à ε près, est en $O(\frac{1}{\varepsilon})$.

Remarques :

- 1- Même si l'erreur est en $O(\frac{1}{p^2})$ dans le théorème 4. 25, il faut noter que le zonotope Z_{p^2} est la somme de $O(p^2)$ vecteurs: l'ordre de convergence est donc de 1 par rapport au nombre de vecteurs. La convergence est donc plus faible que dans le cas où F est une fonction d'une seule variable. Mais, ce résultat reste meilleur que celui prévu par l'approximation des zonoïdes dans la théorie générale (cf le théorème 4. 10).
- 2- Le résultat reste vrai quand on passe à un domaine rectangulaire quelconque.
- 3- En utilisant la méthode d'intégration des rectangles (avec point milieu), on obtient la séquence de zonotopes Z_{p^2} somme des segments:

$$[0, h^2 F(x_i, y_j)], \quad i = 1, \dots, p \text{ et } j = 1, \dots, p,$$

où

$$x_i = \frac{h}{2}(2i - 1),$$

$$y_j = \frac{h}{2}(2j - 1).$$

De la même manière, on peut démontrer que Z_{p^2} converge vers $Z(F)$ quand $p \rightarrow \infty$ et que l'erreur est en $O(\frac{1}{p^2})$. $\rightarrow$

5. Exemples

Une illustration de la convergence d'ordre 1 par rapport au nombre de vecteurs est donnée dans les exemples qui suivent.

Exemple 1 (Sphère)

F une fonction définie sur $[-\pi, \pi] \times [-\pi, \pi]$ à valeurs dans $\mathbb{R}^3$ par:

$$F(x, y) = \begin{pmatrix} \frac{\cos(x)\cos(y)}{4} \\ \frac{\sin(x)\cos(y)}{4} \\ \frac{\sin y(y)}{2\pi} \end{pmatrix}.$$

Illustration de la convergence d'ordre 1 par rapport au nombre de segments:

Nbre Seg.	$haus(Z_p, Z(F))$	ordre
8	0.233701	
32	0.053029	1.069900
128	0.012951	1.016877
512	0.003219	1.004183
2048	0.000804	1.001043
8192	0.000201	1.000260
32768	0.000050	1.000061

Exemple 2

$D = [0, 2] \times [-\pi, \pi]$

$$F(x, y) = \begin{pmatrix} x \cos(y) \\ x \sin(y) \\ 1 \end{pmatrix}.$$

Illustration de la convergence d'ordre 1 par rapport au nombre de segments:

Nbre Seg.	$haus(Z_p, Z(F))$	ordre
8	0.233697	
32	0.053026	1.069935
128	0.012947	1.017016
512	0.003216	1.004738
2048	0.000800	1.003265
8192	0.000198	1.009217
32768	0.000047	1.037043

Exemple 3

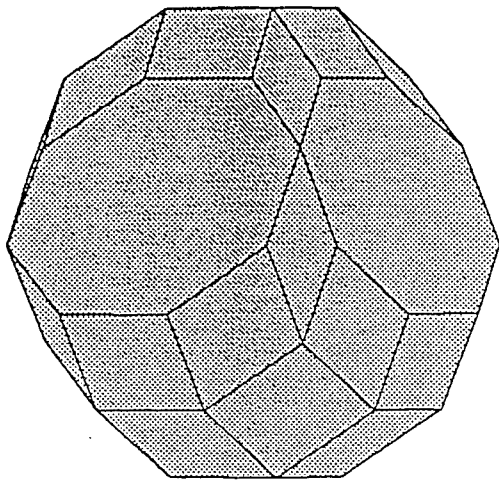
$$D = [0, 3] \times [0, 3]$$

$$F(x, y) = \begin{pmatrix} x^2 - y^2 \\ xy \\ x + y \end{pmatrix}.$$

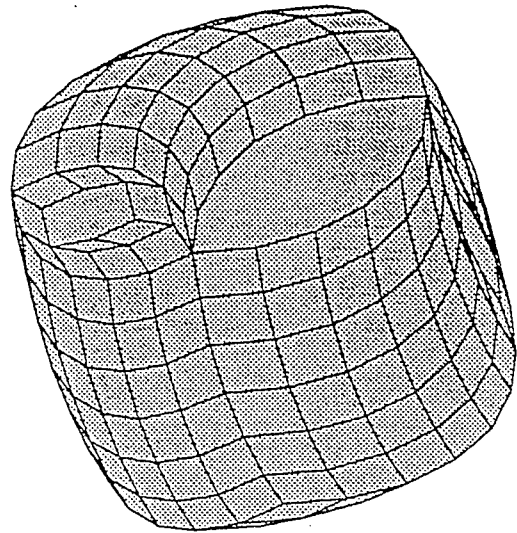
Illustration de la convergence d'ordre 1 par rapport au nombre de segments:

Nbre Seg.	$haus(Z_p, Z(F))$	ordre
8	1.476539	
32	0.369117	1.000035
128	0.092262	1.000139
512	0.023048	1.000554
2048	0.005744	1.002222
8192	0.001418	1.008956
32768	0.000337	1.036978

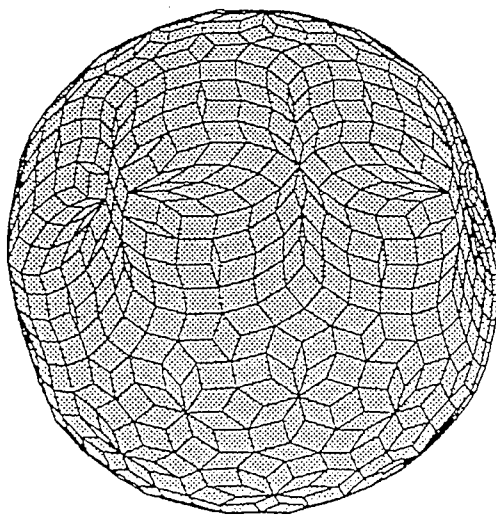
Dans les figures (9), (10) et (11), nous donnons quelques exemples de zonotopes qui convergent vers les zonoïdes définis ci-dessus.



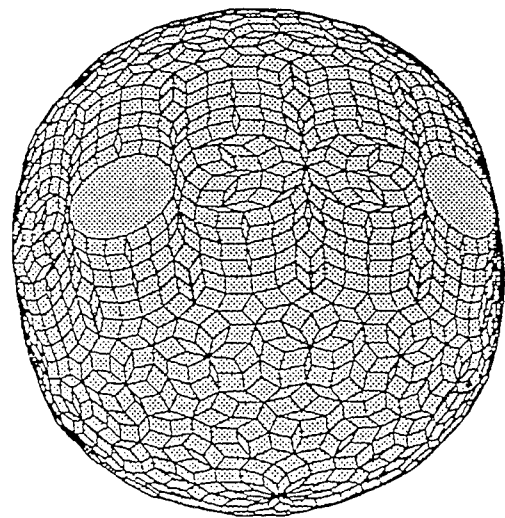
4x2 vecteurs



3x6 vecteurs

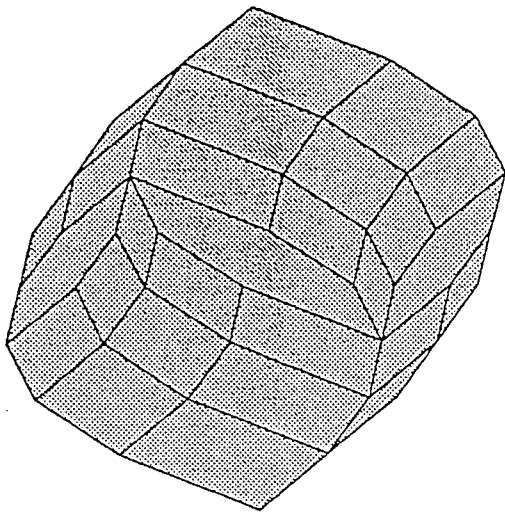


8x4 vecteurs

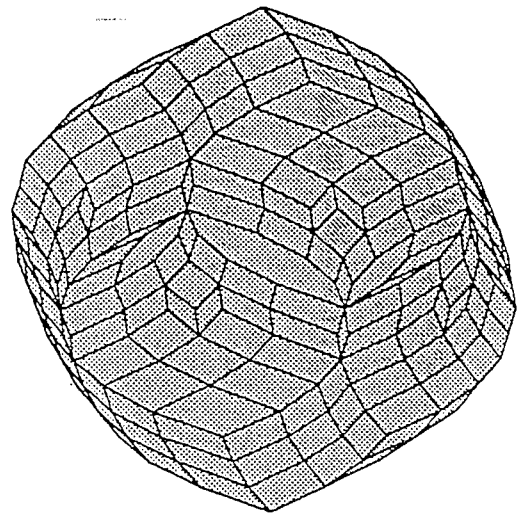


12x4 vecteurs

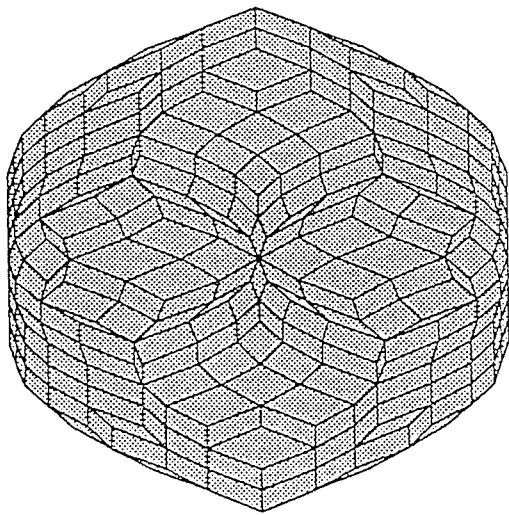
FIGURE .9. Exemple 1, convergence des zonotopes vers la boule unité dans $\mathbb{R}^3$.



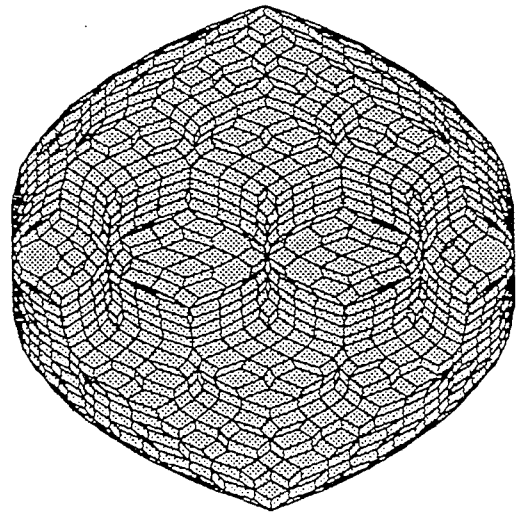
3x3 vecteurs



4x5 vecteurs

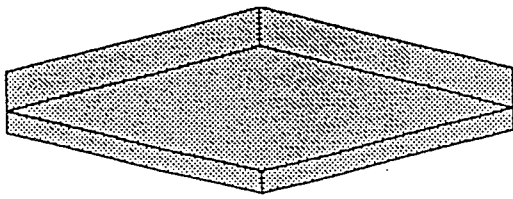


4x6 vecteurs

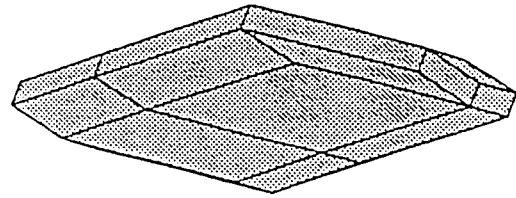


5x10 vecteurs

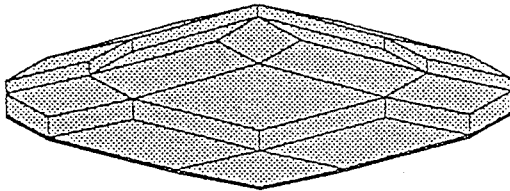
FIGURE .10. Exemple 2, convergence de Z_{p^2} .



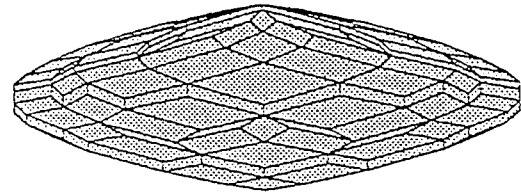
2x2 vecteurs



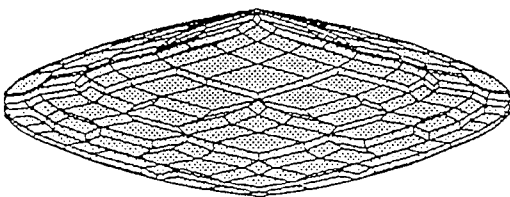
3x2 vecteurs



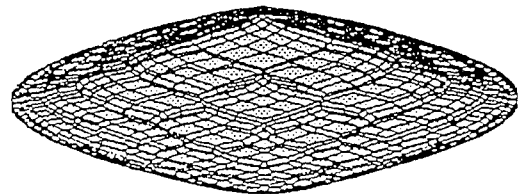
3x3 vecteurs



4x4 vecteurs



5x5 vecteurs



7x7 vecteurs

FIGURE .11. Exemple 3, convergence de Z_{p^2} .

Chapitre 5

Représentation des tables ASTM-IP 53/54 par B-splines régularisées

Introduction

L'objectif de ce travail est de proposer une représentation utilisable sur ordinateur des tables ASTM-IP 53/54. Ces tables sont définies de la manière suivante: La table 53 donne la valeur de la densité à 15 degrés en fonction de la densité observée et de la température. La table 54 fournit une valeur appelée "facteur de réduction du volume" en fonction de la densité à 15 degrés et de la température. L'importance de ces tables est grande, elles sont utilisées dans le calcul des taxes douanières sur les produits pétroliers.

Deux méthodes ont déjà été proposées; nous les appellerons : "méthode IP" et "méthode Chevallier". Le but de ce chapitre est de présenter une troisième approche basée sur les B-splines bi-dimensionnelles. Il s'agit d'une méthode générale qui permet de représenter globalement une fonction de deux variables dont les valeurs en certains points sont données (éventuellement avec erreurs). Elle a permis dans le cas présent de représenter la totalité des données avec un nombre d'erreurs nettement inférieur à celui obtenu pour les méthodes précédentes. Elle nécessite par contre le stockage d'un plus grand nombre de constantes. De façon annexe, la méthode a détecté certaines incohérences dans la table elle-même.

Chacune des 2 tables est divisée en deux parties, le but de ce travail est d'obtenir pour chacune des 4 parties une expression qui, arrondie à 3, 4 ou 5 décimales selon les cas, conduisent à un nombre d'erreurs aussi réduit que possible par rapport aux deux tables.

Il existe entre les deux tables une relation implicite. Cette relation permet théoriquement de déduire la table 53 de la table 54 en faisant une itération de type point fixe. Cette façon de procéder a été suivie aussi bien dans la méthode IP que dans la méthode de Chevallier. L'inconvénient est que cela ne permet pas d'atteindre toute la table 53. Il faut en effet que l'itération converge sans sortir de la table 54. Par exemple dans la partie II, on ne peut atteindre que 6685 données environ sur les 9422. Dans notre approche par B-splines, nous donnerons

une expression mathématique indépendante pour chacune des 4 parties et nous n'utiliserons la relation implicite qu'à des fins de vérification.

Enfin il faut souligner que les nouvelles formules qui sont proposées sont basées sur des considérations purement numériques et ne prennent pas du tout en compte les aspects physiques, contrairement à la méthode IP par exemple.

Dans la partie 1, On donne quelques rappels sur le problème d'approximation par produit tensoriel de B-splines.

Les résultats numériques sont exposés dans la partie 2, en considérant successivement le cas des tables 54 et 53 et en séparant pour chacune les parties I et II. (Une étude globale de la table 53 a été tentée mais elle ne conduit pas à un résultat meilleur).

Dans la partie 3, on examine d'un point de vue statistique le raccordement à $D_{15} = 0.600$ entre les parties I et II de la table 54.

Enfin dans la partie 4, on examine, pour les trois méthodes, le nombre d'erreurs que l'on commet si l'on déduit la formule de représentation de la table 53 de celle de la table 54 au moyen de la relation implicite.

1. B-splines régularisées, problème variationnel:

1. 1. Lissage unidimensionnel par B-splines régularisées:

Soit $[a, b]$ un intervalle de $\mathbb{R}$, et soient (t_k, z_k) ; $k = 1, \dots, ND$ des couples de points de $[a, b] \times \mathbb{R}$ où chaque z_k est la valeur, entachée d'erreur, prise par une fonction inconnue f au point t_k :

$$z_k = f(t_k) + \varepsilon_k$$

où ε_k est une erreur due à une incertitude sur les mesures ou un arrondi.

Au moyen d'un lissage par B-splines régularisées des données z_k , on approche la fonction f sur $[a, b]$ par une fonction s , élément d'un certain espace de fonctions S que nous allons définir après.

Soit:

$$a = u_0 < u_1 < \dots < u_n = b$$

une subdivision régulière de $[a, b]$ de pas $h = (b-a)/n$.

Soit S l'espace des fonctions polynomiales par morceaux de degré d sur chaque intervalle $[u_i, u_{i+1}]$, $i = 0, \dots, n-1$, et de classe C^{d-1} . C'est un espace de dimension $n + d$.

On prend sur S la base, $(B_j)_{j=-d, \dots, n-1}$, des splines de degré d . Le support de chaque spline de base B_j est $[u_j, u_{j+d+1}]$ où j varie de $-d$ à $n-1$. (On a introduit ainsi des noeuds supplémentaires fictifs $u_{-d}, \dots, u_{-1}$ et $u_{n+1}, \dots, u_{n+d}$).

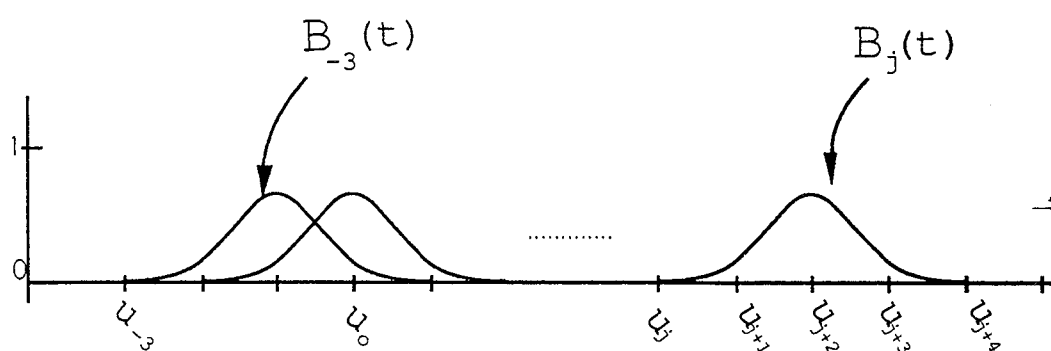


FIGURE .1. Splines de base ($d = 3$).

Toute fonction s de S s'écrit:

$$s(t) = \sum_{j=-d}^{n-1} \alpha_j B_j(t). \quad (5.1)$$

Le lissage par B-splines régularisées se pose donc comme un problème de minimisation au sens des moindres carrés régularisés sur l'espace de fonctions S . La fonction σ cherchée est solution du problème suivant (cf [Laurent87]):

$$\min_{s \in S} \quad \tau \int_a^b (s^{(d)}(t))^2 dt + \sum_{i=1}^{ND} (s(t_i) - z_i)^2, \quad (P)$$

où τ est un réel positif appelé paramètre de régularisation.

En remplaçant s par son expression en (5.1), le problème (P) devient un problème de minimisation sur $\mathbb{R}^{n+d}$ s'écrivant:

$$\min_{\alpha \in \mathbb{R}^{n+d}} \quad \tau \int_a^b \left(\sum_{j=-d}^{n-1} \alpha_j B_j^{(d)}(t) \right)^2 dt + \sum_{i=1}^{ND} \left(\sum_{j=-d}^{n-1} \alpha_j B_j(t_i) - z_i \right)^2 \quad (P')$$

Les conditions d'optimalité s'obtiennent alors en dérivant par rapport aux α_k et s'écrivent:

$$\tau \int_a^b \left(\sum_{j=-d}^{n-1} \alpha_j B_j^{(d)}(t) \right) B_k^{(d)}(t) dt + \sum_{i=1}^{ND} \left(\sum_{j=-d}^{n-1} \alpha_j B_j(t_i) - z_i \right) B_k(t_i) = 0, \quad (5.2)$$

pour $k = -d, \dots, n-1$.

Les α_j , $j = -d, \dots, n-1$ sont donc solution du système, à $(n+d)$ équations, suivant:

$$\sum_{j=-d}^{n-1} \alpha_j \left(\tau \int_a^b B_j^{(d)}(t) B_k^{(d)}(t) dt + \sum_{i=1}^{ND} B_j(t_i) B_k(t_i) \right) = \sum_{i=1}^{ND} z_i B_k(t_i), \quad (5.3)$$

$k = -d, \dots, n-1$.

En posant Ω la matrice de taille $(n+d) \times (n+d)$ définie par:

$$\Omega_{ij} = \int_a^b B_i^{(d)} B_j^{(d)} dt,$$

et B la matrice de taille $(n+d) \times ND$ définie par

$$B_{ij} = B_i(t_j),$$

et en remplaçant dans (5.3), on obtient le système:

$$(\tau \Omega + B^t B) \alpha = B^t z, \quad (5.4)$$

où le coefficient τ est à choisir pour doser correctement l'effet du lissage.

Pour $t \in [a, b]$, soit i l'entier vérifiant $t \in [u_i, u_{i+1}]$. Les valeurs de j pour lesquelles $B_j(t)$ est non nul sont $j = i-d, \dots, i$. Pour calculer $B_j(t)$, on utilise la relation de

réurrence reliant la B-spline de degré d et celles qui sont de degré $d - 1$. En effet, si B_i^m désigne la B-spline de degré m qui a pour support l'intervalle $[u_i, u_{i+m+1}]$ alors la relation de récurrence s'étabit comme suit:

$$B_i^m(t) = \frac{t - u_i}{u_{i+m-1} - u_i} B_i^{m-1}(t) + \frac{u_{i+m} - t}{u_{i+m} - u_{i+1}} B_{i+1}^{m-1}(t). \quad (5.5)$$

$$B_i^0(x) = \begin{cases} 1 & \text{si } x \in [u_i, u_{i+1}] \\ 0 & \text{sinon} \end{cases}.$$

La formation de la matrice et du second membre correspondant au système d'équations (5.4) nécessite l'évaluation, pour chaque $t_k \in [u_i, u_{i+1}]$, des quantités $B_j(t_k)$ pour tout entier j ne les rendant pas nulles.

Dans la pratique, d prend les valeurs 1, 2 ou 3; si p et q sont deux réels qui représentent la localisation de t_k dans l'intervalle $[u_i, u_{i+1}]$:

$$p = \frac{t_k - u_i}{h} \text{ et } q = \frac{u_{i+1} - t_k}{h}$$

en utilisant (5.5), l'évaluation des B_j se fait par:

pour $d = 1$

$$B_{i-1}(t_k) = q$$

$$B_i(t_k) = p$$

pour $d = 2$

$$B_{i-2}(t_k) = q^2/2$$

$$B_{i-1}(t_k) = (1 + 2pq)/2$$

$$B_i(t_k) = p^2/2$$

pour $d = 3$

$$B_{i-3}(t_k) = q^3/6$$

$$B_{i-2}(t_k) = (3q(pq + 1) + 1)/6$$

$$B_{i-1}(t_k) = (3p(pq + 1) + 1)/6$$

$$B_i(t_k) = p^3/6$$

Par ailleurs $B_j^{(d)}(t)$ est constant sur l'intervalle $[u_j, u_{j+1}]$. Si l'on pose :

$B_{i-l}^{(d)}(t) = c_l$, on a:

pour $d = 1$:

$$c_0 = 1/h, c_1 = -1/h$$

pour $d = 2$:

$$c_0 = 1/h^2, c_1 = -2/h^2, c_2 = 1/h^2$$

pour $d = 3$:

$$c_0 = 1/h^3, c_1 = -3/h^3, c_2 = 3/h^3, c_3 = 1/h^3$$

les autres constantes étant nulles.

1. 2. Lissage bi-dimensionnel par produit tensoriel de B-splines régularisées

Soit $[a, b] \times [c, d]$ un domaine rectangulaire de $\mathbb{R} \times \mathbb{R}$.

Soient $(x_k, y_k)_{k=1, \dots, ND}$ des points de $[a, b] \times [c, d]$ et z_k des réels représentant les valeurs, entachées d'erreurs, prises par une fonction f aux points (x_k, y_k) :

$$z_k = f(x_k, y_k) + \varepsilon_k.$$

Soient $a = u_0 < u_1 < \dots < u_{N_x} = b$ une subdivision régulière de $[a, b]$ et $c = v_0 < v_1 < \dots < v_{N_y} = d$ une subdivision régulière de $[c, d]$.

Notons S_x (resp. S_y) l'espace des fonctions polynomiales de degré d sur chaque intervalle $[u_i, u_{i+1}]$, $i = 0, \dots, N_x - 1$ (resp. $[v_j, v_{j+1}]$, $j = 0, \dots, N_y - 1$) et de classe C^{d-1} sur $[a, b]$ (resp. sur $[c, d]$).

Notons S l'espace vectoriel des fonctions définies par le produit tensoriel des éléments de S_x et S_y . S est un espace vectoriel de dimension $(N_x + d)(N_y + d)$.

Si on prend sur S_x (resp. S_y), la base B_{1i} , $i = -d, \dots, N_x - 1$, (resp. B_{2j} , $j = -d, \dots, N_y - 1$) des splines de degré d ; alors $B_{1i}B_{2j}$, $i = -d, \dots, N_x - 1$, $j = -d, \dots, N_y - 1$ constitue une base de S .

Le lissage bi-dimensionnel par produit tensoriel de B-splines régularisées se pose comme un problème de minimisation au sens des moindres carrés régularisés sur l'espace S .

Soit σ la fonction de S approchant f sur $[a, b] \times [c, d]$, alors σ est solution du problème de minimisation:

$$\min_{s \in S} \tau L(s) + \sum_{k=1}^{ND} (l_k(s) - z_k)^2 \quad (P_1)$$

où l_k est l'opérateur linéaire qui à $s \in S$ associe $s(x_k, y_k) \in \mathbb{R}$, et L l'opérateur relatif au lissage défini par:

$$L(s) = \int_a^b \int_c^d (D_x^d D_y^d s)^2 dx dy$$

(où D_x^d est la dérivée partielle d'ordre d par rapport à x).

• **Evaluation du terme de lissage:**

$$\begin{aligned}
L(s) &= \int_a^b \int_c^d \left(\sum_{j,k} \alpha_{jk} D_x^d D_y^d B_{x_j}(x) B_{2_j}(y) \right)^2 dx dy \\
&= \int_a^b \int_c^d \left(\sum_{j,k} \alpha_{jk} B_{1_i}^{(d)}(x) B_{2_j}^{(d)}(y) \right)^2 dx dy \\
&= \sum_{k,k'=-d}^{N_y-1} \sum_{j,j'=-d}^{N_x-1} \alpha_{jk} \alpha_{j'k'} \int_a^b \int_c^d B_{1_i}^{(d)}(x) B_{2_j}^{(d)}(y) B_{x_{i'}}^{(d)}(x) B_{y_{j'}}^{(d)}(y) dx dy \\
&= \sum_{k,k'=-d}^{N_y-1} \sum_{j,j'=-d}^{N_x-1} \alpha_{jk} \alpha_{j'k'} \int_a^b B_{1_i}^{(d)}(x) B_{x_{i'}}^{(d)}(x) dx \int_c^d B_{2_j}^{(d)}(y) B_{y_{j'}}^{(d)}(y) dy
\end{aligned}$$

En posant $N_1 = N_x + d$ et $N_2 = N_y + d$, on définit les matrices:

$$[\Omega_x]_{jj'} = \int_a^b B_{1_i}^{(d)}(x) B_{x_{i'}}^{(d)}(x) dx \quad \text{matrice } N_1 \times N_1$$

$$[\Omega_y]_{kk'} = \int_c^d B_{2_j}^{(d)}(y) B_{y_{j'}}^{(d)}(y) dy \quad \text{matrice } N_2 \times N_2$$

Ω_x et Ω_y sont deux matrices symétriques.

$L(s)$ peut s'écrire alors sous la forme:

$$L(s) = \langle \Omega_x \alpha, \alpha \Omega_y \rangle$$

$$= \langle \alpha, \Omega_x \alpha \Omega_y \rangle$$

où $\alpha = (\alpha_{ij})$ est une matrice $N_1 \times N_2$ (le produit scalaire $\langle \cdot, \cdot \rangle$ représente la somme des produits des termes correspondants de chaque matrice).

Soit Ω la matrice qui à α , pris comme vecteur de $\mathbb{R}^{N_1 \times N_2}$, associe le vecteur $\Omega_x \alpha \Omega_y$ de $\mathbb{R}^{N_1 \times N_2}$. Ω est le produit tensoriel des matrices Ω_x et Ω_y ; c'est une matrice carré d'ordre $(N_1 \times N_2)$, symétrique. L'opérateur de lissage devient:

$$L(s) = \langle \alpha, \Omega \alpha \rangle. \quad (5.6)$$

• Evaluation du terme d'approximation

Le terme d'approximation est défini par:

$$\sum_{k=1}^{ND} (l_k(s) - z_k)^2$$

On a:

$$\begin{aligned} \sum_{k=1}^{ND} (l_k(s) - z_k)^2 &= \sum_{k=1}^{ND} \left(\sum_{i,j} \alpha_{ij} l_k(B_{1i}, B_{2j}) - z_k \right)^2 \\ &= \sum_{k=1}^{ND} \left(\sum_{i,j} \alpha_{ij} B_{1i}(x_i) B_{2j}(y_j) - z_k \right)^2. \end{aligned}$$

Soit B la matrice $ND \times (N_1 \cdot N_2)$ définie par:

$$B_{k,[i,j]} = l_k(B_{1i}, B_{2j}).$$

En considérant α comme un vecteur de $\mathbb{R}^{N_1 \times N_2}$, le terme d'approximation s'écrit alors:

$$\|B\alpha - Z\|^2. \quad (5.7)$$

Z est le vecteur de $\mathbb{R}^{ND}$ dont les composantes sont formées par les z_k .

Le problème de minimisation (P_1) devient alors:

$$\min_{\alpha \in \mathbb{R}^{N_1 \times N_2}} \tau \langle \alpha, \Omega \alpha \rangle + \|B\alpha - Z\|^2 \quad (P'_1)$$

Les conditions d'optimalité, en dérivant par rapport à α , conduisent au système suivant:

$$(\tau \Omega + B^t B) \alpha = B^t Z.$$

qui est un système symétrique d'équations, de taille $N_1 \times N_2$.

2. Traitement Numérique

2. 1. Définitions et notations

On utilisera les paramètres suivants:

- N_x : le nombre d'intervalles utilisés pour la température,
- N_y : le nombre d'intervalles utilisés pour la densité à 15 degrés,
- d : degré de la spline,
- τ : paramètre de régularisation,

et on donnera les résultats suivants:

- Le nombre total des erreurs $NErr$. Pour évaluer le nombre d'erreurs, seules les erreurs supérieures ou égales à 0.5 sur le dernier digit (c'est à dire ayant conduit à une valeur arrondie différente de celle de la table) sont comptées.
- Le nombre d'erreurs dont la valeur est comprises entre 0.5 et 0.6 sur le dernier digit et leur pourcentage par rapport au nombre total d'erreurs.
- Le nombre d'erreurs qui sont supérieures à 1, (i.e celles pour lesquelles la valeur calculée arrondie diffère de la table de plus d'une unité sur le dernier digit).
- L'erreur maximale.
- L'erreur quadratique moyenne qui est définie par :

$$EQM = \sqrt{\frac{\sum_{k=1}^{ND} (e_k)^2}{ND}}$$

où e_k est l'erreur sur la $k^{\text{ème}}$ donnée .

Pour comparer correctement l'erreur quadratique moyenne, il faut remarquer que l'erreur quadratique moyenne dans le cas idéal où on arrondit une fonction connue serait égale à $\frac{1}{2\sqrt{3}} \approx 0.29$ (c'est l'écart-type d'une variable aléatoire uniforme entre

-0.5 et 0.5, c'est à dire $\sqrt{\int_{-0.5}^{0.5} x^2 dx}$).

- L'erreur algébrique moyenne définie par :

$$Emoy = \frac{\sum_{k=1}^{ND} e_k}{ND}$$

La valeur de la moyenne algébrique des erreurs nous permet d'évaluer si les erreurs issues d'une méthode sont bien centrées sur 0.

Pour le calcul de l'erreur algébrique moyenne et de l'erreur quadratique moyenne on a fait intervenir les erreurs sur toutes les données.

- Le facteur de compression défini par:

$$Comp = \frac{\text{nombre de coefficients utilisés}}{\text{nombre de données total}} \times 100.$$

Il représente le taux de compression pour chaque jeu de N_x , N_y et d .

- L'histogramme des erreurs pour montrer la distribution des erreurs autour de la valeur 0.5. Cet histogramme représente les valeurs 0.50, 0.52, ... et -0.50, -0.52, ... (Pour des raisons pratiques d'échelle, les erreurs supérieures à 1.1 ne sont pas présentées).

2. 2. Tables ASTM-IP 53/54

La **table 53** exprime la densité à 15 degrés D_{15} en fonction de la densité observée et de la température T :

$$D_{15} = F(D_o, T).$$

Le domaine des données est défini par les valeurs de la densité observée variant entre 0.420 et 0.689 et une température T entre -46 et 60. Le pas sur la température est 0.5 alors que celui sur la densité observée est soit 0.005 soit 0.001 selon que D_o est inférieure ou supérieure à 0.600, ce qui rend la grille des données non régulière (voir graphique des données). De plus, il y a un manque de données dans certaines régions du domaine formant ainsi des trous plus ou moins importants.

Le nombre total des données est 13722. La précision n'est pas la même pour toutes les données: celles qui correspondent aux valeurs de D_o inférieures à 0.600 sont fournies avec trois chiffres et les autres avec 4 chiffres significatifs. En tenant compte de toutes ces remarques, on peut considérer que le domaine est une union de deux parties:

- la partie I : définie par les valeurs de D_o entre 0.420 et 0.600, le pas étant 0.005 et les valeurs de la température T entre -46 et 60,
- la partie II : définie par D_o entre 0.600 et 0.689, le pas étant 0.001 et les valeurs de T variant entre -25 et 50.

Nous avons traité le problème de deux manières: la première consiste à traiter les deux régions séparément en considérant respectivement les rectangles $[-46, 60] \times [0.420, 0.600]$ et $[-25, 50] \times [0.600, 0.689]$; la deuxième consiste à faire un traitement en considérant le domaine global définie par le rectangle $[-46, 60] \times [0.420, 0.689]$.

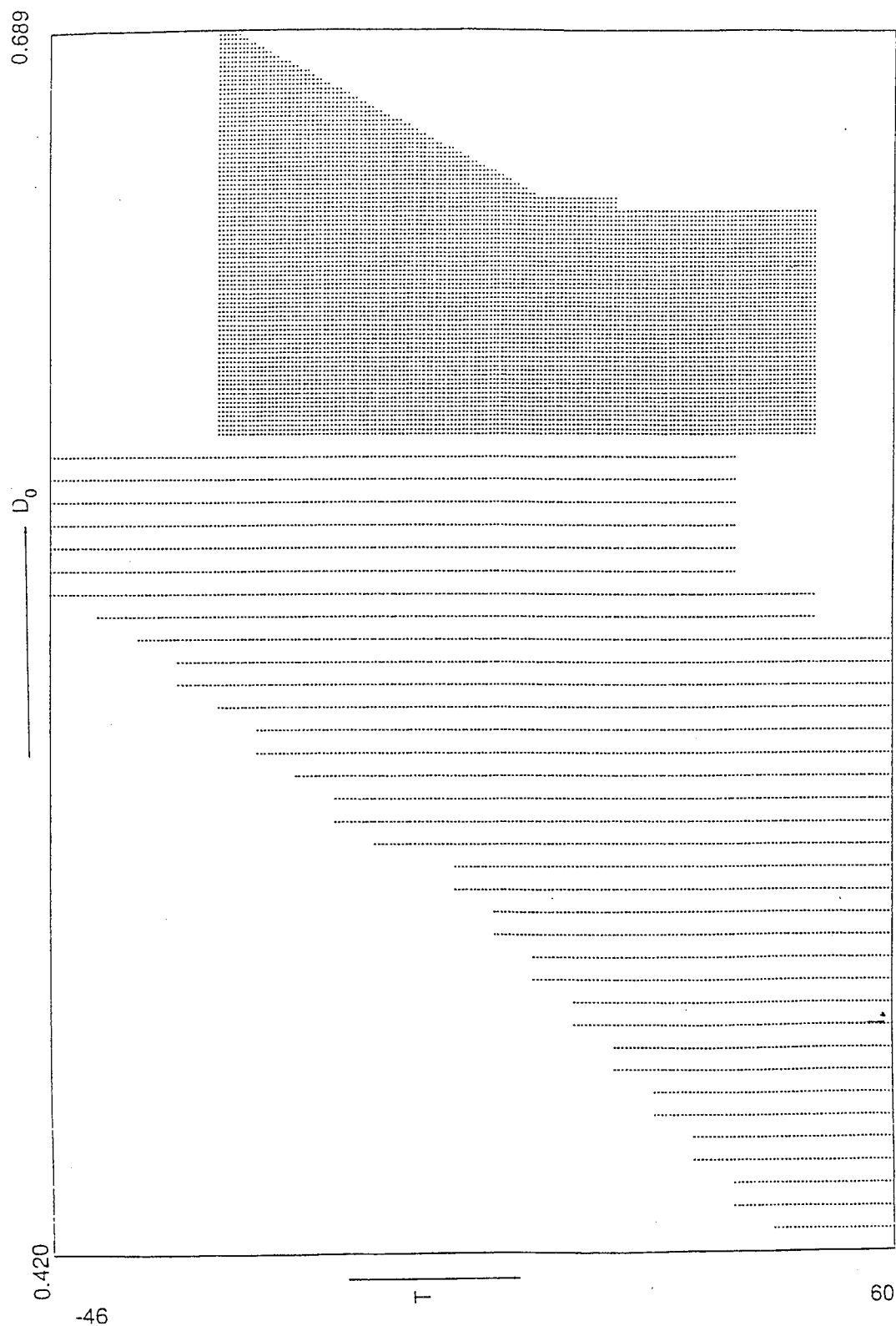
La **table 54** exprime le "Volume correction factor" (VCF) en fonction de la densité à 15 degrés D_{15} et de la température T :

$$VCF = G(D_{15}, T).$$

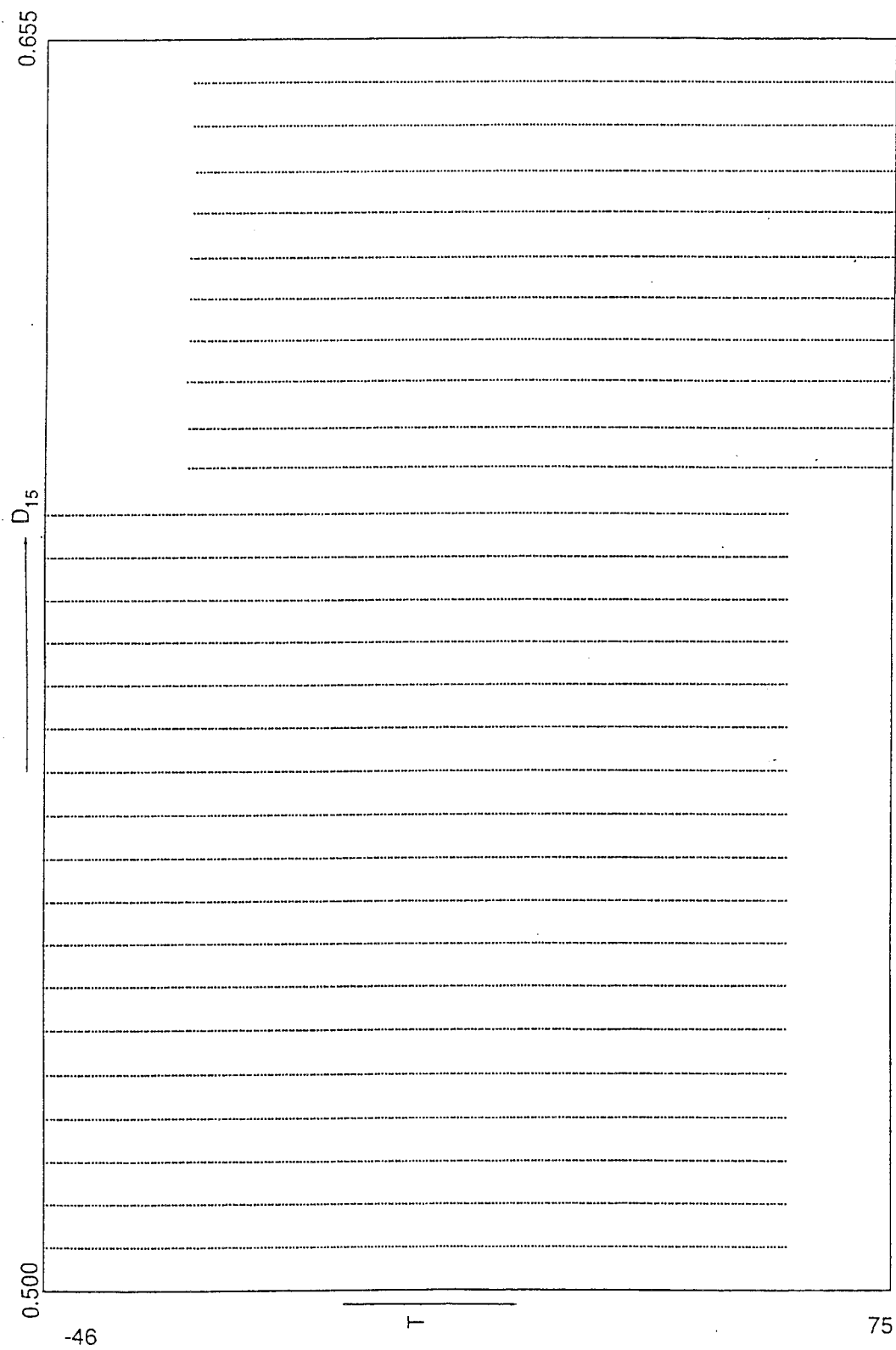
Elle comporte au total 6672 données. Cette table comprend également deux parties : la partie I (4260 données) correspondant à $D_{15} \leq 0.600$ est donnée avec 4 décimales et la partie II (2412 données) correspondant à $D_{15} \geq 0.600$ est donnée avec 5 décimales.

On remarque que pour $D_{15} = 0.600$, la valeur de la VCF est donnée soit avec 4 décimales (partie I) soit avec 5 décimales (partie II).

Pour chacune des deux parties, les couples (D_{15}, T) forment une grille rectangulaire régulière, mais ces deux rectangles sont décalés l'un par rapport à l'autre. Les données dans les deux parties sont disposées sur une grille régulière et ne présentent pas de trous.



Distribution des données (Table 53)



Distribution des données (TABLE 54).

2. 3. Partie I de la table 54:

2.3.1. Méthode IP:

La méthode IP pour le traitement de la partie I, calcule la VCF par approximation des volumes V_T (volume à la température T) et V_{15} (volume à 15 degrés).

La VCF est alors trouvée par la relation

$$VCF = \frac{V_{15}}{V_T}.$$

Pour calculer V_{15} et V_T la méthode utilise 18 constantes.

Le nombre d'erreurs trouvées en utilisant la méthode IP est 1551, ce nombre représentant 36.4 % de l'ensemble des données (4260).

Le nombre d'erreurs comprises entre 0.5 et 0.6 est 345 (22.2 % du nombre total d'erreurs).

L'erreur maximale est 3,56.

Le nombre d'erreurs supérieurs à 1 est 236.

L'erreur quadratique moyenne est 0.59.

L'erreur moyenne est 0.123.

2.3.2. Méthode de Chevallier:

La méthode de Chevallier consiste à faire une approche de la partie I de la table 54 par la fonction suivante:

$$VCF = \exp(-\alpha \cdot \Delta T (1 + A \cdot \alpha \Delta T + B \alpha^2 \Delta T^2 + C \alpha^3 \Delta T^3) + (D \cdot \Delta T^4 + E \cdot \Delta T^2)(F - (600 - D_{15})^\beta))$$

où

$$\alpha = K_0 + \frac{K_1}{D_{15}} + \frac{K_2}{D_{15}^2} + \frac{K_3}{D_{15}^3}$$

et

$$\Delta T = T - 15$$

$$K_0, K_1, K_2, K_3 \text{ et } A, B, C, D, E, F, \beta$$

sont des constantes données (voir rapport de Chevallier Aout 87).

Le nombre d'erreurs pour cette méthode est 1397.

Le nombre d'erreurs comprises entre 0.5 et 0.6 est 357, représentant 25.6 % du total des erreurs.

Le nombre d'erreurs supérieures à 1 est 325.

L'erreur maximale est -2.245.

L'erreur quadratique moyenne est 0.56.

L'erreur moyenne est 0.125.

2.3.3. Méthode B-spline

- **Première stratégie**

Dans cette méthode, on représente la partie I de la table 54, en utilisant un nombre de splines assez important, et donc un nombre important de coefficients pour le calcul de la fonction approchée. Il faut souligner ici que ce nombre reste très petit devant le nombre de données comme le montre les valeurs du facteur de compression (voir exemple).

- **Deuxième stratégie:**

Dans cette méthode, on traite le domaine (partie I) de manière à n'avoir que quelques constantes à stocker (une vingtaine ou une trentaine) et on regarde comment se comportent les erreurs par rapport à la stratégie I et aux autres méthodes qui n'utilisent pas beaucoup de constantes (voir exemple). Ainsi:

N_x prendra la valeur 1 ou 2,
 N_y prendra la valeur 1,2,3,5 ,
 d sera égal à 3.

Exemple (stratégie 1)Nombre de données $ND =$

4260

$$N_x = 5,$$

$$N_y = 17,$$

$$d = 3,$$

$$\tau = 10^{-7}.$$

Nombre d'erreurs total, $NErr =$

504

Nombre d'erreurs comprises entre 0.5 et 0.6

278

ce qui représente parmi l'ensemble des erreurs

55.2 %

Nombre d'erreurs supérieures ou égales 1

0

Erreur maximale

0.971

Erreur quadratique moyenne

0.33

Moyenne des erreurs

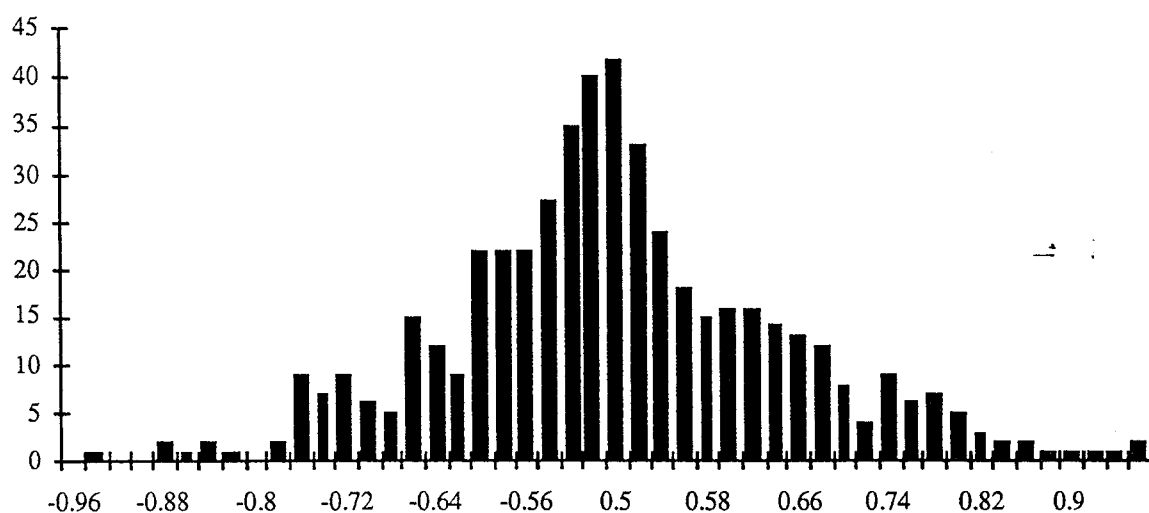
 10^{-12}

Nombre de coefficients pour le calcul de la spline

160

Taux de compression, $comp =$

3.76 %

FIGURE .2. *Histogramme des erreurs (stratégie 1)*

Dans la figure 2, on donne l'histogramme représentant la distribution des erreurs autour de -0.5 et 0.5 (sur le dernier digit).

Exemple (stratégie 2)Nombre de données, $ND =$

4260

$$N_x = 2,$$

$$N_y = 3,$$

$$d = 3,$$

$$\tau = 10^{-5}.$$

Nombre d'erreurs total, $NErr =$

822

Nombre d'erreurs comprises entre 0.5 et 0.6

278

ce qui représente parmi l'ensemble des erreurs

41 %

Nombre d'erreurs supérieures ou égales 1

19

Erreur maximale

-1.45

Erreur quadratique moyenne

0.38

Moyenne des erreurs

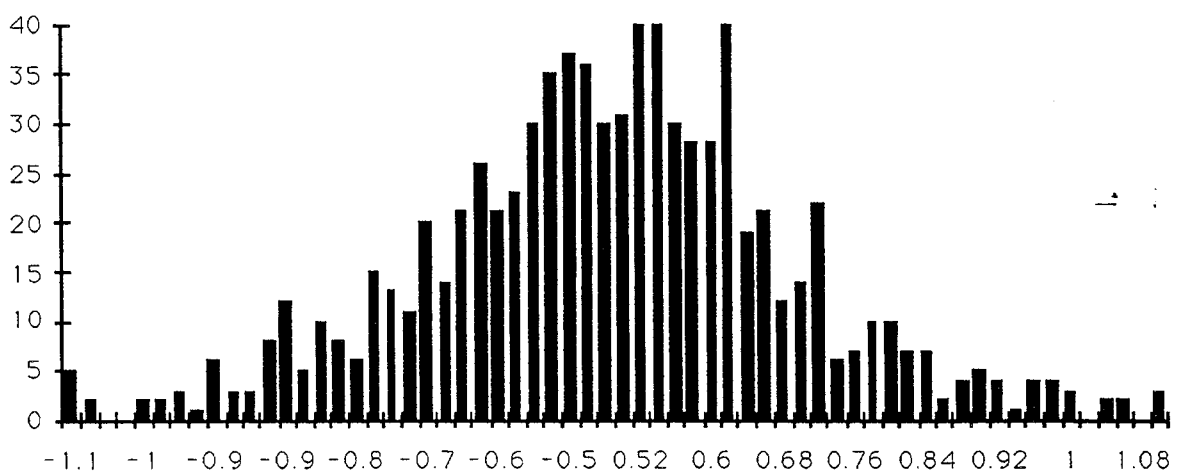
 -1.510^{-4}

Nombre de coefficients pour le calcul de la spline

30

Taux de compression, $comp =$

0.70 %

FIGURE .3. *Histogramme des erreurs (stratégie 2)*

Dans la figure 3, on donne l'histogramme représentant la distribution des erreurs autour de -0.5 et 0.5 (sur le dernier digit).

2.3.4. Conclusions (table 54, partie I):

1- Pour la première stratégie consistant à minimiser le nombre d'erreurs en utilisant un nombre relativement important de constantes, on recommande les valeurs $N_x = 5$ et $N_y = 17$ qui conduisent à 160 constantes. La comparaison avec les deux autres méthodes s'établit comme suit:

	nombre d'erreurs	nombre d'erreurs entre 0.5 et 0.6	nombre d'erreurs > 1	EQM-0.29
Méthode IP	1551	345 (22.2 %)	236	0.30
Méthode de Chevallier	1397	357 (25.6 %)	325	0.27
Méthode B-splines	504	268 (55.2 %)	0	0.04

2- Pour la deuxième stratégie où l'on s'impose de n'utiliser qu'un petit nombre de constantes (du même ordre que pour les méthodes IP et Chevallier), on recommande les valeurs $N_x = 2$ et $N_y = 3$ qui conduisent à 30 constantes. La comparaison entre méthodes s'établit alors comme suit :

	nombre d'erreurs	nombre d'erreurs entre 0.5 et 0.6	nombre d'erreurs > 1	EQM-0.29
Méthode IP	1551	345 (22.2%)	236	0.30
Méthode de Chevallier	1397	357 (25.6 %)	325	0.27
Méthode B-splines	822	337 (41%)	19	0.09

Si l'on observe la répartition des erreurs pour la partie I de la table 54, on remarque une zone le long de la ligne $T = -25$ qui comporte un nombre anormalement élevé d'erreurs. Ceci a été noté dans tous les exemples quelles que soient les valeurs des paramètres N_x, N_y ou τ . Il s'agit donc très probablement d'un défaut de raccordement entre les zones de la table (≤ -25 et ≥ -25) dont les origines devaient être différentes. Ce phénomène devrait être recherché dans l'historique de la fabrication de la table.

2. 4. Partie II de la table 54:

Dans cette partie la température varie entre -25 et 75 et la densité à 15 degrés varie entre 0.600 et 0.655, le rectangle considéré sera alors $[-25, 75] \times [0.600, 0.655]$. Chaque valeur de la VCF est donnée avec 5 chiffres. On a donc une précision dix fois meilleure que dans la partie I. Le nombre total de données est 2412, avec un pas de 0.5 degré en température et 0.005 en densité.

Le traitement de cette partie est relativement plus facile que l'autre, comme le montrent les résultats de différents tests de la méthode B-splines, et des méthodes IP et Chevallier. Le nombre de valeurs à stocker reste faible.

Dans les exemples présentés la valeur de N_x (nombre de splines suivant la température) est toujours prise égale à 1 (différents essais ont montré qu'on ne gagne pas beaucoup en augmentant ce nombre).

2.4.1. Méthode IP:

Pour le traitement de cette partie, la méthode IP, approche les valeurs de la table par la fonction :

$$VCF(D_{15}, T) = 1 + (p_2 - p_1/D_{15})(T - 15) + (p_4 - p_3/D_{15})(T - 15)^2,$$

où p_1, p_2, p_3 et p_4 , sont des constantes qui dépendent de l'intervalle où se trouve D_{15} .

L'approximation utilise une subdivision de 3 intervalles en densité D_{15} et un seul intervalle en température. Donc, au total, on a 12 constantes à stocker.

Résultats:

Le nombre d'erreurs total est 210.

Le nombre d'erreurs comprises entre 0.5 et 0.6 est 141 représentant 67.1 % de l'ensemble des erreurs.

Le nombre d'erreurs supérieures ou égales à 1 est 3.

L'erreur maximale est 1.397.

L'erreur quadratique est 0.33.

L'erreur moyenne est 0.019.

2.4.2. Méthode de Chevallier:

Elle consiste à approcher la fonction VCF dans le rectangle $[-25, 75] \times [0.600, 0.655]$ par la fonction

$$VCF = \exp(-\alpha \Delta T (1 + A \alpha \Delta T + B \alpha^2 \Delta T^2)),$$

où

$$\alpha = K_0 + \frac{K_1}{D_{15}} + \frac{K_2}{D_{15}^2} + \frac{K_3}{D_{15}^3},$$

α est appelé le coefficient d'expansion thermique à 15 degrés du produit. K_0, K_1, K_2, K_3 , et A, B sont des constantes (voir rapport Chevallier sept 88).

Le nombre de constantes utilisées pour la représentation de la table est 6.

Résultats:

Le nombre d'erreurs total est 351.

Le nombre d'erreurs comprises entre 0.5 et 0.6 est 173 représentant 49.3 % de l'ensemble des erreurs.

Le nombre d'erreurs supérieures à 1 est 2.

L'erreur maximale est 1.42.

L'erreur quadratique est 0.37.

L'erreur moyenne est $-5.9.10^{-2}$.

2.4.3. Méthode B-spline:

ExempleNombre de données, $ND =$

2412

$$N_x = 1,$$

$$N_y = 5,$$

$$d = 2,$$

$$\tau = 10^{-5}.$$

Nombre d'erreurs total, $NErr =$

165

Nombre d'erreurs comprises entre 0.5 et 0.6

125

ce qui représente parmi l'ensemble des erreurs

75.3 %

Nombre d'erreurs supérieures ou égales 1

1

Erreur maximale

1.109

Erreur quadratique moyenne

0.31

Moyenne des erreurs

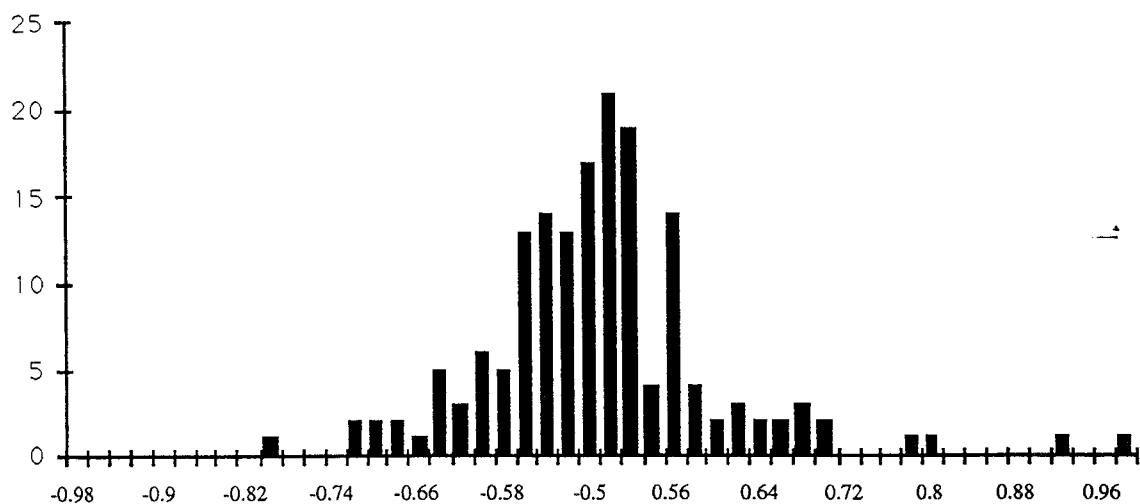
 10^{-12}

Nombre de coefficients pour le calcul de la spline

21

Taux de compression, $comp =$

0.87 %

FIGURE .4. *Histogramme des erreurs*

Dans la figure 4, on donne l'histogramme représentant la distribution des erreurs autour de -0.5 et 0.5 (sur le dernier digit).

2.4.4. Conclusions Partie II table 54:

La partie II est relativement plus facile à traiter comme le montrent les résultats des trois méthodes utilisées. On recommande les valeurs de $Nx = 1$ et $Ny = 5$ pour un degré $d = 2$ qui conduisent à un nombre de constantes égal à 21. La comparaison avec les autres méthodes s'établit comme suit:

	nombre d'erreurs	nombre d'erreurs entre 0.5 et 0.6	nombre d'erreurs > 1	$EQM - 0.29$
Méthode IP	210	141 (67.1 %)	3	0.04
Méthode de Chevallier	351	173 (49.3 %)	2	0.08
Méthode B-splines	165	125 (75.5 %)	1	0.02

2. 5. Partie I de la table 53

Dans cette partie de la table, la température varie dans l'intervalle $[-46,60]$, le pas étant 0.5 degré; la densité observée varie dans l'intervalle $[0.420,0.600]$, le pas étant 0.005. Les données sont fournies avec 3 chiffres. Le nombre total des données est de 4300.

2.5.1. Méthode IP

Le traitement de cette partie par la méthode IP n'est pas direct mais utilise la relation implicite suivante:

$$D_{15} = \frac{D_0}{VCF(D_{15}, T)},$$

où VCF est la représentation proposée par IP pour le traitement de la table 54.

Les valeurs de D_{15} qui sont inférieures à 0.500 ne sont pas prises en considération.

Résultats:

Le nombre de données traitées est 3520.

Le nombre d'erreurs total est 1319.

Le nombre d'erreurs comprises entre 0.5 et 0.6 est 320, représentant 24.3 % de l'ensemble des erreurs.

Le nombre d'erreurs supérieures à 1 est 332.

L'erreur maximale est -3.43.

L'erreur quadratique est 0.63.

L'erreur moyenne est -0.204.

2.5.2. Méthode de Chevallier

Le traitement de cette partie utilise la relation implicite suivante:

$$D_{15} = \frac{D_0}{VCF(D_{15}, T)} \cdot HYC,$$

où

$$HYC = 1 - 0.000023(T - 15) - 0.00000002(T - 15)^2,$$

et VCF est la représentation proposée par Chevallier pour le traitement de la table 54.

Les valeurs qui sont traitées par cette méthode sont celles qui sont comprises entre 0.500 et 0.600.

Résultats:

Le nombre de données traitées est 3241.

Le nombre d'erreurs total est 1196.

Le nombre d'erreurs comprises entre 0.5 et 0.6 est 290, représentant 24.2 % de l'ensemble des erreurs.

Le nombre d'erreurs supérieures à 1 est 236.

L'erreur maximale est 1.887.

L'erreur quadratique est 0.56.

L'erreur moyenne est -0.157.

Remarque 1:

Pour le traitement de cette partie, les données, correspondant à la densité 0.600, sont prises en compte dans l'approximation, mais elles ne sont pas considérées pour l'évaluation du nombre d'erreurs. La colonne $D_0 = 0.600$ appartient à la partie II.

2.5.3. Méthode B-spline:

ExempleNombre de données, $ND =$

4300

$$N_x = 12,$$

$$N_y = 30,$$

$$d = 3,$$

$$\tau = 10^{-8}.$$

Nombre d'erreurs total, $NErr =$

419

Nombre d'erreurs comprises entre 0.5 et 0.6

266

ce qui représente parmi l'ensemble des erreurs

63.5 %

Nombre d'erreurs supérieures ou égales 1

0

Erreur maximale

0.959

Erreur quadratique moyenne

0.32

Moyenne des erreurs

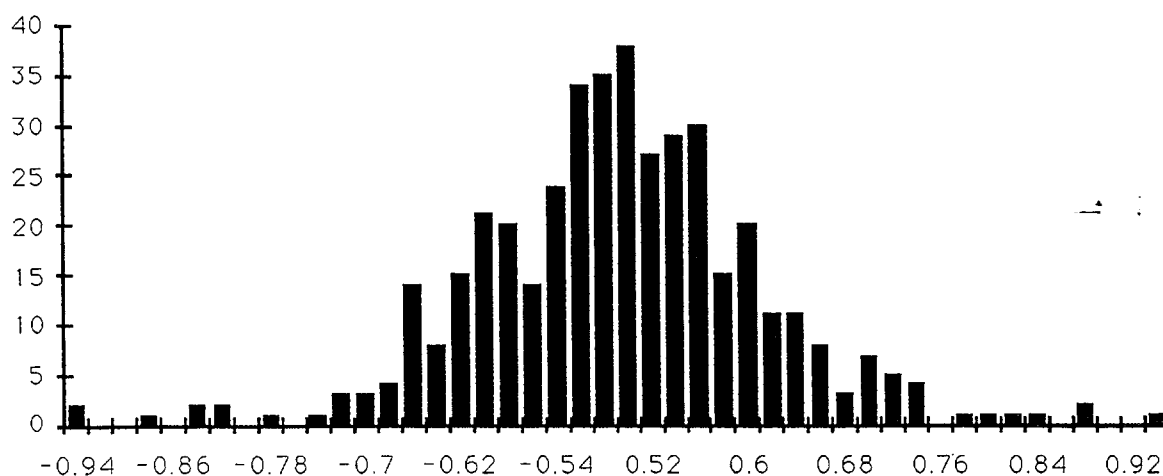
 10^{-12}

Nombre de valeurs pour le calcul de la spline

495

Taux de compression, $comp =$

11.5 %

FIGURE .5. *Histogramme des erreurs*

Dans la figure 5, on donne l'histogramme représentant la distribution des erreurs autour de -0.5 et 0.5 (sur le dernier digit).

2.5.4. Conclusions (Partie I table 53):

Pour cette partie, nous avons retenu les valeurs de $N_x = 12$, $N_y = 30$ et $d = 3$ qui conduisent à 495 constantes. La comparaison avec les autres méthodes s'établit comme suit:

	nbre d'erreurs	nbre d'err. entre 0.5 et 0.6	nbre d'err. > 1	EQM – 0.29	nbre de données traitées
Méth. IP	1319	320	332	0.34	3520
Méth. Chevallier	1196	290	236	0.27	3241
Méth. B-splines	419	259	0	0.04	4300

Remarques :

- Dans la méthode IP, l'écart est relativement grand par rapport aux deux autres méthodes. Ceci provient du fait que l'itération qu'elle définit ne converge pas toujours vers la bonne valeur.
- Pour les résultats relatifs à la méthode IP et Chevallier, les données dont les valeurs de D_{15} sont inférieures à 0.500 ne sont pas traitées.

2. 6. Partie II de la table 53:

Dans cette partie, la température varie dans l'intervalle $[-25, 50]$, le pas étant 0.5 degré; la densité observée varie dans l'intervalle $[0.600, 0.689]$, le pas étant 0.001. Les données sont fournies avec 4 chiffres significatifs.

Le nombre total de données est 9422. Cette partie est plus restreinte que la première mais plus dense.

2.6.1. Méthode IP

La méthode IP utilise une méthode itérative utilisant la formule

$$D_{15} = \frac{D_0 + p_2(T - 15) + p_4(T - 15)^2}{1 + p_1(T - 15) + p_3(T - 15)^2} * HYC,$$

où

$$HYC = 1 - 0.000023(T - 15) + 0.00000002(T - 15)^2$$

et

p_1, p_2, p_3, p_4 sont des constantes qui dépendent de l'intervalle où se trouve D_{15} .

On arrête l'itération quand la valeur courante de D_{15} tombe dans le même intervalle que celui où elle a été prise au cours de l'itération précédente.

Les valeurs de D_{15} qui sont supérieures à 0.654 ne sont pas prises en considération.

Résultats:

Le nombre de données traitées est 8013.

Le nombre d'erreurs total est 606.

Le nombre d'erreurs comprises entre 0.5 et 0.6 est 271, représentant 44.2 % de l'ensemble des erreurs.

Le nombre d'erreurs supérieures à 1 est 225.

L'erreur maximale est 30.012.

L'erreur quadratique est 1.39.

L'erreur moyenne est -0.046.

2.6.2. Méthode de Chevallier

Utilise la même relation implicite que pour la partie I, c'est à dire:

$$D_{15} = \frac{D_0}{VCF(D_{15}, t)} \cdot HYC,$$

où

$$HYC = 1 - 0.000023(T - 15) - 0.00000002(T - 15)^2.$$

Les valeurs qui sont traitées par cette méthode sont celles qui sont comprises entre 0.600 et 0.654.

Résultats:

Le nombre de données traitées est 6788

Le nombre d'erreurs total est 1525.

Le nombre d'erreurs comprises entre 0.5 et 0.6 est 480, représentant 30.8 % de l'ensemble des erreurs.

Le nombre d'erreurs supérieures à 1 est 181.

L'erreur maximale est 1.5.

L'erreur quadratique est 0.44.

L'erreur moyenne est 0.22.

2.6.3. Méthode B-spline:

ExempleNombre de données, $ND =$

9422

$$N_x = 10,$$

$$N_y = 15,$$

$$d = 3,$$

$$\tau = 10^{-6}.$$

Nombre d'erreurs total $NErr =$

258

Nombre d'erreurs comprises entre 0.5 et 0.6

237

ce qui représente parmi l'ensemble des erreurs

91.9 %

Nombre d'erreurs supérieures ou égales 1

0

Erreur maximale

0.75

Erreur quadratique moyenne

0.29

Moyenne des erreurs

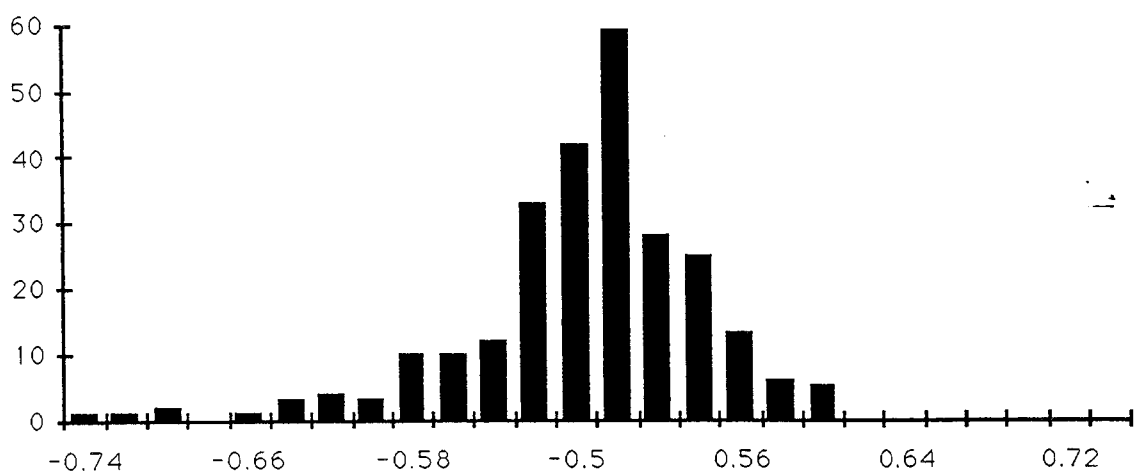
 -10^{-12}

Nombre de coefficients pour le calcul de la spline

234

Taux de compression, $comp =$

2.48 %

FIGURE .6. *Histogramme des erreurs*

Dans la figure 6, on donne l'histogramme représentant la distribution des erreurs autour de -0.5 et 0.5 (sur le dernier digit).

2.6.4. Conclusions (Partie II table 53):

Malgré le grand nombre de données que renferme cette partie, le traitement se fait plus facilement que pour la partie I comme le montrent les exemples traités. Nous recommandons les valeurs $N_x = 10, N_y = 15, \tau = 10^{-6}$ et $d = 3$ qui conduisent à un nombre de constantes égal à 234 . La comparaison avec les méthodes IP et Chevallier s'établit comme suit:

	nbre d'err.	nbre d'err. entre 0.5 et 0.6	nbre d'err. > 1	$EQM -$ 0.29	nbre de données traitées
Méth. IP	606	271	225	1.1	8013
Méth. Chevallier	1548	480	181	0.15	6788
Méth. B-splines	258	237	0	0	9422

On notera que pour la méthode IP, le nombre 8013 correspond au nombre de données que la méthode affirme pouvoir traiter. De ce fait, dans certains cas, l'itération conduit à sortir de la table 54 et produit des erreurs importantes (jusqu'à 30 unités du dernier chiffre).

3. Traitement global

Dans cette partie, on effectue un traitement global par B-splines de toutes les données de la table 53. Le domaine considéré est le rectangle $[0.420, 0.689] \times [-46, 60]$. Ce domaine est caractérisé par la présence de trous plus ou moins importants.

Nombre de données, $ND =$ 13722

$$N_x = 20,$$

$$N_y = 40,$$

$$d = 3,$$

$$\tau = 10^{-9}.$$

Nombre d'erreurs total, $NErr =$ 677

Nombre d'erreurs comprises entre 0.5 et 0.6 463

ce qui représente parmi l'ensemble des erreurs 68.4 %

Nombre d'erreurs supérieures ou égales 1 15

Erreur maximale 1.36

Erreur quadratique moyenne 0.30

Moyenne des erreurs	-3.10^{-13}
Nombre de coefficients pour le calcul de la spline	989
Taux de compression, <i>comp</i> =	7.21 %

3. 1. Conclusions (traitement global des parties I et II de la table 53)

L'importance des trous dans le domaine quand il est considéré globalement, et les différences entre les parties I et II citées précédemment, rendent le traitement assez difficile. Dans les traitements séparés des deux parties on peut voir qu'on arrive à de meilleurs résultats avec moins de constantes.

Nous avons retenu les valeurs $N_x = 20, N_y = 40, d = 3$ et $\tau = 10^{-9}$ qui conduisent à 989 constantes; le nombre d'erreurs étant égal à 677 dont 463 sont comprises entre 0.5 et 0.6; le nombre d'erreurs supérieures à 1 est 15; l'erreur quadratique moyenne par rapport au cas idéal d'un simple arrondi est 0.01. Une comparaison avec les méthodes IP et Chevallier ne peut être faite étant donné que leur traitement n'est pas global.

Sur les exemples traités, on peut voir le comportement du nombre d'erreurs par rapport au nombre de coefficients. Ainsi, dans le cas où $N_x = 5, N_y = 30$ et $D = 3$, donc pour 264 constantes, on arrive à un nombre d'erreurs de l'ordre de 1374 (environ le dixième du nombre de données); en prenant l'exemple de $N_x = 15, N_y = 25$ et $D = 3$, qui conduit à un nombre de constantes égal à 504, le nombre d'erreurs devient 1183, soit une diminution par rapport au cas précédent de 191 erreurs alors qu'on a ajouté 240 constantes supplémentaires. Pour $N_x = 20, N_y = 40, D = 3$, on arrive à 677 erreurs avec un nombre de constantes égal à 989; on a pratiquement multiplié le nombre de constantes par 4 pour avoir un nombre d'erreurs réduit de moitié.

4. Raccordement à $D_{15} = 0.600$ entre les parties I et II de la table 54

Pour étudier le raccordement des deux parties I et II, on procède par comparaison de deux histogrammes. Le premier est celui qui correspond au raccordement assuré par la table 54, ceci en regardant les valeurs de la VCF à $D_{15} = 0.600$ dans la partie I et la partie II. Cet histogramme représente les différences qu'il y a entre les valeurs de la VCF dans la partie II (données avec 5 chiffres) et celles de la partie I (données avec 4 chiffres). Le deuxième est celui qui correspond au raccordement assuré par la méthode des B-splines en prenant la différence des valeurs de la spline d'approximation de la partie II de la table arrondies à 5 chiffres et celle de la spline d'approximation de la partie I arrondies à 4 chiffres.

Théoriquement, les deux histogrammes devraient être équilibrés puisqu'ils devraient correspondre à l'histogramme représentant les différences entre une fonction et son arrondi.

Histogramme correspondant à la liaison à 0.600 entre la partie I et II de la table 54. En Gris les données relatives à la table, en noir celles relatives à méthode Bsplines

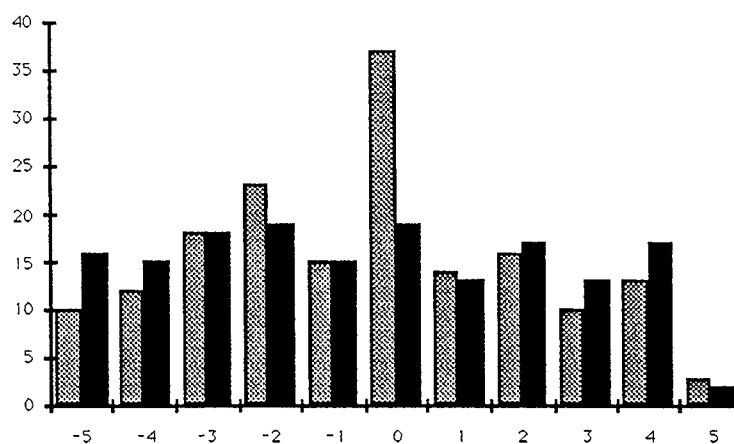


FIGURE .7.

On observe ici une anomalie dans les données. Dans l'histogramme gris la valeur en zéro est double de ce qu'elle devait être. Concrètement, cela signifie que dans la partie II (5 chiffres) en 0.600, on trouve 2 fois plus que la normale de nombres se terminant par un 0.

On peut remarquer que l'histogramme en noir, correspondant à la méthode B-spline est plus équilibré. Le lissage par les B-splines régularisées a amélioré le raccordement entre les parties I et II.

5. Relation entre les tables 53 et 54

La table 53 donne les valeurs de la densité à 15 degrés en fonction de la densité observée D_o et de la température T :

$$D_{15} = F(D_o, T).$$

La table 54 donne les valeurs de le VCF (Volume correction factor) en fonction de la densité à 15 degrés D_{15} et de la température T :

$$VCF = G(D_{15}, T).$$

Il existe entre les deux tables une relation mathématique. Celle-ci permet théoriquement de déduire la table 53 de la table 54. Pour une densité D_o et une température T donnée, on peut effectuer l'itération suivante:

$$D_{15}^{(k+1)} = \frac{D_o}{G(D_{15}^{(k)}, T)}.$$

Si la suite $D_{15}^{(k)}$ converge, la valeur limite $D_{15}^{(\infty)}$ vérifie donc:

$$D_{15}^{(\infty)} = \frac{D_o}{G(D_{15}^{(\infty)}, T)}$$

On obtient alors la valeur D_{15} en appliquant à $D_{15}^{(\infty)}$ le terme correctif HYC (Hydrometer correction):

$$D_{15} = D_{15}^{(\infty)} * HYC,$$

avec:

$$HYC = 1 - 0.000023(T - 15) - 0.00000002(T - 15)^2.$$

Cette méthode a été appliquée par IP et par Chevallier (à la différence pour ce dernier qu'il n'applique pas de la même façon le terme correctif HYC). L'inconvénient majeur de ce procédé réside dans le fait qu'il ne permet de calculer D_{15} en fonction de D_o et T que lorsque l'itération converge (sans sortir de la table 54). On ne peut calculer ainsi qu'une fraction de la table 53.

Le but de ce chapitre est d'examiner pour chacune des trois méthodes ce que donnerait l'application de l'itération précédente pour déduire la table 53 à partir de l'évaluation que chacune des trois méthodes donne pour la table 54. Nous avons estimé qu'il était préférable de faire cette comparaison séparément pour les parties I et II (du fait que les précisions sont différentes pour les deux parties). Par exemple pour la partie I, nous n'avons effectué la comparaison que pour les couples (D_o, T) pour lesquels l'itération converge sans aller dans la partie II. Notons que pour Chevallier cette restriction a été respectée: il ne calcule la table 53 que pour des

couples (D_o, T) qui permettent à l'itération de rester à l'intérieur d'une même partie.

Partie I

Les valeurs de D_{15} qui sont prises en compte dans cette partie, sont celles qui sont supérieures à 0.500.

Les résultats des différentes méthodes utilisées dans cette partie s'établissent comme suit:

	nombre d'erreurs	Nbre d'erreurs entre 0.5 et 0.6	Nbre d'erreurs > 1	EQM – 0.29	Erreur max.
Méth. IP ⁽¹⁾	1103	299	196	0.34	-2.12
Méth. de Chevallier ⁽²⁾	1081	299	169	0.31	-1.85
Méth. B-splines ⁽³⁾	956	277	126	0.32	-1.72

(1) Nombre de données traitées par IP est 3236.

(2) Nombre de données traitées par Chevallier est 3228.

(3) Nombre de données traitées par B-splines est 3043.

Partie II

Les valeurs de D_{15} qui sont prises en compte dans cette partie, sont celles qui sont supérieures à 0.600 et inférieures à 0.655.

Les résultats des différentes méthodes utilisées dans cette partie s'établissent comme suit:

	nombre d'err.	Nbre d'err. entre 0.5 et 0.6	Nbre d'err. > 1	EQM – 0.29	Erreur max.
Méth. IP ⁽¹⁾	233	196 (84.1 %)	0	0	0.70
Méth. de Chevallier ⁽²⁾	335	247 (73.7 %)	1	0.01	-1.022
Méth. B-splines ⁽³⁾	248	227 (91.5 %)	0	0	0.858

(1) Nombre de données traitées par IP est 6685.

(2) Nombre de données traitées par Chevallier est 6685.

(3) Nombre de données traitées par B-splines est 6686.

6. Conclusions

Par rapport aux méthodes existantes (IP et Chevallier) la méthode B-splines a les trois avantages suivants:

- Elle conduit à un nombre d'erreurs d'arrondis nettement inférieur.
- Il n'y a pratiquement pas d'erreurs qui sont supérieures à une unité sur le dernier chiffre.
- Elle permet de représenter la totalité des deux tables.

En revanche elle a l'inconvénient d'utiliser un plus grand nombre de coefficients.

En résumant les résultats donnés en 2.3.4 et 2.4.4, on obtient pour l'ensemble de la table 54:

	nombre d'erreurs	Nbre d'erreurs entre 0.5 et 0.6	Nbre d'erreurs > 1
Méth. IP	1761	486	356
Méth. de Chevallier	1748	666	327
Méth. B-splines	669	414	1

Le nombre total de coefficients utilisés par la méthode B-spline pour la table 54 est de 181.

De même, en résumant les résultats donnés en 2.5.4 et 2.6.4, on obtient pour l'ensemble de la table 53:

	nombre d'erreurs	Nbre d'erreurs entre 0.5 et 0.6	Nbre d'erreurs > 1
Méth. IP	1925	591	557
Méth. de Chevallier	2648	770	417
Méth. B-splines	677	503	0

Dans ce tableau le nombre d'erreurs pour les deux premières méthodes correspond à:

- 11532 données traitées pour la méthode IP,
 - 10029 données traitées pour la méthode de Chevallier
- alors que pour la méthode B-splines il correspond à la totalité des 13722 données.

En tenant compte des "trous" dans la table 53, le nombre total de coefficients réellement utilisés par la méthodes B-splines est de 353 pour la partie I et 195 pour la partie II, soit 548 pour toute la table.

Bibliographie

- [Ale33] A.D. Alexandrov
A theorem on convex polyhedra
Trudy Mat. Inst. Sect. Math. 4, 87 (Russian) (1933).
- [ABC93] A. Antoniadis, J. Berruyer, R. Carmona
Regression Non Linéaire
Economica, (1992)
- [Bai92] R. Baier, F. Lempio
Computing Aumann's Integral
in Proceedings of the IIASA Workshop on Modelling Techniques for Uncertain Systems, Sopron (Hungary), July 6-10, 1992.
- [Ber78] M. Berger
Géométrie 3/ convexes et polytopes, polyèdres réguliers, aires et volumes
Cecic/Fenand Nathan, (1978).
- [Bet83] U. Betke, P. Mc Mullen
Estimating the sizes of convex bodies from projections
Jour. London. Math. Soc. (2) 27 (1983) p. 525-538.
- [Boo78] C. de Boor
A practical guide to splines
Applied Mathematical Sciences, Vol. 27, Springer Verlag, New York (1978).
- [Bol69] E.D. Bolker
A class of Convex Bodies
Trans. Amer. Math. Soc. 145 (1969) p. 323-345.
- [BLM 89] J. Bourgain, J. Lindenstrauss, V. Milman
Approximation of zonoids by zonotopes
Acta Math., 162 (1989), p. 73-145.
- [Ceu94] G. Celeux
Cours de DEA, Université J. Fourier, Grenoble.
- [Chi63] B.L. Chilton, H.S.M. Coxeter
Polar Zonohedra
1963.
- [Cho68a] G. Choquet
Mesures coniques, affines et cylindriques : structures et opérations
C.R. Acad. Sc. Paris 266 (1968) p. 567-569.
- [Cho68b] G. Choquet
Mesures coniques, affines invariantes par isométrie. Zonoformes, zonoèdres et fonctions de type négatif
C.R. Acad. Sc. Paris 266 (1968) p. 619-621.

- [Cia82] P.G. Ciarlet
Introduction à l'analyse numérique matricielle et à l'optimisation
Masson, Paris, (1982).
- [Cox62] H.S.M Coxeter
The classification of zonohedra by means of projective diagrams
J. Math. Pures Appl. (9) 41 (1962) p. 137-156.
- [Cox63] H.S.M Coxeter
Regular Polytopes
2nd Edition, MacMillan, New York (1963).
- [Dao92] O. Daoudi
Représentations des Tables ASTM-IP 53/54, par B-splines Régularisées
Rapport ELF-Aquitaine. n° 92/802 Juin 92.
- [Dao94] O. Daoudi, B. Lacolle, N. Szafran, P. Valentin
Zonoidal Surfaces
, Curve and Surfaces for Geometric Design. P.J. Laurent, A. Le Méhauté, L.L. Shumakker (eds), AK Peters, Wellesly, MA (1994).
- [Ede87] H. Edelsbrunner
Algorithms in Combinatorial Geometry
EATCS Monographs on Theoretical Computer Science, Vol. 10, Springer Verlag (1987).
- [Fed53] E.S. Fedorov
Nachala Ucheniya o Figurah
Leningrad (1953).
- [Fia83] A.V. Fiacco
Introduction to Sensitivity and Stability Analysis in Nonlinear Programming
Vol 16, Academic Press (1983)
- [Fle87] R. Fletcher
Practical Methods of Optimization
2nd ed. Chchester; Bristone; New York; John Welley and Sons (1987)
- [Gir86] D. Girard
Convexes Résidu et Gestion de Mélanges sur une plateforme
Rapport de recherche ELF-Aquitaine.
- [Gir89] D. Girard, P. Valentin
Zonotopes and Mixtures Management
International Series of Numerical Mathematics. Vol. 87, Birkhäuser Verlag (1989).
- [Goo77] P.R. Goodey
Centrally Symetric Convex Sets and Mixed Volumes
Mathematika 24 (1987) p. 193-198.

- [Gor85] Y. Gordon
Some inequalities for Gaussian processes and applications.
Israel J. Math., 50 (1985), 265-289.
- [Grü67] B. Grünbaum
Convex Polytopes
Pure and Appl. Math., vol. 16, Interscience, New York (1967).
- [Hil81] W. Hildenbrand
Short-Run Productions based on Microdata
Econometrica, vol 49 n°5, p.1095-1125.
- [Haj88] B. Hajek
Cooling Schedules for Optimal Annealing.
Mathematics for Operat. Research, 13, (1988) 311-329.
- [Hör54] P.L. Hörmander
Sur la fonction d'appui des ensembles convexes dans un espace localement convexe
Ark. Mat. 3, no 12 (1954), p. 181-186.
- [Kar68] S. Karlin
Total Positivity
Stanford University Press (1968).
- [Kel58] L.M. Kelley, W.O.J. Moser
On the number of ordinary lines determined by n points
Canada J. Math. t. 10 (1958) p. 210-219.
- [Lac91] B. Lacolle, N. Szafran, P. Valentin
Les mélanges binaires : modélisation géométrique et algorithmes
Rapport de recherche IMAG RR841-M février (1991).
- [Lac89] B. Lacolle, P. Valentin
Etude de la faisabilité de plusieurs mélanges
, Rapport de recherche IMAG, N° 776-M, Mai (1989)
- [Lau72] P.J. Laurent
Approximation et Optimisation
Hermann, Paris, (1972).
- [Lau87] P.J. Laurent
Courbes ouvertes ou fermées par B-splines Régularisées
Rapport de Recherche n° 652, Mars, Université de Grenoble.
- [LeM88] V. Le Maire
Optimisation d'un arbre de décision par un algorithme de recuit simulé
Rapport de recherche 854, INRIA-Sophia antipolis Juin 88.
- [Lin75] N.F. Linquist
Approximation of convex bodies by sum of line segments
Portugalia Math. 34 (1975), p.233-240

- [McM68] P. Mc Mullen, G.C. Shephard
Diagrams for centrally symmetric polytopes
Mathematika 15 (1968) p. 139-148.
- [McM71a] P. Mc Mullen
On Zonotopes
Trans. Amer. Math. Soc. Vol. 159 (Sept 1971) p. 91-109.
- [McM71b] P. Mc Mullen, G.C. Shephard
Convex Polytopes and the Upper Bound Conjecture
London Math. Soc. Lecture Notes, vol. 3 Cambridge University Press (1971).
- [McM73] P. Mc Mullen
Representation of polytopes and polyhedral sets
Geometriae Dedicata 2 (1973) p. 83-99.
- [McM74] P. Mc Mullen, R. Schneider, G.C. Shephard
Monotopic polytopes and their intersection properties
Geometriae Dedicata 3 (1974) p. 99-129.
- [McM75] P. Mc Mullen
Space Tiling Zonotopes
Mathematika 22 (1975) p. 202-211.
- [Min83] M. Minoux
Programmation Mathématiques
(tome 1), Dunod, Paris, (1983).
- [Ode90] N. Odeh
Modélisation Mathématique des Propriétés de Mélange : Bsplines et Optimisation avec conditions de forme
Thèse Université J. Fourier, Grenoble (1990).
- [Pre85] F.P. Preparata, M.I. Shamos
Computational Geometry- An Introduction
Texts & Monographs in Computer Science, Springer Verlag (1985).
- [Rock70] R.T. Rockafellar
Convex Analysis
Princeton University Press (1970).
- [Ros60] J.B. Rosen
The Gradient Projection Method for Non Linear Programming
Part 1 : Linear Constraints, Journal S.I.A.M., 8, p. 181-217.
- [Ros61] J.B. Rosen
The Gradient Projection Method for Non Linear Programming
Part 2 : Non Linear Constraints, Journal S.I.A.M., 9, p. 514-532.
- [Rud75] W. Rudin
Analyse réelle et complexe
Masson Paris (1975).

- [Sch69] R. Schneider
Characterization of certain polytopes by intersection properties of their translates
Mathematika 16 (1969) p. 276-282.
- [Sch83] R. Schneider, W. Weil
Zonoids and related topics in convexity and its applications
Birkhäuser Verlag Basel (1983).
- [She67] G.C. Shephard
Polytopes with centrally symmetric faces
Can. J. Math 19 (1967) p.1206-1213.
- [She71a] G.C. Shephard
Decomposable convex polyhedra
J. London Soc. (2), 4 (1971) p.165-175.
- [She71b] G.C. Shephard
Diagrams for positive bases
J. London Soc. (2), 4 (1971) p. 165-175.
- [She71c] G.C. Shephard
Polyhedral diagrams for sections of non negative orthant
Mathematika 18 (1971) p.255-263.
- [She74] G.C. Shephard
Combinatorial properties of associated zonotopes
Can. Jour. Math., 26-2 (1974), p.302-321.
- [She74b] G.C. Shephard
Spaces-Filling Zonotopes
Mathematika 21 (1974) p.261-269.
- [Sla86] K.H. Slaoui
Application des techniques à la gestion des mélanges : histosplines et optimisation
Thèse I.N.P.G., Grenoble (1986).
- [Sza91] N. Szafran
Zonoèdres : de la géométrie algorithmique à la théorie de la séparation
Thèse Université J. Fourier, Grenoble (1991).
- [Val86] P. Valentin
Convex set of mixtures : a geometric approach to separation engineering
Rapport Centre de Recherches ELF-Solaize (1986).
- [Val87] P. Valentin
A geometric foundation to separation engineering
ICIAM 87, Paris, 29 Juin - 3 Juillet (1987).

- [Val90] P. Valentin
Le principe de non croissance de la séparation : application aux modèles d'équations aux dérivées partielles de transport et de diffusion de la chaleur et de la matière
Rapport Interne Centre de recherches ELF-Solaize. (Août 90).
- [Val91] P. Valentin
Is chromatography a separation process? The zonoid answer
Journal of Chromatography, Sciences Publisher BV (Sept 1991).
- [Val92a] P. Valentin
Theory of zonoids I: A mathematical summary
In Advanced in theory of Chromatography (18-30 août 91).
- [Val9b] P. Valentin
Theory of zonoids II: Application to the chromatography processes
In Advanced in theory of Chromatography (18-30 août 91).
- [Vale64] F.A. Valentine
Convex Sets
McGraw-Hill (1964).
- [Van71] B. Van Cutsem
Éléments aléatoires à valeurs convexes compactes
Thèse Université scientifique et médical de Grenoble. Juin 1971.
- [Wei76] W. Weil
Centrally Symmetric Convex Bodies and Distributions
Israel Jour. Math., 24 (1976), p. 352-367.